Final Report of Language Background as a Variable in NAEP Mathematics Performance

CSE Technical Report 429

Jamal Abedi
CRESST/University of California, Los Angeles
Carol Lord
CRESST/California State University, Long Beach
Joseph R. Plummer
CRESST/University of California, Los Angeles

August 1997

Center for the Study of Evaluation
National Center for Research on Evaluation, Standards,
and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90024-1522
(310) 206-1532

The work reported herein was supported in part under the National Center for Education Statistics Contract No. RS90159001 as administered by the U.S. Department of Education, Office of Educational Research and Improvement.

ACKNOWLEDGMENTS

We would like to thank all of the researchers—faculty, staff, and students—who informed this study. This project would have been impossible to accomplish without the efforts of a large group of people. The authors owe all of them a debt of gratitude. While the members of the group often disagreed on how best to approach the numerous problems, this was because of their dedication and enthusiasm for the project, and because of their genuine caring for those groups of students who are negatively impacted by their language background.

Frances Butler	Joan Herman
Patricia Snyder	Sara Cushing Weigle
Larry Fish	Leigh Burstein
Josie Bain	Alfredo Artiles
Derek Mitchell	Katharine Fry
Missy Krasner	Victoria Taylor

Los Angeles Unified School District and all the test administrators and teachers who helped us

CONTENTS

FINAL REPORT OF LANGUAGE BACKGROUND AS A VARIABLE IN NAEP MATHEMATICS PERFORMANCE

EXECUTIVE SUMMARY

Jamal Abedi, CRESST/University of California, Los Angeles
Carol Lord, CRESST/California State University, Long Beach
Joseph R. Plummer, CRESST/University of California, Los Angeles

Introduction

In an attempt to respond to the growing national concern about students' language background and its effects on performance, the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at UCLA has undertaken a study that examines the linguistic features of National Assessment of Educational Progress (NAEP) test items. The goal of the study is to begin to identify linguistic features in NAEP mathematics items that may affect the performance of students with language backgrounds other than English. Because little is known about how language affects content area assessment, the study is considered to be exploratory, but is also viewed as a first step in providing a language-sensitive framework for constructing and reviewing content area assessments.

The Studies

We first reviewed the literature in the areas of mathematics performance of language minority students, the role of language in solving math word problems, and measurement of linguistic complexity in general. The literature review guided us in deciding which language variables could be productively applied to the NAEP data. The research itself had two separate phases: (1) analyses of extant data, and (2) field research.

Phase 1

In Phase 1 of the study, we examined the NAEP data from the 1990 and 1992 main assessments. Items from the 8th-grade NAEP math tests and questionnaire items were analyzed using a linguistic categorization scheme. A multiple discriminant analysis was applied to composite scores to examine the effects of language background variables. In this multiple discriminant analysis, language background variables were used as grouping variables and composite test scores were used as discriminating variables. The results clearly revealed lower math proficiency scores for the subjects who predominantly spoke a language other than English in the home. This relationship was more evident for longer items.

Next, the effect of linguistic complexity on students' performance on NAEP math items was analyzed by creating item parcels based on linguistic complexity, using pragmatic criteria including difficulty of vocabulary, abstract or culture-specific content, and number of complex structures in a sentence. A repeated measures design was applied to the parcel scores. The results of the analyses conducted on the language background variables showed a highly significant difference between the scores of the two parcels. Students who spoke more of a language other than English in the home performed significantly lower than students who spoke only English in the home, and the difference was greater for the linguistically complex items, $F_{1,1170} = 56.42$, $p < .01$.

Lastly, we examined the proportions of omitted or not-reached items by students' language background. Groups were formed based upon whether the student reported speaking a language other than English in the home *always*, *sometimes*, or *never*. The groups were then compared on omitted/not-reached items. In nearly all cases, the students who always spoke a language other than English in the home had much higher percentages of omitted/not-reached items than the students who spoke only English in the home.

Phase 2

In Phase 2 of the study, we examined the role of linguistic complexity in students' performance on NAEP math items. Based on the literature and expert knowledge, we identified linguistically complex NAEP items. The set of linguistic features employed for this phase of the study was limited to features actually occurring in the small corpus of released NAEP math items available to us. The features chosen included familiarity/frequency of non-math vocabulary, length of nominals (noun phrases), voice of verb phrase, conditional clauses, question phrases, and abstract or impersonal presentations. We then prepared modified versions of these linguistically complex items so that the revised items contained simpler language but retained their original math content. The linguistically complex items and their revised counterparts were administered to a group of mostly 8th-grade students in the greater Los Angeles area to find out, in fact, whether linguistic complexity had any impact on students' math performance. The study's item pool was limited to a subset of the 1992 released math items.

Student Perceptions Study

Three separate studies were conducted in Phase 2 of the project and are reported separately. The first study, which will be referred to as the Student Perceptions Study, consisted of interviews with a group of 38 8th-grade students in the greater Los Angeles area, including native and non-native speakers of English with a range of math skill levels. The purpose of the interviews was to investigate the hypothesis that linguistically simplified items are, in fact, perceived as easier to understand by students. The students were presented the original (linguistically complex) math items and their revised (less linguistically complex) counterpart items in a structured interview format. Subjects consistently reported a strong preference for the revised items over the original items. Student preference for the revised items seemed to support the notion that the math items could be linguistically simplified in meaningful ways for the test taker. The interview results supported our plan to test a larger group of students to determine whether the observed differences in student responses to the language of the math items would translate into actual differences on math test scores.

Accuracy Study

The second study in Phase 2 will be referred to as the Accuracy Study. In this study, 39 8th-grade classes (1,031 students) were selected, with oversampling of limited English proficient (LEP) students. Released items from the 1992 main assessment were then re-examined for linguistic complexity based on the information obtained from the Student Perceptions Study. From these released items, 20 were identified as linguistically complex and were then modified to reduce linguistic complexity. The two sets of items (20 original and 20 revised) were placed into two booklets (Form A and Form B) along with 5 linguistically noncomplex items. In addition to the 25 math items, each booklet contained a 12-item language background questionnaire that was specifically designed for the substudy. Also, information on students' math background, ESL program participation, and socioeconomic status (SES) (as measured by participation in a free lunch program) was collected from schools.

In the data from the Accuracy Study, students' math performances on the original and revised items were compared. In general, the results of this study were consistent with the literature and indicated that (a) student background in math (as indicated by the level of math class) had a significant impact on student math scores in this study; (b) students in ESL programs had lower scores in math than non-ESL students; (c) males and females performed at about the same level; and (d) there were some differences in students' math performance with respect to ethnicity.

No analyses performed on revised versus original items yielded statistically significant results, except for those linked to math class level. However, certain trends were observed. As these trends suggested interesting possibilities, we investigated them in detail. We computed percent of improvement of students' math performance as a result of the revision of math items. For each level of math class, percentage of improvement was computed by subtracting the mean of original item scores from the mean of the revised item scores for the same set of items and then dividing the difference between the two means by the mean of original item scores. The revision of items had differential impact on students' math performances. Students in low- and average-level math classes exhibited the greatest improvement. The trend decreased over the intermediate to high categories, and for the highest level math classes (high math, honors, and algebra) there was no improvement. Students in different levels of math classes benefited differently from the revisions.

Because of the initially mixed results from the Accuracy Study, it was decided to perform analyses using HLM procedure. We created two models. In Model 1, we used the composite scores of the 10 original items in Booklet A as the outcome variable; students' membership in native/not-native English speaker groups and students' participation in free lunch program were used as subject-level data; and type of math class and an aggregate of free lunch program were used as level-2 variables in our HLM model. For Model 2, we used the same variables as level-1 and level-2 variables with 10 revised items in Booklet B (sister items of the 10 original items in Booklet A). A comparison between the two models revealed changes/improvements due to revision of items.

Speed Study

Based on results from the Accuracy Study, we examined the effect of linguistic modifications on the time a student required to answer/complete the math test items. Two more booklets were developed for this third study. The 20 original items were placed in Booklet A and the 20 revised items were placed in Booklet B. The five noncomplex items were eliminated from these booklets. The same language background questionnaire that was included in the booklets for the Accuracy Study was included in these booklets. One hundred forty-three 8th-grade students in the greater Los Angeles area were selected (mostly ESL math students) because it seemed that those students would benefit more from linguistically simplified versions of items. However, some students from high-level math classes and algebra classes were also included. Of the 143 students who participated in this study, 76 students answered the original items (Booklet A) and 67 answered the revised items (Booklet B). Students were given ten minutes to answer the 20 math questions. (In contrast, in the Accuracy Study, the majority of students were given enough time to answer all 25 questions.)

Native speakers (M = 4.76, SD = 2.75) performed slightly higher than non-native speakers (M = 3.65, SD = 2.25) on the speed test; but the difference was not significant. However, there were large differences between performances of students in different ESL, math class, and school lunch program categories. We could not apply analysis of variance in many cases because of extremely unproportional cell frequencies. For those analyses with appropriate frequencies, students in different math classes performed differently. For the Low math class category, the mean score was 3.68 (*SD* = 2.48) and for the High math class category, the mean was 5.18 (*SD* = 2.56). Analyses of variance revealed no significant differences between the subgroups of type of math class, $F_{2,64} = 1.76$, $p = .18$. School lunch program participation also seemed to have some impact on students' performance on the revised items. A range of differing degrees of participation in such programs was reported. The greatest degree of involvement was labeled AFDC, and no involvement in such programs was labeled No Lunch Code. For categories on this variable, means ranged from 3.14 (*SD* = 1.96) for No Lunch Code to 5.75 (*SD* = .500) for AFDC. However, ANOVA results yielded no significant results in this case, $F_{2,59} = 1.03$, $p = .36$.

Results

The analyses of NAEP data indicated some effects of students' language backgrounds on their math performance in junior high school. When items were categorized by their length, students who spoke a language other than English at home performed significantly lower than students who always spoke English at home; the difference was more pronounced on long items. Analysis also showed that the rates of omitted/not-reached math items for non-native English speakers were higher than those for native speakers. These results clearly indicate that confounding of language and performance occurs on NAEP math items.

Original and linguistically simplified items were administered in the Accuracy Study and the Speed Study. No statistically significant results were found overall, but students in low and average math classes scored higher on the simplified versions, consistent with similar findings in previous studies.

A number of problems emerged during the study, including limited access to the NAEP item pool, an unequal distribution of items across the NAEP content area subscales, and a lack of reliable measures of English proficiency. It was also observed that classes that were supposedly linguistically homogeneous were not necessarily so; although NAEP policy is to avoid testing ESL students, NAEP administrations are in fact testing students whose ability in English may be weak.

In addition to analyzing and discussing the language background questionnaire items independently, we also used these items along with the background data gathered from schools in analyzing students' math performances.

The results of analyses on the language background questions were consistent across the two field studies. Following is a summary of some of the findings from the language background questions:

1. Non-native English speakers tend to use their native language more with their parents and grandparents than with their siblings and friends.

2. Beginning ESL students showed more signs of concern in the area of understanding, speaking, reading and writing English.

3. All students reported that they have more problems understanding teachers' explanations, textbooks, and the texts of tests in the area of math than in the areas of science or social studies.

4. Native English speakers self-reported a higher level of proficiency in English than non-native speakers.

5. Males and females reported about the same level of proficiency in English and the "other language."

6. The most apparent differences between groups of students was across the categories of ESL class placement codes; differences were found on their self-reported level of English proficiency (understanding, speaking, reading, and writing) and on their understanding their teacher's explanation, textbook, and text of their exams.

Beginning ESL students in most cases reported a considerably lower level of English proficiency. However, the number of students in this category was so small in many instances that no valid interpretation was possible.

The most salient results of our analyses were significant differences in students' performances across categories of type of math class. When variability due to the type of math class was controlled, there was very little variability left to warrant further attention.

For the speed section of the study, there were higher rates of response on the revised items. These improvements were more evident with the language minority students. Unfortunately, the small number of students in this part of study did not allow us to do any in-depth analyses.

Conclusion

The results of our analyses on the original, revised, and total scores in general indicated that students in the ESL categories, particularly in the lower levels, show considerably lower math performance than other students. This is a great sign for concern and it requires special attention. There do not seem to be major differences between these ESL low-performance groups of students and other groups of students based on SES or other variables that could explain such differences. Therefore, one must conclude that language is a very important element in such cases. That is, language and performance are confounded. The exact nature of the confounding factors remains elusive.

The results of our analyses also suggested that revising math items to make them less linguistically complex helped some students, particularly those in low- and average-level math classes; previous studies have shown math and reading proficiency to be correlated, and it is likely that the reading and language skills of many of these students were also at the low or average level. In order to do math word problems, students must learn the special vocabulary and structures peculiar to the math word problem genre. In addition, general proficiency in language is necessary if the student is to learn from teachers and books in the mathematics classroom. General proficiency in language is also necessary for a true assessment of the student's knowledge in NAEP mathematics tests. Solving math word problems presents an additional challenge for the student whose language proficiency is limited, and the added cognitive load can impact individual performances negatively. Thus, the language of math items may disproportionately impact the scores of less language-proficient students, whether they are native speakers or non-native speakers. Other approaches emphasizing more representational rendition of content might facilitate performance of students with lower proficiencies in English.

Summary

To summarize, the study clearly shows that ESL students are at a significant disadvantage in mathematics content area assessments. We found that there was a small overall improvement in math scores on the revised versions of the NAEP math items, although such improvement was unimpressive. The lack of statistically significant improvement was due, we feel, to a number of limitations, including the small size of the item pool available. It remains prudent to continue searching for interactions among linguistic and socioeconomic background variables that will shed light upon the increasingly important issue of the role of language in content area assessment.

FINAL REPORT OF LANGUAGE BACKGROUND AS A VARIABLE IN NAEP MATHEMATICS PERFORMANCE

Jamal Abedi, CRESST/University of California, Los Angeles
Carol Lord, CRESST/California State University, Long Beach
Joseph R. Plummer, CRESST/University of California, Los Angeles

INTRODUCTION

Along with the recent development and advocacy of alternative assessments, there has been a growing realization of the importance of language in content areas that have not traditionally been considered to be confounded by linguistic variables. Because language skills affect a person's ability to process and understand facts and concepts and to demonstrate understanding to others, tests that attempt to measure content knowledge must be sensitive to the language ability of the test taker. The continuing increase in the number of language minority students in classrooms nationwide has forced the issue of language impact on assessments to the forefront. In an attempt to respond to the growing national concern about language background of students and its effects on performance, the National Center for Research on Evaluation, Standards, and Student Testing/Center for the Study of Evaluation (CRESST/CSE) undertook a study examining the linguistic features of the National Assessment of Educational Progress (NAEP) math test items and investigating the significance of language-related variables for NAEP's assessment in the content area of mathematics. Because NAEP is including direct samples of language from students in its move to open-ended and extended open-ended questions, there is increasing need for study of the effects of language background and the linguistic characteristics of items on students' performance on the NAEP test.

Recent reports from the National Council of Teachers of Mathematics (NCTM, 1989, p. 214) and the New Standards Project (see the 1992 report from the Learning Research and Development Center at the University of Pittsburgh and the National Center on Education) include discussion of the role of language in mathematics assessment. The latter study discusses the creation of a new set of national standards for education; a better set of national standards would lead to improved assessment, which in turn would be more likely to achieve the desired goal of "thinking" curricula. The assessment strategies being developed for these proposed standards, however, downplay the role of language differences across students from differing language backgrounds. The issue remains whether it is indeed advisable to consider mathematics a subject in which language plays a minimal role. It is this very issue, particularly as it relates to national assessment, that motivates the current study.

The study addressed the language of mathematics problems in current NAEP tests. Definitions of mathematical competency, as proposed by the NCTM, for example, have included the ability to solve problems in real world situations and to communicate about mathematics. Such definitions of competence have implications for mathematics instruction and assessment. Although these issues are important, they are beyond the scope of this study.

Issues and Goals

Do extant NAEP data confirm that language background significantly affects math item performance? If so, does revision of the language in NAEP math items also have an effect?

There is an important, additional concern when addressing equity in national assessments—how will such assessments affect limited English proficient (LEP) students (see Baker, 1991)? Students who are not included in national assessments because of lack of appropriate instruments will fail to benefit from the presumed desirable effects of assessment. Native language testing is not an adequate strategy for solving the LEP problem because of the linguistic diversity of the student population and the variety of curricula under which they are taught. Language and classroom culture are, therefore, areas that need increased attention if all students are to be provided opportunity to learn and if appropriate assessment is to be undertaken.

Thus, there is a need to analyze the linguistic dimensions and variables that can confound content assessment. Linguistic factors can affect NAEP test performance in a number of ways. Students may not understand an item because of the wording of the question, or they may not have the language capabilities to provide answers to open-ended questions, or both. We need to determine the amount of variation in test performance due to language background. The goal of this study, then, was to begin to identify linguistic features in NAEP mathematics items that may affect the performance of students from a variety of language backgrounds.

This study was conducted in two separate phases: (1) analyses of extant data and (2) field research. In Phase 1 of the study, we examined the NAEP data from the 1990 and 1992 main assessments. Items from the 8th-grade assessment were categorized according to the length of the item, a convenient index of linguistic complexity. Long items were defined as multiple-choice questions with extended stems (two lines or longer) and/or extended answer choices (longer than one line), and short items were defined as items with stems and choices of one line or less. Composite scores were computed for long and short items. Multiple discriminant analyses (DA) were applied to the composite scores to examine the effects of language background variables. In these analyses, language background variables were used as grouping variables, and composite test scores were used as discriminating variables. For each of the language background variables, one discriminant analysis was conducted. In addition, 8th-grade items were analyzed using a linguistic categorization scheme.

In Phase 2 of the study, we examined the impact of linguistic complexity of items on students' math performance. Based on the literature, expert knowledge, and the particular linguistic characteristics of released NAEP math items, we identified items with potentially problematic language. We then prepared modified versions of these linguistically complex items so that the revised items were less linguistically complex but retained their original math content. We administered the linguistically complex items and their revised counterparts to 8th-grade students in the greater Los Angeles area to find out whether, in fact, linguistic complexity had any impact on students' perceptions and math performance. For this phase of the study we were limited to the 1992 released math items.

Phase 2 included three studies: a Student Perceptions Study, which employed interviews to obtain students' reactions to the language of math items; an Accuracy Test Study, which compared scores on a multiple-choice test for simplified and nonsimplified items; and a Speed Test Study, in which student performance under severe time limitations was compared for simplified and nonsimplified items.

Review of the Literature

A number of studies investigating the relationship of language and mathematics performance were reviewed for this study and informed the research process. Researchers have examined the role of language in students' understanding of mathematics story problems, focusing on linguistic complexity, mathematics vocabulary, and translations between English prose and mathematical symbols—for English-proficient students as well as for LEP students. A number of studies have focused on the specialized vocabulary and syntactic constructions peculiar to the mathematics domain—the so-called "mathematics register." Other studies have considered the possibility that the general level of complexity of the language in story problems may play a role in students' comprehension.

Recent research has drawn attention to the importance of language in student performance on mathematics word problems. Nationally, children perform 10% to 30% worse on arithmetic word problems than on comparable problems presented in numeric format (Carpenter, Corbitt, Kepner, Linquist, & Reys, 1980). The discrepancy between performance on verbal and numeric format problems strongly suggests that factors other than mathematical skill contribute to success in solving word problems (Cummins, Kintsch, Reusser, & Weimer, 1988). Previous research has addressed three issues: (a) mathematics performance of language minority students, (b) the role of language in solving math word problems, and (c) measurement of linguistic complexity in general. We review this research here.

Mathematics Performance of Language Minority Students

Language minority students (including Native American and Hispanic students) score lower than White students on standardized tests of mathematics achievement in elementary school, as well as on the SAT and the quantitative and analytical sections of the Graduate Record Examination. Although there is no evidence to suggest that the basic abilities of minority students are different from those of White students, the achievement differences between minority and majority students are pronounced (Cocking & Chipman, 1988). Students with limited English proficiency may perform less well on tests because they read more slowly (Mestre, 1988).

Cocking and Chipman (1988) describe a study in which bilingual students with Spanish as the dominant language scored higher on the Spanish version of a math placement test than on the English version. In other studies (e.g., Macnamara, 1966), bilingual students showed lower performance when the language of instruction was the students' weaker language; evidence suggests that bilingual students keep pace with monolingual students in mechanical arithmetic but fall behind in solving word problems.

The literature linking language background with mathematics performance shows support for the idea of differential effect. What is the relevance of language to mathematical problem solving? A number of studies addressing this issue are reviewed below.

The Role of Language in Solving Math Word Problems

Text comprehension is a crucial step in the problem-solving process. This step calls for an understanding of ordinary English, the conventions of the word problem genre, and the special vocabulary and language structures of math problems (the mathematics register). We begin with the investigation of the latter by researchers from a variety of perspectives.

In 1960, Spencer and Russell (1960) claimed that the difficulties in reading mathematics are due to its specialized language and terminology (Cocking & Chipman, 1988). A call for the analysis of mathematical language was made by Aiken (1971, 1972) in a review of studies showing correlations between high reading ability and high arithmetic problem-solving ability. Munro (1979) examined syntax and vocabulary used in mathematics contexts and observed that they may differ from ordinary language; Rothman and Cohen (1989) noted the importance of the language and vocabulary of mathematics. Ginsburg (1981) found that the vocabulary children have for expressing math and number concepts differs widely.

Cummins et al. (1988) claim that word problems constitute tests of verbal sophistication as well as logico-mathematical knowledge. The problem below—

There are 5 birds and 3 worms.

How many more birds are there than worms?

—was answered correctly by 17% of nursery school children and 64% of first graders. However, when the last line was changed to

Suppose the birds all race over and each one tries to get a worm! How many birds won't get a worm?

the scores improved dramatically to 83% for nursery school children and 100% for first graders (Hudson, 1983). In other studies as well, changing the language of the problem to make the relationships clearer raised student performance (De Corte, Verschaffel, & DeWin, 1985; Riley, Greeno, & Heller, 1983). As pointed out by Cummins et al. (1988), if children fail to solve certain problems because they do not possess the conceptual knowledge required to solve them, one would not expect minor wording changes to improve solution performance; yet this is precisely what is observed. The results suggest that children find certain problems difficult because they cannot interpret key words and phrases in the problem text. According to Cummins et al. (1988), certain verbal formats allow contact to be made with superschema knowledge leading to a possible problem solution, while others do not.

De Corte et al. (1985) point out that word problems given to schoolchildren are often stated briefly and sometimes ambiguously because of presuppositions in the text. Experienced problem solvers have developed semantic schemata that serve to compensate for the omissions and ambiguities in the problem statement. In contrast, children and inexperienced problem solvers tend to rely more on text-driven processing. De Corte et al. (1985) hypothesize that rewording verbal problems so that the semantic relations are made more explicit without affecting the underlying semantic and mathematical structure facilitates constructing a proper problem representation and, by extension, finding the correct solution.

A number of studies have focused on the translation from English prose to numbers and mathematical symbols, a translation from textual to symbolic representations (Clement, Lochhead, & Monk, 1981; Kaput & Clement, 1979; Lochhead, 1980; Lochhead & Mestre, 1988; Mestre 1988; Mestre & Lochhead, 1983; Rosnick, 1981; Rosnick & Clement, 1980; Spanos, Rhodes, Dale, & Crandall, 1988). Researchers found, for example, that high school students and their teachers, as well as college students and faculty, made similar errors in translating sentences such as *There are six times as many students as professors;* they represented the statement as [6S = P] rather than [S = 6P], apparently by following a strategy of direct sequential mapping of words onto symbols.

Linguistic features of the mathematics register may present special difficulties for non-native speakers of English. Spanos et al. (1988) analyzed transcriptions of student descriptions of interpreting and solving mathematics items. They identified potential difficulties with comparative structures, prepositional phrases, article usage, conditionals, long nominals (noun phrases), and passive voice constructions, as well as unfamiliar cultural content and vocabulary items that have different meanings in the mathematics context.

Children's understanding of "more" and "less" in Grades 2 through 10 was studied by Jones (1982). He found that second-language learners lagged behind native speakers in their understanding of these terms in mathematical statements. The lag ranged from 2 to 4 years, depending on the type of context, and affected the children's ability to solve certain types of math word problems.

Orr (1987) argues that a major cause of performance errors in mathematics is dialect differences across communities. She focuses on differences between Black English Vernacular (BEV) and the language of mathematics. She cites numerous dialect differences, including methods of clause formation, math vocabulary, comparatives, and preposition usage. However, Baugh (1988) has challenged Orr's understanding of the linguistic complexity of structures such as comparative forms in BEV. Orr's study points to the need for research into the role of dialect variation in students' understanding of mathematics word problems. Work in this area has begun (see Lucas & Borders, 1994), but more is needed.

Cross-cultural conflicts may also affect the way language and mathematics interrelate. For example, an interview study of Native American students whose first language was Crow found that the students considered much mathematics vocabulary (taught in English) to be associated with school activities and that most out-of-school activities did not employ such vocabulary. Indeed, much mathematics vocabulary in the Crow language appears to be relatively new, according to Davison and Schindler (1988); further, they claim that the lack of cultural relevance of mathematics concepts was a problem for these students.

Mestre (1988) compared bilingual Hispanic 8th-grade students with monolingual students with the same level of mathematical sophistication and concluded that language deficiencies can lead to the misinterpretation of word problems. Mestre identified four proficiencies in language that interact to produce knowledge in the mathematics domain: proficiency with language in general, proficiency in the technical language of the domain, proficiency with the syntax and usage of language in the domain, and proficiency with the symbolic language of the domain. He advocates a less stilted and formal style that would nevertheless retain the precision and rigor necessary for mathematical discourse.

Among other factors indicative of potential linguistic complexity, an obvious candidate is the length of the problem statement. Lepik (1990) looked at a large number of structural and linguistic features in algebra word problems, including word length, number of words, number of sentences, and sentence length. He found the highest correlation between the number of words in the problem statement and problem-solving time; however, he did not find a significant relationship between any of the linguistic variables he considered and the proportion of correct responses. None of the variables correlating length of prompt with student achievement reached significance in Lepik's study, in contrast to the findings of Jerman and Rees (1972), who found a significant correlation between length of prompt and number of correct responses.

Researchers tend to agree that familiarity with features of mathematical language and the arithmetic word problem genre contributes to success with word problems, and that potential difficulties may arise for non-native speakers or speakers of nonmainstream dialects. In addition, there have been a number of studies on the impact of general language complexity on word problem performance. These studies are reviewed below.

Measurement of Linguistic Complexity

Research on the language of mathematics word problems has not been limited to the particular features of the mathematical prose register. Because the solution of a math problem requires clarity and precision of thought, a correct solution is more likely when the problem statement is comprehensible. The pursuit of overall comprehensibility has led researchers to examine more general aspects of linguistic complexity such as sentence length, word length, and word frequency (as incorporated in readability measures, for example); sentence structure and clause type; concrete versus abstract language; and culture-specific vocabulary and content. We consider these in this section.

The most direct way to test the comprehensibility of a language passage is to test how well a group of subjects actually comprehends it. This can be expensive and time-consuming. Furthermore, it is difficult to make valid and reliable comparisons among different passages (Lorge, 1939). In order to measure the difficulty of textbooks, readability formulas have been devised, using as a point of reference a set of passages that were tested directly with schoolchildren to determine their suitability for use at various grade levels (McCall-Crabbs *Standard Test Lessons in Reading,* 1925). Among the most widely used early readability formulas were the Dale-Chall formula (Dale & Chall, 1948) and Rudolph Flesch's Scale of Reading Ease (1948). The Dale-Chall formula computed a reading grade score for a passage, using (a) the average sentence length in number of words and (b) the percentage of words not on a list of 3,000 high-frequency words. The Flesch scale computed difficulty level from (a) the average number of words per sentence and (b) word length as measured by the number of syllables per 100 words. These formulas have validity coefficients of .70 with the *Standard Test Lessons in Reading* (Klare, 1974). Since then, most readability formulas, including one developed by Edward Fry in 1961, use similar word and sentence difficulty measures.

Not all long words are difficult or unfamiliar, but the most frequent and familiar words tend to be short (Zipf, 1949). For example, the 56 most frequent words in the Brown University Standard Corpus of Present-Day Edited American English are monosyllabic, and as frequency of occurrence decreases, the words tend to be longer (Kucera & Francis, 1967). Thus, word length can serve as an index of word familiarity. And, because word length is easily measurable, it is a convenient index for use in readability formulas.

Adams (1990) presents a tripartite cognitive model for reading, including phonological, orthographic and meaning components, with interaction effects between the three. The chief component of written word recognition is the reader's familiarity with the word. A reader who encounters a familiar word will interpret it quickly and will spend less cognitive energy in analyzing its phonological component. Carroll, Davies, and Richman (1971) analyzed lexical frequencies in books used in classrooms from Grades 3 to 8. They found that the number of different words was rather large—approximately 86,000—but that the vast majority of these words were encountered only infrequently. Word frequency estimation, therefore, becomes important in estimating the difficulty of a written sentence.

MacDonald (1993) has investigated written sentences in which there was a need to resolve lexical and grammatical category ambiguities (the word *trains,* for example, can be a noun or a verb). Her results show that word frequency in the lexicon, both within and across grammatical categories, was one of the primary factors contributing to the resolution of such ambiguities. An alternative to reliance on standardized passages for measuring comprehension is the Cloze procedure, in which words in a passage are deleted at intervals, for example, every fifth word (Taylor, 1953). Using Cloze items to assess comprehension difficulties of reading passages, Bormuth (1966) identified a number of linguistic variables that correlate with passage difficulty, including mean word depth, the ratio of verbs to conjunctions, and letter redundancy, as well as words per sentence and syllables per word. The concept of word depth is a sophisticated measure of syntactic complexity based on a tree diagram of the linguistic structure of a sentence (MacGinitie & Tretiak, 1971; Wang, 1970; Yngve, 1960). Bormuth found a correlation of .86 between sentence length and word depth; consequently, sentence length was supported as an index of complexity in computing readability. Thus, although sentence length may not be a cause of difficulty, it serves as a convenient index for syntactic complexity and can be used to predict comprehension difficulty.

Readability formulas have been criticized for failing to identify actual causes of difficulty. For example, subordinate clauses may contribute more to complexity than coordinate clauses (Hunt, 1965, 1977; Wang, 1970). Botel and Granowsky (1974) argue that sentence length offers little indication of the grammatical make-up and complexity of a sentence. They point out that certain linear arrangements of sentence constituents are more complex than others. Guided by transformational grammar theory, child language research, and their own intuitions, they propose a syntactic complexity formula that incorporates a weighting system for linguistic structures. In their system, the simplest structures include simplex sentences, questions, and coordinate clauses joined by *and;* among the more complex structures are passives, comparatives, and subordinate clauses. Freeman (1978) notes the limitations of any assessment of comprehensibility that counts surface features and does not consider abstract underlying forms from a transformational-generative viewpoint.

Finegan (1978) shows that the meaning-preserving linear rearrangement of syntactic constituents within a sentence can lead to substantially different degrees of comprehension. The readability formulas miss differences of this sort. In fact, they cannot discriminate between a well-organized sentence and the same words randomly ordered. Furthermore, the readability formulas measure the difficulty of an average sentence in a passage; they do not address discourse features that contribute to understanding by indicating relationships among sentences.

There is some disagreement in the literature regarding the value of readability measures in assessing the linguistic difficulty of mathematics items. Paul, Nibbeling, and Hoover (1986) showed that readability had little effect on students' ability to solve word problems. Using a number of popular measures of readability, they created forms of tests of equal mathematical difficulty but with readability quotients at, below, and above four different grade levels. They found that readability, as determined by the Dale-Chall (1948), Fry (1977), Harris-Jacobson (1973-74), and Spache (1953) formulas (which respectively employ number of syllables per hundred words, mean sentence length, and difficulty of vocabulary as primary determinants), was not an appropriate method for determining grade-level appropriateness of story problems. Kane (1968) notes the special nature of mathematics text as a mix of ordinary English and mathematical English; because the readability formulas are validated on ordinary English, it is inappropriate to apply them to text containing symbols and mathematical jargon (Kane, 1970). Perera (1980) also questions the applicability of readability formulas, since they fail to account for difficulties arising from unusual sentence structures or compressed language. Mathematics texts could receive lower readability scores than they in fact warrant because of the terse nature of the prose.

Addressing the problem of readability of mathematics texts, Noonan (1990) claims that studies show that a pupil's success in mathematics is closely related to his or her ability to read and interpret written material, and that many pupils who could attain success in mathematics are being handicapped because of a weakness in their reading skills. He notes that the simplification of mathematics texts by shortening the sentences and including diagrams does not necessarily mean that the texts will be more comprehensible.

Related to the readability of a text is the facility with which the reader processes the written symbols. This is not merely an issue of aptitude at reading but, for second-language learners, a question of different modes of access to the lexicon. Cross-cultural studies of written word perception (e.g., Rayner & Pollatsek, 1989) have shown that there are a number of different ways of processing written text across cultures. Even among alphabetic systems, there are differences in the directness of the correspondence between the orthographic system and the underlying phonological system. For students accustomed to a shallow orthography (that is, a written system with close correspondence between letters and phonemes, such as Spanish), reading a deep orthography (where the correspondence is less direct, and the spelling shows morpheme relationships, as in English) is more complicated.

Focusing on language in math story problems, Larsen, Parker, and Trenholme (1978) defined levels of syntactic complexity ranging from easy to hard in terms of sentence structure patterns ranging from simple to complex to compound to compound/complex. The authors used three tests at different levels of syntactic complexity but equal mathematical difficulty. The easy form used a sequence of two or three simple sentences followed by a question; the moderate form used a complex sentence, either as one of the problem statements or as the interrogative; the hard test included a compound and a complex interrogative sentence. The level of vocabulary, number of computations required to solve the problem, the difficulty of the computations, and verbal cues were all controlled. They found that low-achieving 8th-grade students' scores on word problems were significantly lower on problems containing structures of greater complexity. This study suggests that, although language complexity in math items may not show up as a salient problem for the total population, it may nevertheless be a significant factor for a specific group of students. In Phase 2 of the current study, a comparison of the performance of high-achieving and low-achieving math students yielded parallel results.

Many studies stop short of making specific recommendations for text simplification. An exception is a study by Shuard and Rothery (1984), cited in Noonan (1990). Among their explicit recommendations for ameliorating language problems in mathematics text are the following:

1. Use short sentences.

2. Use simple words.

3. Remove unnecessary expository material.

4. Keep to the present tense and particularly avoid the conditional mode. For instance, *"Given that butter costs 47p a block, what would be the cost of 5 such blocks?"* can be replaced by *"Butter costs 47p a block. How much do 5 blocks cost?"*

5. Avoid starting with sentence clauses. For example, *"Draw a circle of radius 4.2 cm."* is more readable than *"Using a radius of 4.2 cm., draw a circle."*

The practical recommendations of Shuard and Rothery are consistent with some of the research findings on readability, namely: Longer sentences tend to be more complex syntactically and, therefore, more difficult to comprehend; passages with words that are familiar (simple semantically) are easier to understand; words that are short (simple morphologically) are likely to be more familiar and, therefore, easier; long items (for example, with "unnecessary expository material") tend to pose greater difficulty; and complex sentences tend to be more difficult than simple or compound sentences. Implicit in their recommendation (item 4 above) is the suggestion that propositions are more difficult to process as conditional clauses than as separate sentences. Practical recommendations such as these could find support in results of future research.

Studies report better performance when reasoning tasks are presented in concrete formats (e.g., envelopes and postage requirements) rather than abstract formats (e.g., letters and numbers) (Cummins et al., 1988). Adults tend to use strategies more appropriate to the task when the rule statement is concrete rather than abstract (Reich & Ruth, 1982).

Saxe (1988) notes that word problems may cause special problems for the language minority student, and that tests may be biased toward sensitivity to the mainstream culture. According to Mestre (1988), research illustrates that the poor comprehension skills of Hispanic bilinguals adversely affect their interpretation of math word problems. Bilingual Hispanic students possessing the same level of mathematical and computational sophistication as their monolingual peers often solve word problems incorrectly. In one study, Hispanic 9th-grade students were asked to solve word problems selected from their algebra textbook. One problem contained the culture-specific vocabulary items *stock* and *share;* another contained *revolving charge account, monthly payment,* and *interest*. When the students were asked to tell an interviewer what they thought these terms meant, their responses showed that they had little idea. Nevertheless, the students attempted to solve the latter problem by combining the problem's four monetary quantities to obtain an (incorrect) answer. According to Mestre (1988), because problem solving involves reading and comprehending the problem under consideration, ability to understand written text is of paramount importance.

The reading and language problems of children from low-income homes are discussed by Chall, Jacobs, and Baldwin (1990). For these children, reading difficulties in the intermediate grades lead to trouble in other subjects, such as mathematics, that are learned in part from printed text. Chall et al. (1990) found that children from low-income families had particular difficulty with words that were less familiar, longer, or more specialized. The effect of unfamiliar vocabulary in math word problems became a focus in Phase 2 of the current study, and in fact emerged as one of the linguistic features showing an effect on student performance.

A large body of literature addressing additional background variables not directly related to the research questions of this study was reviewed. Because these studies indirectly relate language background with student performance, they are mentioned here. Some studies include reports on significant relationships between students' self-concept and locus of control (defined in Sue & Sue, 1990) and their academic progress (Burger, 1992; Byrne, 1984; Chadha, 1989; Cone & Owens, 1991; Flynn, 1991; Hagborg, 1991; Laffoon, Jenkins-Friedman, & Tollefson, 1989; Luthar & Zigler, 1992; Lyon & MacDonald, 1990; Maqsud & Rouhani, 1991; Marsh, Walker, & Debus, 1991; Winefield, Teggemann, Goldney, & Winefield, 1992). Other studies discuss cultural background and differential schemes for acquisition of particular content area knowledge (Roseberry, Warren, & Conant, 1992; Saxe, 1988; Travers, 1988).

Research Perspective for This Study

Certain research findings particularly influenced the focus of this study. Specifically, the many previous studies identifying sentence length as an index of linguistic complexity suggested the use of this variable in our Phase 1 analyses. Previous studies found that, for longer math items, students took longer to answer them and answered fewer correctly. A number of previous studies of language difficulty identified features including word frequency/familiarity, passive versus active voice, clause type, linear sequence within a sentence, and concrete versus abstract presentation; these studies influenced our selection of certain linguistic features for close examination in both Phase 1 and Phase 2 of this study.

The range of linguistic factors relevant in mathematics assessment is very broad. A common theme across many of the studies reviewed is that the complexity of linguistic features in math items must be considered to be a separate issue from the mathematical complexity of those items. Another theme in the literature is that a student's language background is an important factor in mathematics assessment. The latter theme was the impetus for the investigation in Phase 1 of this study.

In the following sections we present the two phases of the language background study conducted by CRESST/CSE—Phase 1: Analyses of Existing Data and Phase 2: Field Research—in the sequence in which they were conducted. Phase 2 of the study included three field studies: a Student Perceptions Study, an Accuracy Test Study, and a Speed Test Study. For each section we present methods, results, and discussion. We conclude with a general discussion of both Phase 1 and Phase 2 findings.

PHASE 1: ANALYSES OF EXISTING DATA

The primary intent of Phase 1 of the study was to analyze the 1990 and 1992 NAEP mathematics test data for Grade 8 to determine whether certain language background variables showed any impact on students' performance on the NAEP mathematics test. It was hypothesized, first, that there would be a relationship between the length of items and performance. Do long items negatively affect performance? Second, it was hypothesized that the linguistic complexity of items would have a negative impact on performance. Does linguistic complexity contribute to poorer performance on NAEP math items? Third, it was hypothesized that students who come from a language background other than English would have more difficulty processing the items, a difficulty that would be evidenced by a higher number of items omitted or not reached. Do students who report speaking a language other than English in the home have a significantly greater number of items omitted/not reached?

Methodology

To address these questions, a series of analyses were performed. Included among them were an analysis based on length of items, an analysis based on a general assessment of the linguistic complexity of the items, and an analysis of items omitted/not reached. The results from each of the analyses will be discussed using one or more of the NAEP background variables from the 1990 and 1992 data (see Tables 1 and 2 for descriptions; see Appendix A, Tables 1-25).

Table 1 presents the names and the descriptions of the background variables that were used for grouping subjects in a Discriminant Analysis (DA) for the 1990 data. Table 2 presents the same information for the 1992 data. The methods used for each DA will be described; then, the results from each analysis will be presented in turn.

Analysis Based on Length of Items

For this analysis, we created two composite scores; the first was a composite of the items with long stems and/or long choices, and the second was the composite of items with short stems and short choices. We categorized items based on the length of the stem and/or the answer choices. If the length of the stem was two lines or more, and/or the answer choices were longer than one line, we categorized them as "long items." If the length of the stem was shorter than one line and the answer choices were shorter than one line, we categorized them as "short items." Medium-length items were discarded from the analysis because they were considered an arbitrary category.

After creating the two item parcels (long and short), we then obtained the average scores for the parcels and used those average scores in a series of discriminant analyses in which selected background variables were used as grouping variables. The background variables from the 1990 and 1992 data were student-reported and included gender, questions regarding the amount and type of reading material in the home, and questions addressing how often a language other than English is spoken and read in the home.

Analysis Based on the Linguistic Complexity of Items

For this analysis, we categorized mathematics items according to degree of linguistic complexity, using pragmatic criteria such as difficulty of vocabulary, abstract or culture-specific content, and number of complex structures in a sentence. We created two composites of items, one with greater complexity and one with lesser complexity. Analyses of variance repeated measures design was applied to the parcel scores that were created from this linguistic categorization of these items.

Analysis of Omitted/Not Reached Items With Language Background

Percentages of omitted/not reached items were computed with respect to NAEP language background variables (see Tables 28-33). We obtained the percentages of omitted/not-reached items for all the 8th-grade students in the three separate categories of LANGHOM (Group 1 always spoke English in the home, Group 2 sometimes spoke a language other than English in the home, and Group 3 always spoke a language other than English in the home).

Analysis of the omitted/not-reached items by linguistic characteristics was also performed. The simplest linguistic feature of the items—the item length—was used also as a criterion for this categorization. Two item parcels were created, one consisting of items that had a long stem and/or long answer choices, and the other consisting of items that had short stems and short answer choices. The two parcels' percentages of omitted/not-reached items were compared across the student subgroups that were formed based on the LANGHOM language background variable. This analysis was conducted on the booklet level so that the results from the individual booklets could be used as cross-validation data. The analyses were performed on randomly chosen booklets 1, 2, 15, 19, and 24 for the 1990 and 1992 data. (It must be noted, however, that some of the items were common across several booklets.)

Results

The overall results from the analyses were similar in that each indicated a significant role for language background variables on extant NAEP data. The results for each of the individual analyses follow.

Results of Analysis Based on Length of Items

Students who always spoke a language other than English in the home scored lower than students who always spoke English at home, especially on longer test items. The impact of other variables, such as gender, was not influential.

Table 3 summarizes the results of the DA for the long and short items by gender for the 1990 data. (Gender was selected for the purpose of checking validity of sample by comparison with extant NAEP data, in order to show that categorization by variables other than language, such as gender, did not show significant impact on math performance.) Similarly, Tables 4 through 9 present the results of the DA for the second, third, fourth, fifth, seventh, and eighth background variables (respectively, see Tables 1 and 2) used in this analysis for Booklet 8.[1] Tables 10 through 16 present similar results for the 1990 data, Booklet 9; and Tables 17 through 23 present the results for Booklet 10, 1990.

For the 1992 data, we performed similar analyses on the same set of background variables. However, there were some major differences between the 1990 and 1992 data structures. In the 1990 data, all of the math items were distributed across 10 booklets; in the 1992 data, the items were distributed into 26 booklets. As a result, the total number of subjects per booklet was less on the 1992 data. Thus, for the 1992 data we had to combine several booklets to obtain enough subjects to permit the categorization of subjects into groups based on their background variables. Table 24 summarizes the results of the DA for Booklets 1, 2, and 15 from the 1992 assessment, and Table 25 presents the results of the DA for Booklets 10, 19, and 24 for the 1992 assessment.

These results clearly indicate that language background does in fact affect test performance. This is evident from the categorization of the questions/items by the simple linguistic factor of length. Student performance was similar across booklets on most variables, including gender. However, when the subjects were grouped based on certain relevant language background variables, significant results and interesting trends were observed. Table 15 presents the results of the DA on the background variable "How often is a language other than English spoken in your home?" for the 8th graders who received Booklet 9. Students were grouped into three categories according to their response to this background variable question: (1) *never* (i.e., never spoke any language other than English in the home), (2) *sometimes*, and (3) *always* (i.e., always spoke a language other than English in the home). As Table 15 indicates, the average math score of the long items for Group 1 (i.e., never spoke any language other than English in the home) was .43. For Group 2 (*sometimes*) the average was .41, and for Group 3

---

[1] Since all booklets were considered to be parallel, booklet selection was random.

(*always*) the average was .32. The averages for the short items for Groups 1, 2, and 3 were .53, .51, and .46 respectively. Function 1 significantly discriminated between the three groups of subjects (r = .130, χ^2 = 20.87, and p < .01). These results clearly reveal lower math proficiency scores for the subjects who speak more of a language "other than English" in the home. This relationship was also more evident for the longer items than the shorter items (see Table 15).

The results of the analyses on Booklets 8 and 10 were consistent with the results that were obtained for Booklet 9. The results obtained for the 1990 data were also consistent with the results that were obtained for the 1992 data. For cross-validation purposes, we compared the results of the three booklets in the 1990 data. The comparisons led us to believe that the booklet results could be considered replications. Results shown in Tables 3 through 9 were compared with results in Tables 10 through 16, and each in turn with results in Tables 17 through 23. We also compared the 1990 results (Tables 3 through 23) with the 1992 results (Tables 24 and 25). These comparisons revealed consistencies in the results obtained from analyses of performance of the different groups of students that took the test at different points in time.

Results of Analysis Based on the Linguistic Complexity of Items

Analyses showed significant differences with respect to language background between student scores on complex items and less complex items. On the parcels with greater linguistic complexity, the performance of students who spoke more of a language other than English in the home was significantly lower than the performance of students who spoke only English in the home. Language was found to be more important than other background variables such as gender. The mean for Group 1 (never spoke a language other than English) was .58, and the mean for Group 2 (always spoke a language other than English) was .42 (t = 2.97, p < .01) (see also Tables 26 and 27).

In addition to analyses by language in the home, we conducted analyses by gender for purposes of comparison; differences with respect to gender were not expected, and were not found. We used gender and language spoken in the home (LANGHOM) as between-subjects variables and parcel composite scores as within-subjects variables. Table 26 (reproduced here from Appendix B) presents the results of an analysis of variance repeated measures on parcel scores of linguistically complex and noncomplex items by gender for 8th-grade students

Table 26

Analysis of Variance Summary Table, 1992, 8th Grade, Block 8

Source of variation	SS	*df*	*MS*	*F*	*P*
Between subjects					
A (sex)	.16	1	.16	1.43	.23
Subject within group					
Within subjects					
B (problems)	3.24	1	3.24	56.42	.00
AB (sex x problem)	.26	1	.26	4.44	.04
B x subject within group					

on Booklet 8. As Table 26 indicates, there was no significant difference between the two composite parcel scores by gender, $F = 1.43$, $p = .23$. The within-subjects main effect (linguistically complex versus nonlinguistically complex parcel scores) was significantly high, $F = 56.42$, $p = .01$. The gender-by-parcel scores interaction was also significant, $F = 4.44$, $p = .04$. This significant interaction could introduce some difficulty in interpreting the between- and within-subjects main effects.

For cross-validation purposes, we ran the same analyses on the four other booklets (booklets 3, 12, 13, and 15). Table 27 (reproduced below) presents the results of the analyses on Booklet 15, for example. As Table 27 indicates, the within-subjects main effect (linguistically complex versus noncomplex) was significant, $F = 67.96$, $p < .01$; however, the main effect of gender, $F= 23.71$, $p < .01$, and the interaction effect between gender and linguistic complexity were also significant, $F = 5.73$, $p < .05$.

Analysis of Omitted/Not Reached NAEP Math Items

For students who always spoke a language other than English in the home, results showed higher percentages of test items omitted or not reached. Table 28 presents the percentages for items either omitted or not reached, separated into three groups (see Appendix C for Tables 28-33). The first column of statistics lists the percentage for those students who *always spoke English in the home*. The second column lists the results for those students who *sometimes spoke a language other than English in the home*, and the third column lists the results for those students who *always spoke a language other than English in the home*.

Table 27

Analysis of Variance Summary Table, 1992, 8th Grade, Block 15

Source of variation	SS	*df*	*MS*	*F*	*P*
Between subjects					
A (sex)	2.65	1	2.65	23.71	.00
Subject within group					
Within subjects					
B (problems)	3.33	1	3.33	67.96	.00
B x subject within group					
AB (sex x problem)	.28	1	.28	5.73	.02
AB x subject within group					

When comparing columns 1 and 3 (the two extremes), in almost all cases the students who always spoke a language "other than English in the home" had much higher percentages of omitted/not-reached items than the students who always spoke English in the home. These results are consistent with what was initially hypothesized; they reveal the impact of language on students' math performance.

Tables 29 through 33 present the results of the analyses for Booklets 1, 2, 15, 19, and 24 respectively for the 1992 data. As the results indicate, for every booklet, students who always spoke a language "other than English" in the home had a much higher percentage of omitted/not-reached items on longer items than those students who spoke "only English" in the home. For the shorter items, the differences between the two groups (students who always spoke English in the home versus students who never spoke English in the home) were not as large as the differences between the two groups for the longer items.

Discussion of Analyses of Existing Data

The results of the analyses of extant data indicate that there was a significant effect upon NAEP math student performance from language background. This effect was evident in studies based upon the length of the item, the linguistic complexity of the item, and in an analysis of items omitted/not reached. Students who always spoke a language other than English in the home scored lower, particularly on long items and linguistically complex items;

compared to students who always spoke English at home, their tests showed more items omitted or not reached.

In Phase 1 of the study, the initial focus was intentionally narrow; analyses were conducted using simple, pragmatic measures of item length and complexity. Based on the combination of strong statistical findings from Phase 1 and continuing questions about the linguistic characteristics of NAEP math items, we hypothesized that NAEP math items might be revised in such a way as to ameliorate the disadvantage faced by students who could be negatively affected by specific linguistic factors within the items. In order to investigate this hypothesis, a sequence of field studies was undertaken in Phase 2. The sources and the nature of the linguistic disadvantages faced by limited-English-speaking students and non-native speakers of English are potentially very diverse. A number of salient factors were addressed in Phase 2; however, it should not be inferred that these are the only possible linguistic factors affecting student performance.

PHASE 2: FIELD RESEARCH

This section describes the field research phase of the study, which consisted of three studies investigating the effects of language simplification in mathematics word problems. We describe here the general approach, the selection of mathematics items, and the identification of linguistic features for investigation. Then each of the three field studies is described.

Three separate field studies were conducted in Phase 2. For the first field study, the Student Perceptions Study, we interviewed 8th-grade students in the greater Los Angeles area. These students were given the original (linguistically complex) NAEP math items and their revised sister items (with simplified language; see below) in a structured interview format to discover the students' perceptions and preferences.

The second field study in Phase 2 was the Accuracy Test Study.[2] In this study, 39 8th-grade classes (1,031 students) were selected, with an intention to include LEP students. Released items from the 1992 main mathematics assessment were examined for linguistic complexity, and the most linguistically

---

[2] Previously, this study was known as the Performance Study, a name that was changed to avoid confusing it with a performance assessment.

complex items were identified. A revised set of items with simpler language but original math content was then created. The two sets of items were placed into two booklets, along with five released NAEP items judged to contain relatively simple language as controls. The distribution of the items into the two booklets was based on the content of the items, the difficulty level of the items and other related factors. We tried to make the two booklets as similar as possible, although we were not attempting to produce parallel forms. In addition to the 25 math items, each booklet contained a language background questionnaire that was specifically designed for the study. It was hypothesized that the simplified language of the revised items would help students achieve more correct answers on the revised items than on the original items; further, an analysis of the impact of different types of linguistic changes might suggest which linguistic features are most likely to affect student performance.

The data obtained from the Student Perceptions Study showed that the majority of the students understood and preferred the revised items over the original items when asked which set they would choose in a timed test setting. Based on these results, we decided to examine the effects the linguistic modifications had on the time a student required to answer/complete the math test items. Two additional booklets (Forms A and B) were developed for this third study, the Speed Test. For this study, the 20 original NAEP items were placed in one booklet and the linguistically simplified items were placed in a separate booklet. The students were allowed only 10 minutes to work on the test. The same language background questionnaire used for the Accuracy Test Study was included. Our hypothesis was that students would be able to complete more simplified items than original ones, and that this trend would be more evident for students who did not speak English as their primary language.

Selection of Math Items

NAEP math items available for field research included 1992 items from 4th-, 8th-, and 12th-grade blocks. Because previous studies by the Technical Review Panel (TRP) and other researchers have focused primarily on 8th-grade items, and because some of the 8th-grade items were found in 4th- and 12th-grade blocks as well, they appeared to be more appropriate for this study. Consequently, the 8th-grade items became the focus of the study and were used for all three field research studies. The 69 released items from the 1992 8thgrade blocks were reviewed for linguistic features that could be potentially problematic for 8th graders.

Identification of Linguistic Features

The process of identifying the potentially problematic linguistic features in the NAEP math items was threefold and iterative. The research literature, expert knowledge, and the actual characteristics of the NAEP items led to the identification of the features. First, sources in the research literature provided guidance (e.g., Spanos et al., 1988, discussed in the Introduction section above, and other studies discussed below). Second, the process was informed by project staff's knowledge of the types of linguistic features likely to cause problems for adolescents and for learners of English as a second language. Finally, the set of linguistic features selected for study was limited by that subset of structures appearing in the 69 NAEP items that constituted the corpus for the investigation.

Each of the 69 items was read and the mathematical operations attempted. Items in which the language was considered potentially difficult for students to understand were flagged and analyzed; linguistic features likely to contribute to the difficulty were identified and categorized.

Simplified forms of linguistically complex items were drafted in order to make these items easier for students to understand. The process was iterative in that project staff worked back and forth between revising the items and refining the categorization scheme, guided by the review of the literature, expert knowledge, and Phase 1 analyses of NAEP items. The set of potentially difficult linguistic features found in the Phase 2 test items and the strategies used to revise them are given below; however, it should be noted that this categorization is by no means exhaustive, as the set of linguistic features is limited to features actually occurring in the small corpus of released NAEP math items referenced above. From this set of features, only the most salient and frequent language problems were selected for investigation in the field study.

Changes in seven categories were made to the language of the original NAEP items: familiarity/frequency of nonmath vocabulary, voice of the verb phrase, length of nominals (noun phrases), conditional clauses, relative clauses, question phrases, and abstract or impersonal presentations. Changes in each of these areas are described and illustrated below.

Familiarity/Frequency of Non-Math Vocabulary

Potentially unfamiliar, low-frequency lexical items were replaced with more familiar, higher frequency lexical items.

Original: *A certain reference file contains approximately six billion facts*.

Revision: *Mack's company sold six billion hamburgers*.

In the student's world, the concepts of "company" and "hamburger" are probably more familiar, and are probably encountered more frequently, than "reference file" and "facts."

The vocabulary of a typical 8th grader is not equivalent to that of a typical adult; the teenage years are a period of continuing growth in vocabulary (Dunn, 1959). School-age children and adults add new words to their lexicons, but they also expand their definitions of familiar words (McNeill, 1970). Consequently, not all 8th graders show comparable vocabulary development. If a student does not understand all the words in a test item, he/she may not understand what the item is asking and may be unable to solve it. If an item contains unfamiliar vocabulary, it may take the student longer to read and understand the item, and the student may be at a disadvantage compared to other students on a timed test. The accuracy and speed of written word recognition depend on the reader's familiarity with the word in print (Adams, 1990). A task places greater demands on a student if his attention is divided between employing math problem-solving strategies and coping with difficult vocabulary and unfamiliar content (Gathercole & Baddeley, 1993). In the student interviews conducted for this study (see Student Perceptions Study, below), some students commented on the presence of "complicated words" and stumbled when they tried to read aloud items that contained words that seemed to be unfamiliar to them.

In revising the items, estimates of familiarity/frequency of vocabulary were made based on established word frequency sources as well as staff judgments of the students' familiarity with the words and concepts. For example *The American Heritage Word Frequency Book* (Carroll et al., 1971), based upon 5 million words from textbooks and library materials for Grades 3 through 9, and the *Frequency Analysis of English Usage: Lexicon and Grammar* (Francis & Kucera, 1982), based on the one million-word Brown University Corpus, listed the word "company" as occurring more frequently than "reference" or "file," a result that was consistent with our intuitions. However, both sources listed the word “hamburger” as much less frequent than “fact.” Nevertheless, we used the word “hamburger” in our revision because the word and concept are frequently encountered in the student culture in contexts (not included in the corpora) such as signs, advertisements and menus, as well as in spontaneous language through television and radio media, and through spontaneous conversation. No useful frequency list of spoken English for this age group is available yet; Hall, Nagy, and Linn (1984) provide spoken word frequencies for a corpus of one million words spoken by and to four- and five-year-old children, and here “hamburger” is more frequent than “fact.” Existing frequency sources are of limited usefulness; although Francis and Kucera give information according to grammatical category, Carroll, Davies, and Richman do not, so they obscure the fact that “carpet” is less frequent as a verb than as a noun. And, contrary to the experience of contemporary teenagers, both sources list “video” as less frequent than “census.”

Voice of Verb Phrase

Verbs in the passive voice were replaced with verbs in the active voice.

Original: *A sample of 25 was selected.*

Revision: *He selected a sample of 25.*

People find passive verb constructions more difficult to process than active constructions (Forster & Olbrei, 1973) and more difficult to remember (Savin & Perchonock, 1965; Slobin, 1968). Passive constructions occur less frequently than active constructions in English (Biber, 1988). Children learning English as a first language have more difficulty understanding passive verb forms than active verb forms (Bever, 1970, deVilliers & deVilliers, 1973). Typically, children do not develop a full understanding of passive forms until their elementary school years; some passive forms do not appear until age 11 (Owens, 1988). In addition, passive constructions can be complicated for non-native speakers of English for a number of reasons. First of all, passives in most languages are used much less frequently than in English, and in more restricted contexts (Celce-Murcia & Larsen-Freeman, 1983). Also, passives tend to be used much less frequently in conversation than in certain types of formal writing, such as scientific writing (Celce-Murcia & Larsen-Freeman, 1983). For these reasons, non-native speakers may not have had much exposure to the passive voice and may not be able to process passive sentences as easily as active sentences. Adolescent native speakers, as well, may have difficulties with the passive voice because of lack of exposure to this structure.

Length of Nominals

The number of prenominal modifiers in a noun phrase was reduced, as in the example below, where the list of nouns and adjectives preceding "president" was shortened in the revised form.

Original: . . . *last year's class vice president* . . .

Revision: . . . *vice president* . . .

Postnominal modifiers, including prepositional phrases and participles following a noun, were reduced or recast.

In processing novel nominal compounds, people use lexical information as well as knowledge of the world and the context to rule out implausible readings. Faced with the task of interpreting a long nominal, a student with a limited English vocabulary will be at a disadvantage. It may take her longer to interpret the phrase, and her interpretation may be incorrect. Long nominal compounds are inherently syntactically ambiguous, and a reader's comprehension of a text may be impaired or delayed by problems in interpreting them (Halliday & Martin, 1993; Just & Carpenter, 1980; King & Just, 1991; MacDonald, 1993).

Postmodifiers can be similarly ambiguous; for example, in a noun phrase followed by two prepositional phrase modifiers, such as "the man in the car from Mexico," the man may be from Mexico, or the car may be from Mexico. Adding more modifiers multiplies the possibilities for ambiguity.

Romance languages such as French, Spanish, Italian, and Portuguese make less use of compounding than English does, and when they do employ the device, the rules are different; consequently, students whose first language is a Romance language may have difficulty interpreting compound nominals in English (Celce-Murcia & Larsen-Freeman, 1983).

Conditional Clauses

Some conditional *if* clauses were replaced with separate sentences. In some instances the order of the *if* clause and the main clause was reversed.

Original: *If x represents the number of newspapers that Lee delivers each day . . .*

Revision: *Lee delivers x newspapers each day.*

The semantics of the various types of conditional clauses in English are subtle and hard to understand even for native speakers (Celce-Murcia & Larsen-Freeman, 1983). As the example above illustrates, in some contexts the presence of an overt conditional marker, such as *if,* is not needed to signal a conditional relationship between two events or situations; readers can infer the relationship without the marker (Mann & Thompson, 1986). Non-native speakers may omit function words (such as *if*) and may employ separate clauses without function words (Klein, 1986). Separate sentences, rather than subordinate *if* clauses, may be easier for some students to understand (Spanos et al., 1988). Statistically, languages of the world prefer conditional clauses in iconic order—that is, preceding main clauses rather than following them. In fact, some languages do not allow sentences with the conditional clause in last position (Haiman, 1985). Consequently, sentences with the conditional clause last may cause difficulty for some non-native speakers.

Relative Clauses

Some relative clauses were removed or recast.

Original: *A report that contains 64 sheets of paper . . .*

Revision: *He needs 64 sheets of paper for each report.*

In this example, the original version contains information in a relative clause, whereas the revised item contains the same information in a separate, simple sentence. While the number of sentences in the item is increased, the number of clauses per sentence is reduced. Shorter sentences with lower information density levels are more easily processed by students.

Since relative clauses are less frequent in spoken English than in written English, some students may have had limited exposure to them (in fact, Pauley and Syder, 1983, argue that the relative clauses in literature differ from those in spoken vernacular language). This fact of little exposure, along with the complexity of the form, is reflected in the late acquisition of relative clauses; although children learning English as a first language acquire a command of most grammatical structures in their preschool years, they do not develop a full structural knowledge of relative clauses until their school years (Tager-Flusberg, 1993). Students who learn English as a second language may find that English employs unfamiliar devices such as relative pronouns instead of particles or affixes. In English, relative clauses follow the noun, but relative clauses precede the noun in other languages such as Chinese and Japanese; furthermore, English relative clauses differ from those in some languages in that a pronominal reflex (pronoun) may be absent (Schachter, 1974). Relative clauses in English may be difficult for a non-native speaker to interpret if his first language employs patterns that are different from those of English.

Question Phrases

Some question structures were changed from complex question phrases to simple question words—for example, from WH-NP to simple WH forms.

Original: *At which of the following times . . . ?*

Revision: *When . . . ?*

The complex question phrase in the original version was replaced with a single question word in the revision. The single-word structure is simpler syntactically, and the placement of the question word at the beginning of the sentence gives it greater salience. The longer question phrases occur with lower frequency, and low-frequency expressions will in general be harder to read and understand (Adams, 1990).

Abstract or Impersonal Presentations

In some instances, an abstract presentation mode was made more concrete.

Original: *The weights of three objects were compared using a pan balance. Two comparisons were made. . . .*

Revision: *Sandra compared the weights of three objects using a pan balance. She made two comparisons. . . .*

In this example, the problem statement was made more story-like by the introduction of "Sandra." (Abstract or nonsituated items may employ the passive voice, but not all passive constructions are abstract or nonsituated; abstract/impersonal presentations may also employ modals or generic nominals, for example.)

Recent studies suggest that information presented in narrative structures is understood and remembered better than information presented in expository text. A problem expressed in concrete terms may be easier for students to understand than an abstract problem statement (see, for instance, Lemke, 1986).

The three field research studies in Phase 2 utilized math items containing one or more instances of one or more of the linguistic features described above. The following three sections describe the field studies: first, the Student Perceptions Study; next, the Accuracy Test Study; and last, the Speed Test Study.

Student Perceptions Study

The Student Perceptions Study investigated the following questions in interviews with 8th-grade students. For mathematics items with parallel mathematics content, do students respond differently to items that contain different linguistic structures? Do students find linguistically simpler items easier to comprehend? Do they show a preference for items with simplified language?

Method

Interviews were conducted with 8th-grade students; their comments on various pairings of mathematics items were audio tape recorded. Items were selected from the 69 released NAEP 8th-grade 1992 items and were revised as discussed above. Student comments on original and revised versions of items were analyzed.

Subjects. A total of 38 students at four school sites in the Los Angeles area were interviewed. The students represented a cross-section of ethnic and language backgrounds, and their current grades in math class ranged from A to D. The native languages, in addition to English, included Spanish, Cambodian, and Vietnamese.

Interview procedure. Each recorded interview lasted 10-15 minutes. After a brief introductory conversation, the student was asked to read a pair of math items—the original item and the corresponding revised item—and was asked questions such as "Which one do you think would be easiest to do?" or "If you were in a hurry on a test and you had to choose just one of these to do, which one would you choose?” A sample protocol for the interview is provided in Appendix D.

Stage 1 interviews. Four items representative of typical complicated features were selected by project staff, and a linguistically simplified version was written for each item. For each linguistically simplified version, an effort was made to retain the original math concept while changing names, numbers, diagrams, non-math vocabulary, and sentence structures. For each item, both versions—the original and the revised—were presented to each student in individual interviews. The original versions contained linguistic features such as:

passive verb forms: *A sample of 20 was selected . . .*

unfamiliar vocabulary: *A certain reference file . . .*

conditional subordinate clauses: *If two batteries in the sample were found to be dead . . .*

In the revised versions,

passives were deleted: *He selected a sample of 20 . . .*

unfamiliar vocabulary was replaced: *Mack’s company . . .*

conditional subordinate clauses were replaced with separate sentence structures: *He found three broken skateboards in the sample*.

For some items, the revision changed more than one structure; for example, in the last item above, the conditional subordinate clause was removed, and, in addition, the passive verb form and unfamiliar vocabulary were changed.

Stage 2 interviews. A second stage of interviews was conducted in order to obtain student feedback on other potentially complex language features. Three new items and a reworked version of the fourth item from the Stage 1 interviews were presented to 17 8th-grade students from two schools in the greater Los Angeles area. For these items, the original math concepts were retained in the revisions, as were some names, numbers and diagrams. As in Stage 1, complex vocabulary and sentence structures were changed.

Results

For the Stage 1 interviews, a total of 19 students from two schools participated. The majority of these students picked the revised version for items

1 and 2. See Table 34 for Stage 1 interview results. Table 35 (reproduced below) presents the choices made by the students in the Stage 2 interviews. In Stage 2, the majority of the students chose the revised version of all items, including the modified version of item 4. Following is a discussion of the reasons students gave for their choices.

Discussion of Student Perceptions Study

As the recorded responses from both Stage 1 and Stage 2 interviews demonstrate, in general, the students chose the revised items as easier to understand or preferable to the original items in terms of language. Indeed, the responses of many students showed an awareness of the linguistic features in the items.

Student comments about the items were of three types: (a) general difficulty in understanding, (b) length of items, and (c) complexity of vocabulary. In addition, difficulties with vocabulary and syntax could be inferred from problems students had in reading the items aloud. Some examples of student comments follow.

1. Many students reported a global judgment that the language in the revised item was easier to interpret. They said such things as:

"Well, it makes more sense."
"It explains better."
"Because that one's more confusing."
"It seems simpler. You get a clear idea of what they want you to do."

Table 35

Stage 2 Interview Results: Students' Choices (N = 17)

Item #	Original item chosen	Revised item chosen
4[a]	3	14
5	4.5[b]	12.5
6	2	15
7	2	15

[a] Modified (piloted for a second time) version of item #4.
[b] One student was ambivalent about his choice.

2. Some students made specific reference to time pressure as a factor in taking tests; some commented on the length of the items, as in the following examples:

> "It's easier to read, and it gets to the point, so you won't have to waste time."
>
> "I might have a faster time completing that one 'cause there's less reading."
>
> "Less reading; then I might be able to get to the other one in time to finish both of them."
>
> " 'Cause it's, like, a little bit less writing."

3. Some students commented on the difficulty of vocabulary items. They indicated, as in the examples below, that the vocabulary in the revised items was more familiar to them.

> "This one uses words like 'sector' and 'approximation,' and this one uses words that I can relate to."
>
> "It doesn't sound as technical."
>
> "I can't read that word."
>
> "Because it's shorter and doesn't have, like, complicated words."

In addition to explicit student comments about the items, further insight about problems with vocabulary and syntax was gained from having students read both versions of each item aloud. When a student is reading, pauses for unfamiliar words or constructions are likely to disrupt the flow of comprehension (Adams, 1990). Some students stumbled on words such as "certain," "reference," "entire," and "closet." In reading aloud an original item containing a passive verb construction, one student substituted an active verb form; the item contained the verb phrase "would be expected," but the student read it aloud as "would you expect to find," replacing a less familiar construction with a more familiar one. The student read the revised version as it was written.

The student responses showed clear differences between the original and the revised item in each pair. Student preference for the revised items gave support to the notion that the math items could be linguistically simplified in meaningful ways for the test taker. The interview results supported the plan to test a larger group of students to determine whether the observed differences in student responses to the language of the math items would be translated into actual differences in math test scores.

Accuracy Test Study

The purpose of the second field study, the Accuracy Test Study, was to examine the impact of revision of selected linguistic features in NAEP math test items on the number of test items answered correctly by students. A test consisting of word problems containing potentially difficult linguistic structures and also items with simplified language was administered, along with a questionnaire on the student's language background. The following sections describe the method used, present the results, and discuss the findings.

Method

Test forms containing original NAEP items and those same NAEP items with simplified language were administered along with a questionnaire on student language background. Student scores on simplified items were compared with scores on original items. Subjects, instruments, and procedure are described here.

Subjects. For the Accuracy Test, 1,031 8th-grade students from 39 classes in 11 schools from the greater Los Angeles area were selected. Schools were selected for participation to provide a range of language, socioeconomic, and ethnic backgrounds.

Information was obtained from school personnel on students' English proficiency, language background, grade level, type of math class, grades in math class, gender, ethnicity, and socioeconomic status (SES) (see Tables 36-42 in Appendix G). Information on students' English language proficiency from the Home Language Survey (HLS, a survey administrated by the school district) was obtained from school personnel. In Los Angeles schools, the HLS (see Appendix N) is administered to determine whether a language other than English is spoken in the home. Based on the HLS results, English language assessment tests are administered, leading to a classification of the student's English proficiency. These classifications were obtained, where available, for students in the study. The results indicated that approximately 31% of the students had been assigned to ESL categories ranging from Initially Fluent in English (4.8%) to Redesignated Fluent (8.7%) to Limited English Proficiency (LEP) (9.2%) (see Table 41). Most students were 8th graders (95%); 5% were in Grade 7 or 9. Types of math classes included honors algebra, algebra, high mathematics, average mathematics, low mathematics, and ESL mathematics. The student group was 54% male, 46% female.

Data on ethnicity were obtained from the schools[3]; 35% were Latino, 26% were White, 19% were African American, 16% were Asian American, and 5% were listed as Others or Declined to State (see Table 37). Estimating from the limited data available, roughly 36% of the students were categorized as low socioeconomic status (SES) on the basis of participation in free school lunch programs or in Aid to Families with Dependent Children (AFDC) programs.

Instruments. For the Accuracy Test Study, instruments included a language background questionnaire and mathematics test items.

Language background questionnaire. Each test booklet contained a two-page language background questionnaire (LBQ). The drafting of the LBQ was informed by a review of existing language background questionnaires, including the NAEP 1992 background questionnaire and the National Education Longitudinal Study (NELS: 88) background questionnaire. Items from previous questionnaires as well as new items were included, and comments from project staff were solicited and incorporated.

A draft LBQ was piloted at a middle school in the greater Los Angeles area, in two 8th-grade classes composed of students who were considered ready for transition from special classes for limited English proficient students into mainstream mathematics classes. The 29 students returning the questionnaire reported a variety of home languages including Spanish, Cambodian, Khmer and several others. All of the returned questionnaires indicated that the students spoke a language other than English, and most students indicated that their other language was spoken at home all or most of the time. The results of these data informed the revision of the LBQ. The LBQ is provided in Appendix I.

Mathematics test forms. From the 69 released 8th-grade NAEP items, 20 items were selected. These items were those judged most likely to impede a student's performance on a test because of language that could be misunderstood, could confuse the student, or could present difficulties that might distract the student's attention from the math content of the item. A simplified version of each of the items was written, following the procedure

[3] Different agencies use different categorial descriptors for ethnicity. The original descriptors from each agency have been retained in the table.

outlined above (see Identification of Linguistic Features section). In the revision process, the language was simplified, but the quantities, numerals, and visuals were retained from the original, so that the math content of the revised items paralleled that of the original items.

In order to ensure that the mathematical content of both versions of each item was equivalent, two experts in mathematics education independently reviewed each pair of items. They were asked to determine whether the two items differed in mathematical content or were equivalent with respect to the mathematical concepts being assessed. One math expert found no differences between the original and revised items in mathematical content; the other math expert pointed out three instances in which the situation in the revised item might be construed as slightly different. Changes were made in those three items to make certain the math content in each pair was parallel.

Two different forms of the mathematics test were created. Booklet A contained 10 original items; the revised versions of these items were placed in Booklet B. Ten additional original items were placed in Booklet B, and the revised version of these were placed in Booklet A. Thus, each form contained 10 original and 10 revised items. In addition, from the 69 NAEP items, 5 items were selected in which the language was judged to have the least potential for misunderstanding or confusion. These 5 items were included in both forms of the test to provide a check for the equivalence of the groups of students taking each of the two forms of the test. Thus, each test form contained a total of 25 math items. The original versions of the 20 released NAEP math items plus the 5 control items are provided in Appendix F. The revised versions of the 20 items also are provided in Appendix F. The test design is summarized in Table 43 (reproduced below).

Table 43

Design of Large-Scale Field Test

No. of items	Item type	Form A	Form B
10	Linguistically complex	Original	Revised
10	Linguistically complex	Revised	Original
5	Non-linguistically complex	Original	Original

Original test items were assigned to booklets according to four criteria: type and number of linguistic complexities, presence or absence of a diagram or other visual aid, mathematical classification of the item content according to NAEP categories, and difficulty of the item. The measure of item difficulty used was the item difficulty index (p value) of each item from an earlier NAEP administration for a sample consisting of 8th-grade students (1992 main assessment in math).

In creating the test booklets, a rough balance of all of the above criteria across the two booklet forms was sought, so that, for example, each booklet had an equal number of original items containing the passive voice, the same number of original problems dealing with algebra, and the same number of original items containing visual aids. The average difficulty of the original items in each booklet was roughly equal. We were not, however, attempting to produce parallel forms.

Items were randomly ordered within the test forms, with the same random ordering used for both booklets. For a small number of multiple-choice items, the order of response options was altered in both booklets to achieve an appropriate balance of correct responses (A, B, C, D, or E).

Procedure. Tests were administered by a team of ten retired teachers and principals experienced in test administration. Test administrators attended a half-day training session, and testing sites were monitored in random visits by members of the project staff. In each mathematics classroom, administrators distributed test booklets, alternating Booklet A with Booklet B; 50.9% received Booklet A, and 49.1% received Booklet B. Students were given approximately one hour to complete the test.

Results

In this section we first report descriptive findings from the administration of the language background questionnaire, then descriptive findings from the Accuracy Test administration with respect to overall performance levels, and, finally, research question results.

Language background questionnaire descriptive results. Each Accuracy Test booklet included a two-page language background questionnaire (LBQ). Student responses are summarized here; see Appendix K for details.

The LBQ, as part of the Accuracy Test booklet, was administered to a total of 1031 students, 61% of whom spoke a language other than English at home and/or with family members. Spanish was the principal second language, cited by 376 (or 60% of the non-native English speaking) students. A diverse group of languages was included. On several items in the LBQ, students were asked to report their use of and ability in the other language using a Likert scale. In general, students reported that they spoke their "other language" more often with their parents and grandparents than with their friends in school. They reported that their aural comprehension of the other language was superior to both their comprehension of written material in that language and production of that language.

Other Likert-type questions on the LBQ asked students to self-report their level of comprehension of content area materials (math, science, and social studies) in English, including understanding of teacher's explanations, textbooks, and questions on tests. Students in general reported that they had more difficulty understanding their teachers' explanations in math as compared to other content areas. Math textbooks were also reported as more difficult to understand than textbooks in other subject areas (compare Tables 69 through 74 in Appendix K). For test question comprehension, math was again reported as more difficult than other subject areas, but the difference was quite small.

Also, students were asked to self-report on their abilities in English. Each of the modalities (i.e., understanding, speaking, reading, writing) was questioned separately. Not surprisingly, students in the "beginning ESL" category (as determined by LAUSD) obtained the lowest mean scores.

Student responses on the above questions were grouped on a number of demographic and other background characteristics, including gender, ethnicity, type of math class, and participation in free lunch program. Complete results of these analyses are included in Appendix K. The results of analyses of variance generally indicate that the subgroups under each of the background variables performed about the same except for the ESL groups.

Following is a summary of descriptive statistics on students' math performance on the Accuracy Test and various background factors. In general, the results of this study were consistent with the literature and indicated that:

1. native speakers of English scored higher than non-native speakers;

2. the students in ESL programs performed at a lower level than the non-ESL students;
3. males and females performed at about the same level;
4. the performance of students in different types of math classes was significantly different;
5. there were differences in students' math performance with respect to ethnicity; and
6. generally, students in a lower category of SES performed less well in math than students in higher SES categories.

Analysis of student scores on the Accuracy Test were made on all 25 test items (10 original, 10 revised, and 5 control items). Analyses were also conducted for original and revised items separately; in both cases, the results were consistent with those reported for the total 25 items. A comparison of student scores for Booklets A and B on all 25 items, and on the five control items, which were the same in Booklet A and Booklet B, showed almost identical results. The mean score on all 25 items for the entire group was 15.43 and the standard deviation was 5.90.

The native English speaking group (as determined by the LBQ) had a higher mean score in math (M = 16.36, SD = 5.74) than the non-native English speaking students (M = 14.42, SD = 5.90), a difference of about a third of standard deviation, $F_{1,1015} = 28.20$, $p < .01$ (see Table 93 in Appendix L).

Differences were found between ESL and non-ESL students (based on ESL codes assigned by the schools); means ranged from 6.41 for beginning ESL students to l6.52 for non-ESL students (see Table 90, Appendix L).

Analysis of variance indicated no significant difference between male and female students on the Accuracy Test, $F_{1,1029} = 1.3497$, $p = .25$ (see Table 94, Appendix L for M, SD, and N). Analyses according to ethnic group are provided in Table 89, Appendix L.

As one might expect, students in higher level math classes received higher scores; the means of the two booklets ranged from 7.84 for the ESL math classes to 21.33 for honors algebra (high level) classes—a difference of approximately 2.3 standard deviations (see Table 91). Analysis of variance performed on the mean scores of the math class subgroups revealed a significant difference between the subgroups, $F_{5,1025} = 150.66$, $p < .01$ (see Table 91, Appendix L).

Differences were found according to ethnic group. White students had the highest mean score (*M* = 18.86, *SD* = 4.73), next were Asian American students (*M* = 18.49, *SD* = 4.70), next were African American students (*M* = 13.21, *SD* = 5.15) and next were Latino students (*M* = 12.52, *SD* = 5.41) (see Table 89, Appendix L). Analysis of variance showed a significant difference between performance of the ethnic groups, $F_{3,981} = 112.60, p < .01$.

Differences were found according to student SES level (see Table 92, Appendix L). A rough index of SES was devised from school information on participation in free lunch and AFDC programs, as discussed above in the Methodology section on Subjects. Mean scores ranged from 13.78 (*SD* = 5.35) for the "free lunch" group to 18.96 (*SD* = 3.71) for the "full payment" category, which differ by nearly one standard deviation. However, for the majority of students, the difference was not that great (see Table 92).

Accuracy Test Study research question results. The Accuracy Test field study addressed the impact of revision of selected linguistic features in math items on the number of items answered correctly by students. As indicated earlier, we, in many cases, placed the original form of an item in one booklet and its revised form in another booklet to avoid any problems due to answering the same items twice, such as transfer of learning. This, however, created another problem for us. That is, data for the original and revised items were not directly comparable because they were obtained on two different group of subjects. In order to compare student performances on original versus revised items, we had to make sure that the overall math performance of the students who answered the set of original items was not statistically different from that of students who answered the sister items. We compared the performance of students who took Booklet A with those who answered items in Booklet B. This comparison was made for all 25 math items. Table TOT1 (reproduced below) presents the results of this analysis. As Table TOT1 indicates, there was no significant difference between the performance of students who answered items in Booklet A and those who answered items in Booklet B on all 25 items ($t = .18$, $df = 1029$, $p = .857$). The results of the analysis in Table TOT1 also show no significant differences between the two groups of students on the 5 control test items ($t = .10$, $df = 1029$, $p = .919$). Similarly, the results indicate that the students' performances in the two groups are alike on the first set of items (original in A, revised in B; $t = -1.19$, $df = 1029$, $p = .235$). There was also no significant

Table TOT1
Across-Sampling Group Equivalency Statistics

	N	Mean	*SD*
Total score by booklet			
Booklet A	525	15.40	6.003
Booklet B	506	15.46	5.802
t = .18, df = 1029, p = .857			
Control/Nonproblematic items by group			
Group 1	525	3.32	1.292
Group 2	506	3.33	1.293
t = .10, df = 1029, p = .919			
Set 1 items by group			
Group A	525	5.70	2.804
Group B	506	5.90	2.604
t = -1.19, df = 1029, p = .235			
Set 2 items by group			
Group A	525	6.37	2.532
Group B	506	6.23	2.526
t = .90, df = 1029, p = .367			

difference between the two groups on the second set of items (original in B and revised in A; t = .90, df = 1029, p = .367). Thus, the results indicated that the two groups of students who answered items from two different booklets were from the same population and could be compared across booklets.

Using two sets of ten items each, we found that mean student scores (number correct) were greater for the revised items than for the original items in both cases—that is, the students did better on the simplified versions—but the results did not reach statistical significance. (For one set the revised mean score minus the original mean score was 6.371 – 6.229 = 0.142; for the other set the difference between the means was 5.905 – 5.705 = 0.200).

Because of the large variability in students' math scores among different schools and different levels of math class, a multilevel analysis approach was considered appropriate. In this analysis, a two-level model was employed. At the first level, students were the units of analysis; at the second level, math classes were the units of analysis using the HLM software (Bryk, Raudenbush, Seltzer, & Congdon, 1989).

Because our version of HLM software performs only univariate analysis, two separate analyses were performed. In the first analysis (Model 1), the dependent variable was the first set of 10 items (original in Booklet A and revised in Booklet B), which is called SET1. The results of analyses for the SET1 variable are summarized in Table HLM1 below. In the second analysis (Model 2), the dependent variable was the second set of 10 items (original in Booklet B and revised in Booklet A) called SET2. The results for this model are presented in Table HLM2 below. Results of these two analyses are very similar, and lead to the same general conclusion, which will be discussed later in this section. In each analysis, the booklet (variable BOOK) was considered the "treatment" effect; a statistically significant BOOK effect would indicate a nontrivial effect of item revision.

The general strategy for our HLM analysis was to begin with a relatively simple model, examine the resulting statistics, then examine more complex models with additional variables added as appropriate. However, for neither dependent variable was a statistically significant booklet effect found, and analyses concluded rather early.

As Table HLM1 shows, the mean for the items in revised form (M = 5.905, SD = 2.604) is higher than the mean for the original items (M = 5.705, SD = 2.804), as anticipated, but the gain was about 0.2 points out of 10 which is very small. In this model the Level 1 predictor was booklet. The results of analyses for Model 1 as shown in Table HLM1 indicate that the grand mean is significantly different from zero (t = 16.9, $p < .01$) but the effect of BOOK on SET1 is not (t = -1.5, $p > .05$). There is still unexplained variance of the means of SET1 among classes (t_{00} = 4.117, χ^2 = 1347.9, $p < .01$). However, the relationship between BOOK and SET1 (which, we have seen, is virtually zero, t_{11} = .021, χ^2 = 37.1, $p > .05$) does not vary among classes.

Similarly, in Model 2 the Level 1 predictor was booklet. The results of analyses for this model as presented in Table HLM2 indicate that the mean for the items in revised form (M = 6.371, SD = 2.532) is higher than the mean for items in original form (M = 6.229, SD = 2.604), as anticipated, but the gain was small: 0.15 points out of 10.

Table HLM1

Two-Level Hierarchical Linear Model (Model 1), DV = SET1

	M	*SD*	Cases
	Statistics		
Booklet A	5.705	2.804	525
Booklet B	5.905	2.604	506
	Model parameters		
g_{00}	5.5812	t = 16.9[a]	
g_{10}	- .0916	t = -1.5, nonsig.	
$var(r_{ij}) = s^2$	3.2316		
$var(u_{0j}) = t_{00}$	4.1170	$\chi^2(38) = 1347.9$[a]	
$var(u_{1j}) = t_{11}$	.0210	$\chi^2(38) = 37.1$, nonsig.	
Total valid cases: 1031			

Note. SET1 is the composite of 10 items that were in original form in Booklet A and revised form in Booklet B.
[a] $p < .01$.

Table HLM2

Two-Level Hierarchical Linear Model (Model 2), DV = SET2

	M	*SD*	Cases
	Statistics		
Booklet A	6.371	2.532	525
Booklet B	6.229	2.526	506
	Model parameters		
g_{00}	6.1088	t = 21.2[a]	
g_{10}	.0939	t = 1.5, nonsig.	
$var(r_{ij}) = s^2$	3.3915		
$var(u_{0j}) = t_{00}$	3.0998	$\chi^2(38) = 943.8$[a]	
$var(u_{1j}) = t_{11}$	.0122	$\chi^2(38) = 31.5$, nonsig.	
Total valid cases: 1031			

Note. SET2 is the composite of 10 items that were in revised form in Booklet A and original form in Booklet B.
[a] p < .01.

The results in Table HLM2 indicate that the grand mean is significantly different from zero (t = 21.1, p < .01), but the effect of BOOK on SET2 is not (t = 1.5, p > .05). There is still unexplained variance of the means of SET2 among classes (t_{00} = 3.10, χ^2 = 943.8, p < .01). However, the relationship between BOOK and SET2 does not vary among classes (t_{11} = .0122, χ^2 = 31.5, p > .05).

Both analyses lead to the same general conclusion: A comparison of the intercepts (means) of the original and revised items revealed that for most of the classes (Level 2 unit) intercepts for the revised items were higher than the intercepts for the original items. However, there is no significant overall booklet (i.e., treatment) effect on SET1 and SET2 (the two sets of items). In other words, item revision is not a significant treatment effect. Furthermore, the booklet effect does not vary significantly among classes—in other words, there is no variability in this effect that could be explained by additional class-level variables. (There is still considerable interclass variability in mean scores that could be investigated by additional analyses, but this is not the aim of the present study.)

Discussion of Accuracy Test Study Results

For the total student sample, the improvement on the total set of revised items was not significant, and consequently it would be inappropriate to claim significant results for subsamples of students or items. However, it is interesting to note that, for certain subgroups of students and items, tests showed occasional significance that was consistent with findings in previous studies reported in the literature. Performance by certain student groups and results on subsets of items are discussed here.

Overall results for the total student group are shown in Figures 1, 2, and 3, comparing students' performance on original and revised items.

Impact of item revision on students at different achievement levels. Certain groups of students benefited more from the item revisions than others did. Larsen, Parker, and Trenholme (1978) tested 8th graders on math items with high, moderate, and low levels of syntactic complexity but equal mathematical difficulty; they found that, for the more complex items, scores were significantly lower for the low-achieving 8th graders. To see whether our findings were consistent with those of Larsen et al. (1978), we classified the students in terms of achievement in mathematics. Assuming that a student's

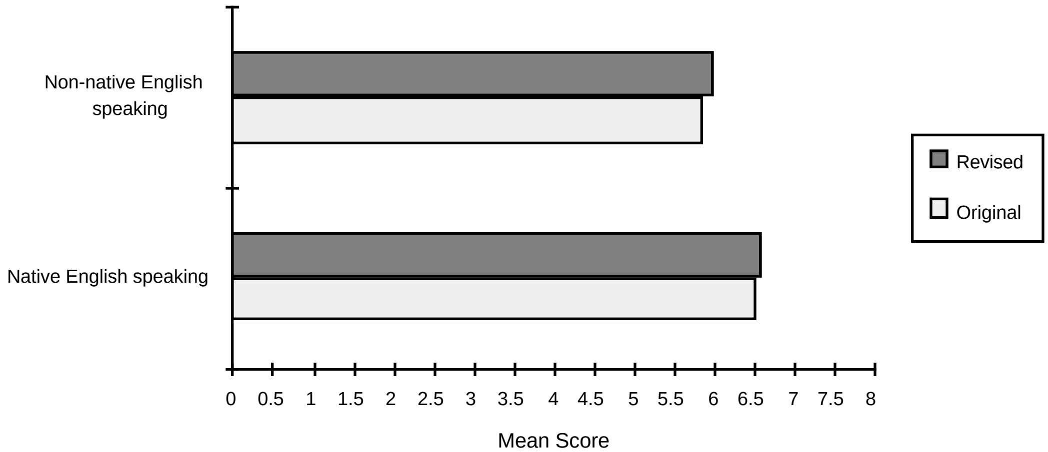


Figure 1. Comparing students' performance on original and revised items for the first set of 10 items.

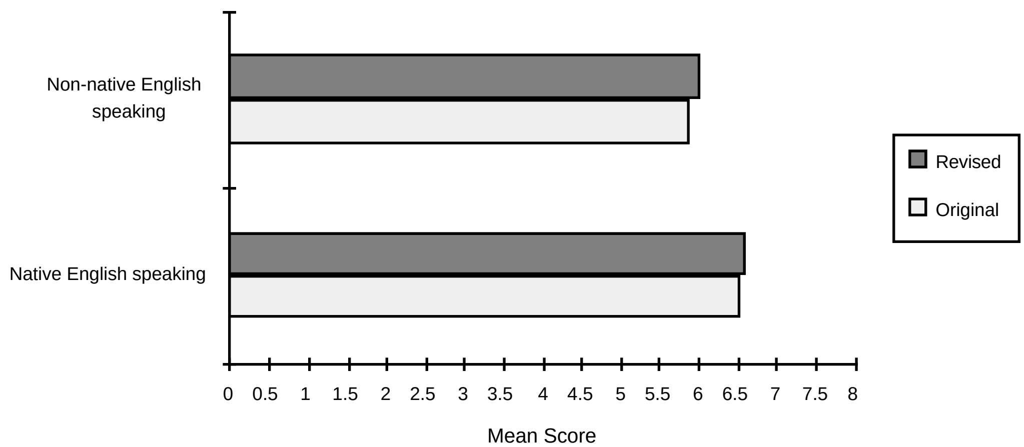


Figure 2. Comparing students' performance on original and revised items for the second set of 10 items.

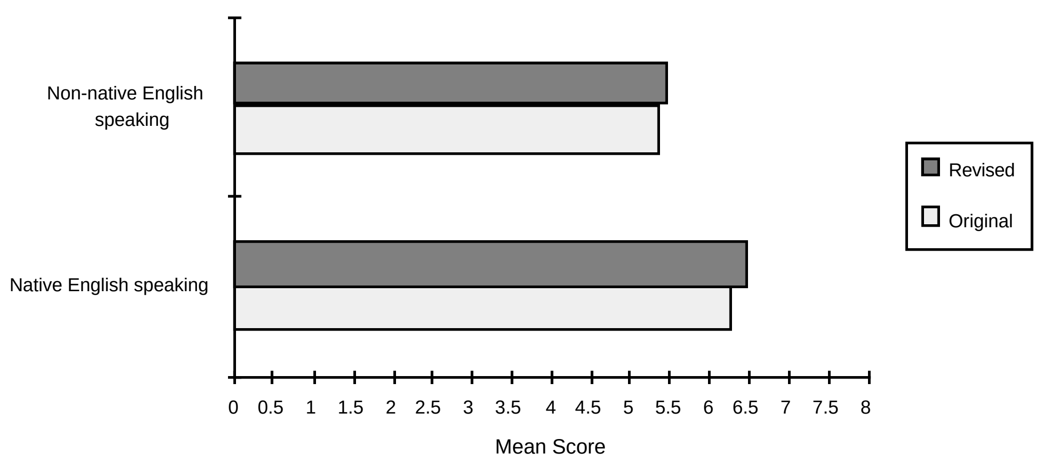


Figure 3. Comparing students' performance on original and revised items for all 20 items.

prior achievement in math was reflected in his/her current level of math class, we separated the students according to type of math class (honors algebra, algebra, high math, average math, low math, remedial/basic, and ESL math). We compared mean scores on original and revised items for students in each type of math class; typically, the mean score for revised items was higher. We then calculated the difference in scores attributable to the language simplifications as a percentage of improvement over the original score (that is, we found the difference by subtracting the mean for original items from the mean for revised items, and then divided the difference by the mean for original items).

Simplifying the language had a differential impact on performance. Students in the lowest categories of math class (ESL) showed slight improvement in their math performance on the revised items. In the next category of math class (remedial/basic), students exhibited more improvement, and even greater improvement was found for the next categories (low and average math classes). The trend did not continue for higher levels of math classes, however; in fact, for the honors algebra class the language simplifications had a small negative effect. Percentages of improvement were 4.9% for the ESL classes, 10.4% for the low math classes, 7.1% for the average classes, and then 0.5%, 0.1% and -0.8% for the high math, algebra, and honors algebra classes.

These differences indicate that the language simplifications had greater impact for students in low and average math classes. Since language ability is, in general, a predictor of math performance, it is possible that the language simplifications had little effect on the algebra and honors students' performance because these high-performing students also had strong language ability and had no problem understanding the original items. Although the original items were longer and more complex linguistically, they did not slow down the top students. If the students in low and average math classes had correspondingly low or average language comprehension skills, the small changes in the revised items could well have led to greater comprehension and relatively greater improvement in their scores.

The differences observed here are consistent with previous research studies showing relationships between reading ability and arithmetic problem solving ability (Aiken, 1971, 1972; Noonan, 1990). They are also consistent with the view that inexperienced problem solvers, lacking highly developed semantic schemata for problem solving, rely more on the text (De Corte et al., 1985); if this is indeed the case, we would expect that the complexity of the text would be a more significant factor for inexperienced problem solvers. Our results support this view.

Impact of changes in specific linguistic features. In some instances, revisions of unfamiliar/infrequent vocabulary and passive voice structures resulted in better student performance. These results are consistent with previous studies and point to the need for a closer examination of the difficulties these features pose for students.

In revising each math item, we typically made more than one change in the wording of the original item; each change was classifiable as one of the seven types discussed above (see the section Identification of Linguistic Features). Among the most frequent types of changes were simplification of unfamiliar/infrequent vocabulary and rephrasing of passive voice constructions. Both of these linguistic features have been addressed in previous studies (as discussed above), and both figure prominently in previous discussions of readability and linguistic complexity. To determine the extent to which simplifications of these types affected student performance, we identified the 6 items with substantial vocabulary simplifications and the 11 items with passive voice construction revisions, and compared original and revised item scores. For the total student group, the scores on items with vocabulary simplifications were significantly better ($t = 2.54$, $df = 1029$, $p < .05$) than on the parallel original items for a 4-item group in one of the 10-item sets (revised-item mean 2.389, original-item mean 2.210). For another item with vocabulary simplification (from the other 10-item set), none of the students in low math classes answered the original item correctly, but 15% answered the revised version correctly. (Interestingly, for the same item, honors algebra students did slightly better on the original version, although that result did not reach significance.)

For items with passive voice changes, the score differences were not significant for the total student group. However, in one of the 10-item sets, the 6 items with passive voice revisions did show significantly higher scores on the revised versions for students in average math classes (original-item mean 2.705, revised-item mean 3.149; $t = -3.03$, $df = 403$, $p < .01$).

In the items on the Accuracy Test, the number of changes in any single linguistic feature type depended on the number of times that feature occurred in the test set of 20 NAEP items. For some of the feature types, there were only a few instances in the corpus; consequently, there were not enough instances to tease apart the relative influences of each type of change. We could begin to assess the impact of unfamiliar/infrequent vocabulary and passive voice changes because these two features occurred more frequently in the corpus.

Speed Test Study

The purpose of the third field research study was to examine the effects of item linguistic complexity on the time it took students to read, understand and answer the test items.

Method

The general methods used for this study are discussed in the previous Methods section.

Subjects. A total of 143 students from two schools in the greater Los Angeles area were tested in the Speed Test Study. Most of the students (82.5%) were 8th graders; 13.3% were in Grade 7, and 4.2% were in Grade 9. Of the total group, 110 (77%) indicated that they spoke a language other than English. Of this number, 103 (94%) indicated that the other language was their first language (see Table 38, Appendix G). Table 40 (Appendix G) shows distributions of student and level of math classes. As Table 40 indicates, most (67.8%) were in ESL mathematics classes, but some were in high math (17.5%) and algebra classes (14.7%). Gender is shown in Table 36 (Appendix G); 48% of subjects for this part of study were male.

Ethnicity is shown in Table 37 (Appendix G): 67.8% were Latinos, 19.6% were African American, 6.3% were White, 4.9% were Asian American, and for 1.4% information on ethnicity was unavailable.

Instruments. For the Speed Test Study, two new test booklets were prepared. Booklet A contained the 20 original NAEP items from the Accuracy Test Study, and Booklet B contained the 20 revised items with simplified language; the mathematics content was not revised and the five control items were not used. Each test booklet contained the language background questionnaire, as in the Accuracy Test (described above).

Procedure. Two test administrators attended a second training session where purposes and procedures were reviewed and a practice administration was held. Students were given 10 minutes to answer 20 mathematics items. For the Speed Test, 76 (53.1%) students received Booklet A and 67 (46.9%) received Booklet B.

Results

Here we first report descriptive analyses of the language background questionnaire responses and descriptive analyses of the students' math performance, followed by research question results.

Descriptive analyses of language background questionnaire responses. For the 143 students in the Speed Test, the descriptive analysis of responses to the LBQ produced results quite similar to those for the Accuracy Test for most of the descriptive categories (see Appendix K, discussed above, for frequencies and percentages for the various language background questions in the LBQ for both Accuracy Test and Speed Test portions of the field research phase of the study). We found some differences between the Accuracy Test and Speed Test groups, however. Generally, students in the Speed Test reported a lower overall proficiency in their other language and a slightly lower level of understanding of teachers' explanations, textbooks, and question items in subject content areas. For the Speed Test, students in beginning ESL classes reported understanding math textbooks better than science and social science books. There were major differences between different ESL groups on other content area comprehension questions (see Appendix K).

Analyses of Speed Test performance results. Student performance on the Speed Test was analyzed with respect to a number of background variables (using data obtained from school personnel, as in the Accuracy Test), including (a) native vs. non-native speakers, (b) ESL classification, (c) gender, (d) mathematics class level, (e) ethnicity, and (f) SES. These background variables were analyzed with respect to number of items correct on Booklet A (original items), Table 107; number of items correct on Booklet B (revised items), Table 108; number attempted on Booklet A, Table 109 and; number attempted on Booklet B, Table 110 (see Appendix M for Tables 107-110). Most differences did not show statistical significance, often because of small sample size. Some significant results were found, however; native speakers of English got more items correct on Booklet A (original items) than non-native speakers (t = 1.96, df = 73, $p < .05$). The number correct on Booklet A also varied according to the student's ESL category, with more advanced students answering more items correctly. And, as might be expected, the number of items answered correctly varied according to the level of the student's math class placement.

Research results for Speed Test Study. Roughly half the students answered Booklet A, original NAEP math items, and the others answered Booklet B, parallel items with language simplified. For both groups of subjects we obtained number of items attempted and number of items answered correctly. There were higher rates of response on the revised items. These improvements were more evident for the language minority students. Unfortunately, the small number of students in this part of the study precluded any in-depth analysis. Means, standard deviations, and number of cases were obtained for number of items attempted and number of correct responses. For the Speed Test, the mean number of original items attempted was 9.24, and the mean number of revised items attempted was 9.54, for a difference of .3, which was not significant (t = -.60, df = 141, $p > .05$).

Discussion of Results of Speed Test Study

The number of cases across the categories was very small for many of the subgroups. This was particularly true for the ESL categories. There were also some unexpected findings on this part of study—for example, the results indicated some gender differences. As described above for the Accuracy Test, male students tended to benefit more from the revisions than female students; such results were not part of our initial focus. Although, as mentioned above, we did not gather a verbal skills measure for the subjects, if the boys were indeed slightly below the girls in verbal comprehension or fluency, it would help to explain why simplifying the language enabled the boys to finish more problems, with a consequently slightly greater effect of revisions on the boys' scores than on the girls' scores.

GENERAL DISCUSSION

In this section we discuss some of the implications of our findings, and we note problems encountered in this study.

The Importance of Language

The analyses of existing data showed some effects of student language background on mathematics test performance. The performance of students who spoke a language other than English at home was significantly lower than the performance of students who spoke only English at home; when items were categorized by length, the difference was more evident for the long items than the short items. Additionally, the number of omitted/not-reached math items was higher for students who spoke a language other than English at home. These results clearly indicate the confounding of language and performance. A lack of familiarity with mathematical terminology will limit a student's test performance; in addition, general language proficiency is required for reading test items and formulating written responses, as well as reading textbooks and understanding teachers' explanations. General language proficiency and knowledge of the specialized language of mathematics are both important; a deficiency in either one constitutes a burden for the student and can negatively impact his/her individual performance.

In Phase 1 of this study we compared students' performance on 1992 NAEP items by classifying the items with respect to linguistic complexity. The performance of language minority students was lower than that of other students, and the difference was greater for the more complex items. In the field study, we analyzed student performance with respect to self-reported background data on language background and with respect to school ESL classifications; we found that students in the ESL categories showed considerably lower math performance than other students. This is a cause for concern; it requires special attention. There do not seem to be major differences between these low-performing ESL students and other groups of students with respect to socioeconomic status or other variables that could explain such differences. Therefore, one must conclude that language is a very important element in such cases—that is, language and performance are confounded. The exact nature of the confounding factors remains elusive.

In this study we did not have all the necessary ingredients to "unconfound" or fully explain the differences in students' math performance across categories of language-related variables. Among the major problems we encountered was the limitation on the number and types of items available to us, but even more important was the degree of complexity involved in categorizing or even typifying language minority students. We want to bring to the attention of policy makers this very important issue, one which may affect any study dealing with language minority students. The lack of an operational and commonly accepted definition of language minority and/or ESL students in our schools is a major obstacle for any analysis of language minority students.

In Phase 2 of the study, we compared student performance on original math items and comparable items with simplified language. We found that students in low and average math classes benefited the most from the revisions, with scores on revised items showing improvements of 7% to 10% over scores on original items. When these students are confronted by a math problem, complexities in the language may constitute an additional obstacle, adding to the cognitive burden of dealing with the problem statement. For these students, simplifying the language may ease the cognitive load just enough to result in an improvement on their test scores. Revisions in the language of the test items did not appear to affect the performance of students in the highest level of math classes, however. These students typically have a good command of both language and mathematics, and for them the complexity of the language is apparently not a factor in solving the problem.

Analysis of student responses to the language background questionnaire showed a range of findings. Students reported that they have more problems understanding teachers' explanations, textbooks, and tests in the area of math than in science or social studies. The most apparent difference between groups of students on their self-reported level of English proficiency (understanding, speaking, reading, and writing), as well as on their understanding of teachers' explanations, text books, and exams, was with respect to ESL classifications.

The results of our analyses also showed significant differences in students' performances across categories of ethnicity, school lunch program and variables not directly related to language. The most noticeable of these differences, however, was across categories of type of math class. When variability due to the type of math class was controlled, there was very little variability left to warrant further attention.

Comparisons of the results of HLM analyses on the original versus revised items (comparing intercepts and slopes of the two models using original and revised scores as outcome variables) revealed that even with the small set of items, the revisions showed changes in students' performance. There were some interesting trends found from the results of HLM analyses. For the models with the revised items as outcome variables, the language-related variables were shown to be more effective than the model with original item score as the outcome variable. However, none of these trends reached statistical significance.

Problems Encountered

In carrying out this study, we encountered problems due to the limited number of mathematics items available to us and the difficulty of obtaining valid measures of students' English proficiency.

The most critical problem encountered during this project was the limited number of NAEP items available for the project team to work with. The project staff was given access to 69 released items. This relatively low number of items was a significant problem because it limited the types and number of linguistically complex items and the range of linguistic features that could be used in the study. In our field studies, each form of the test contained 10 revised items for analysis. However, a 10-item test is not amenable to subscale analyses. Furthermore, some of these 10 items tended to show extreme *p* values, indicating that these items were either too hard or too easy for subscale comparisons with their original counterparts. If we had had access to a larger pool of items, we could have avoided using items with extreme *p* values. Access to a larger pool of items might have significantly affected the findings of this study.

The limited number of items precluded subscale analyses. The five NAEP math content area subscales are numbers and operations; measurement; geometry; data analysis, statistics, and probability; and algebra and functions. For the items available for our study, the distribution across the five content areas was not proportional to that of the NAEP math test; for example, of the 69 items supplied, only two were from the geometry area, and one third were from the measurement area.

Another problem encountered in this study was the difficulty of obtaining valid measures of students' English proficiency. Limited English proficiency (LEP) classifications are not uniform from school to school, due in part to a lack of effective language proficiency measures for K-12 students. Consequently, students' LEP classifications at one school may be different from those at another school. In addition, the school district's information about the language backgrounds of students may be incomplete, outdated, or invalid. With this in mind, we formulated the language background questionnaire used in this study. However, self-reported information on English proficiency may not always be reliable, particularly if the survey instrument is a document written in English. Better access to more accurate language background information is needed if we are to draw valid conclusions from studies like this one, and if we are to have confidence in the results of our performance assessment procedures.

While the current NAEP policy is to exclude students in ESL and Bilingual Sheltered programs, it is possible that this is not being achieved in all cases. In regions other than the greater Los Angeles area, some schools may not be able to accurately classify students and provide appropriate programming for those who need it. It may be that such students are inadvertently being included in NAEP assessments.

CONCLUSION

The results from this study were mixed. Analyses of existing NAEP data and results of the Student Perceptions Study revealed significant effects of language background on performance. For the Accuracy Test Study and the Speed Test Study, results showed no significant differences. Results of overall HLM analyses did not show a significant effect for the linguistically simplified items. Nonetheless, it is worth noting that for one major subgroup, comprising nearly half the sample, significant differences were found. In those mathematics classes considered average and below average, students performed significantly better on revised over original items.

The lack of statistically significant improvement overall in the Accuracy and Speed Test Studies was due in part, we feel, to the limitations discussed in the previous section. Most important perhaps was the lack of consistency in ESL/LEP designations in the LAUSD school system; this situation may well be generalizable to a number of urban and suburban school systems. From the results, it should not be inferred that language complexity is irrelevant for assessment in subject areas such as mathematics. Language is a complex system, and studies such as this one can help us identify linguistic features that can affect performance for some students, thus enabling us to improve the validity of our assessments.

The precise nature of the interaction between linguistic dimensions and other background variables is complex and warrants further research. Before such research begins, we suggest more immediate priorities for investigation. As described in the previous section, problems exist in the definition and assignment of ESL/LEP categories. The significance of this problem cannot be overstated. There is a real need for research in this area. It is possible that NAEP, despite its best efforts, is testing students whose first language is not English and whose English language comprehension/production is weak. Because of these problems, NAEP would benefit from a study examining how effectively the policy is being implemented nationwide.

Other areas in which research might be fruitful would be the replication of the kind of study done here on a larger scale, with greater access to NAEP items and in a larger range of communities. Since NAEP is committed to employing more open-ended and extended-open-ended questions as a format, the linguistic issues confronted in this study will presumably become more critical in the future. For instance, the training of raters for open-ended items will need to include awareness of how to recognize ESL/LEP errors and to distinguish errors in general language from errors in mathematics content.

Although the portion of this study that dealt with the identification of complex language was largely exploratory in nature, it provided useful clues in the search for linguistic features that can negatively affect performance for certain groups of students. Data from this study were consistent with previous research suggesting that unfamiliar/infrequent vocabulary and passive voice constructions may affect comprehension for certain groups of students, and that average and low-achieving students may be at a relatively greater disadvantage in answering mathematics items with complex language. These studies should be replicated and refined. It is also possible that future studies, with larger numbers of other targeted linguistic features such as those described in this study, will reveal similar effects. Meanwhile, it remains prudent to continue searching for interactions among linguistic, socioeconomic and other background variables to shed light upon the growing issue of the role of language in content area assessment.

Ultimately, this study shows that the interaction between language and mathematics achievement is real. This interaction must be a critical consideration in future mathematics assessment research and practice.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Aiken, L. R. (1971). Verbal factors and mathematics learning: A review of research. *Journal for Research in Mathematics Education, 2*, 304-113.

Aiken, L. R. (1972). Language factors in learning mathematics. *Review of Education Research, 42*, 359-385.

Baker, E. (1991). *Issues in policy, assessment, and equity*. Washington, DC: Department of Education, Office of Bilingual Education and Minority Languages Affairs.

Baugh, J. (1988, August). [Review of *Twice as less: Black English and the performance of Black students in mathematics and science*]. *Harvard Educational Review*, *58*, 395-404.

Bever, T. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-353). New York: John Wiley.

Biber, D. (1988). *Variation across speech and writing*. New York: Cambridge University Press.

Bormuth, J. R. (1966). Readability: A new approach. *Reading Research Quarterly*, *1*, 79-132.

Botel, M., & Granowsky, A. (1974). A formula for measuring syntactic complexity: A directional effort. *Elementary English*, *1*, 513-516.

Bryk, A. S., Raudenbush, S. W., Seltzer, M., & Congdon, R. T. (1989). *An introduction to HLM: Computer program and users' guide*. Fairplay, CO: Scientific Software.

Burger, J. M. (1992). Desire for control and academic performance. Special issue: The psychology of control. *Canadian Journal of Behavioral Science*, *24*,147-155.

Burstein, L. (1993). *TRP investigations of validity of NAEP measures*. Paper presented at the TRP/NCES meeting, Washington, DC, July 14, 1993. (Available from University of California, Los Angeles, Center for the Study of Evaluation, Graduate School of Education and Information Studies.)

Byrne, B. M. (1984). The general/academic self-concept nomological network: A review of construct validation research. *Review of Educational Research*, *54*, 427-456.

Carpenter, T. P., Corbitt, M. K., Kepner, H. S., Jr., Linquist, M. M., & Reys, R. E. (1980). Solving verbal problems: Results and implications from national assessment. *Arithmetic Teacher*, *28*, 8-12.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage word frequency list*. Boston, MA: Houghton Mifflin.

Celce-Murcia, M., & Larsen-Freeman, D. (1983). *The grammar book: An ESL/EFL teacher's book*. Rowley, MA: Newbury House.

Chadha, N. K. (1989). Causal antecedents of self-concept, locus of control and academic achievement: A path analysis. *Psychologia, An International Journal of Psychology in the Orient*, *32*, 265-268.

Chall, J. S., Jacobs, V. S., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind.* Cambridge, MA: Harvard University Press.

Clement, J., Lochhead, J., & Monk, G. (1981). Translation difficulties in learning mathematics. *American Mathematical Monthly, 88,* 286-290.

Cocking, R. R., & Chipman, S. (1988). Conceptual issues related to mathematics achievement of language minority children. In R. R. Cocking & J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics* (pp. 17-46). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cocking, R. R., & Mestre, J. P. (Eds.). (1988). *Linguistic and cultural influences on learning mathematics*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cone, A. L., & Owens, S. K. (1991). Academic and locus of control enhancement in a freshman study skills and college adjustment course. *Psychological Reports*, *68*, 1211-1217.

Cossio, M. (1977). The effects of language on mathematics placement scores in metropolitan colleges. *Dissertation Abstracts International*, *38*(7), 4002A.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, *20*, 405-438.

Cummins, J. (1984). *Bilingualism and special education*. San Diego, CA: College Hill Press.

Dale, E., & Chall, J. S. (1948). A formula for predicting readability. *Educational Research Bulletin, 27,* 11-20; *28*, 37-54.

Davison, D. M., & Schindler, S. E. (1988). Mathematics and the Indian student. In J. Reyhner (Ed.), *Teaching the Indian child: A bilingual/multicultural approach*. Billings, MT: Bilingual Education Program.

De Corte, E., Verschaffel, L., & De Win, L. (1985). Influence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology*, *77*, 460-470.

deVilliers, J., & deVilliers, P. (1973). Development of the use of word order in comprehension. *Journal of Psychological Research*, *2*, 331-341.

Dunn, L. M. (1959). *Peabody Picture Vocabulary Test*. Circle Pine, MN: American Guidance Service.

Finegan, E. (1978, December). *The significance of syntactic arrangement for readability* . Paper presented to the Linguistic Society of America, Boston, MA.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*, 221-233.

Flynn, T. M. (1991). Achievement, self-concept and locus of control in Black pre-kindergarten children. *Early Child Development and Care*, *74*, 135-139.

Forster, K. I., & Olbrei, I. (1973). Semantic heuristics and syntactic trial. *Cognition*, *2*, 319-347.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, MA: Houghton Mifflin.

Freeman, G. G. (1978). Interdisciplinary evaluation of children's primary language skills. (ERIC Microfiche, ED157341)

Fry, E. (1977). Fry's Readability Graph: Clarification, validity, and extension to level 17. *Journal of Reading, 21,* 242-252.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ginsburg, H. (1981). The clinical interview in psychological research on mathematical thinking: Aims, rationales, techniques. *For the Learning of Mathematics*, *1*(3), 4-11

Hagborg, W. J. (1991). Group counseling with adolescent special education students: Challenges and useful procedures. *Adolescence*, *26*, 557-563.

Haiman, J. (1985). *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.

Hall, W. S., Nagy, W. E., & Linn, R. (1984). *Spoken words: Effects of situation and social group on oral word usage and frequency*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Halliday, M. A. K., & Martin, J. R. (1993.) *Writing science: Literacy and discursive power*. Pittsburgh, PA: University of Pittsburgh Press.

Harris, A. J., & Jacobson, M. D. (1973-1974). Some comparisons between the Basic Elementary Reading Vocabularies and other word lists. *Reading Research Quarterly, 9*, 87-109.

Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development*, *54*, 84-90.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (Research Rep. No. 3). Urbana, IL: National Council of Teachers of English.

Hunt, K. W. (1977). Early blooming and late blooming syntactic structures. In C. R. Cooper & L. Odell (Eds.), *Evaluating writing: Describing, measuring, judging*. Urbana, IL: National Council of Teachers of English.

Jerman, M., & Rees, R. (1972). Predicting the relative difficulty of verbal arithmetic problems. *Educational Studies in Mathematics*, *4*, 306-323.

Jones, P. L. (1982). Learning mathematics in a second language: A problem with more and less. *Educational Studies in Mathematics, 13*, 269-87.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixation to comprehension. *Psychological Review*, *87*, 329-354.

Kane, R. (1968). The readability of mathematical English. *Journal of Research in Science Teaching*, *5*, 296-298,

Kane, R. (1970). The readability of mathematics textbooks revisited. *The Mathematics Teacher*, *63*.

Kaput, L., & Clement, J. (1979). Letter to the editor. *The Journal of Children's Mathematical Behavior*, *2*, 208.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*, 580-602.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*, 109-129.

Klare, G. R. (1974). Assessing readability. *Reading Research Quarterly*, *10*, 62-102.

Klein, W. (1986). *Second language acquisition*. Cambridge: Cambridge University Press.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day English*. Providence, RI: Brown University Press.

Laffoon, K. S., Jenkins-Friedman, R., & Tollefson, N. (1989). Causal attributions of underachieving gifted, achieving gifted, and nongifted students. *Journal for the Education of the Gifted*, *13*, 4-21

Larsen, S. C., Parker, R. M., & Trenholme, B. (1978). The effects of syntactic complexity upon arithmetic performance. *Educational Studies in Mathematics, 21,* 83-90.

Learning Research and Development Center, and the National Center on Education and the Economy. (1992). *The New Standards Project, 1992-1995: A proposal*. Pittsburgh, PA: University of Pittsburgh.

Lemke, J. L. (1986). *Using language in classrooms*. Victoria, Australia: Deakin University Press.

Lepik, M. (1990). Algebraic word problems: Role of linguistic and structural variables. *Educational Studies in Mathematics*, *21*, 83-90.

Lochhead, J. (1980). *Toward a scientific practice of mathematics education*. (ERIC Document #294747)

Lochhead, J., & Mestre, J. (1988). The language of algebra. *1988 Yearbook of the National Council of Teachers of Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Lorge, I. (1939). Predicting reading difficulty of selections for children. *Elementary English Review*, *16*, 229-233.

Lucas, C., & Borders, D. G. (1994). *Language diversity and classroom discourse*. Norwood, NJ: Ablex.

Luthar, S. S., & Zigler, E. (1992). Intelligence and social competence among high-risk adolescents. *Development and Psychopathology, 4*, 287-299.

Lyon, M. A., & MacDonald, N. T. (1990). Academic self-concept as a predictor of achievement for a sample of elementary school students. *Psychological Reports*, *66*, 1135-1142.

MacDonald, M. C. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, *32*, 692-715.

MacGinitie, W. H., & Tretiak, R. (1971). Sentence depth measures as predictors of reading difficulty. *Reading Research Quarterly*, *6*, 364-377.

Macnamara, J. (1966). *Bilingualism in primary education*. Edinburgh: Edinburgh University Press.

Maqsud, M., & Rouhani, S. (1991). Relationships between socioeconomic status, locus of control, self-concept, and academic achievement of Batswana adolescents. *Journal of Youth and Adolescence*, *20*, 107-114.

Mann, W. C., & Thompson, S. A. (1986.) Relational propositions in discourse. *Discourse processes*, *9*, 57-90.

Marsh, H. W., Walker, R., & Debus, R. (1991). Subject specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology*, *16*, 331-345.

McNeill, D. (1970). *The acquisition of language: The study of developmental psycholinguistics*. New York: Harper & Row.

Mestre, J. P. (1988). The role of language comprehension in mathematics and problem solving. In R. R. Cocking & J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics* (pp. 201-220). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mestre, J., & Lochhead, J. (1983). The variable-reversal error among five cultured groups. In J. Bergeron & N. Herscovics (Eds.), *Proceedings of the Fifth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 180-188). Montreal: International Group for the Psychology of Mathematics Education.

Munro, J. (1979). Language abilities and math performance. *Reading Teacher*, *32*, 900-915

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

Noonan, J. (1990). Readability problems presented by mathematics text. *Early Child Development and Care*, *54*, 57-81.

Orr, E. W. (1987). *Twice as less: Black English and the performance of Black students in mathematics and science.* New York: W. W. Norton.

Paul, D. J., Nibbeling, W. H., & Hoover, H. D. (1986, May). The effects of adjusting readability on the difficulty of mathematics story problems. *Journal for Research in Mathematics Education, 17*, 163-171.

Pauley, A., & Syder, F. H. (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics*, *7*, 551-579.

Perera, K. (1980). The assessment of linguistic difficulty in reading material. *Educational Review*, *32*, 151-161.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.

Reich, S. S., & Ruth, P. (1982). Wason's selection task: Verification, falsification and matching. *British Journal of Psychology*, *73*, 395-405.

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The* *development of mathematical thinking* (pp. 153-196). New York: Academic Press.

Roseberry, A. S., Warren, B., & Conant, F. R. (1992). Appropriating scientific discourse: Findings from language minority classrooms. *The Journal of the Learning Sciences*, *2*, 61-94.

Rosnick, P. (1981). Some misconceptions concerning the concept of variable. *The Mathematics Teacher*, *8*, 50-56.

Rosnick, P., & Clement, J. (1980). Learning without understanding: The effect of tutoring strategies on algebra misconception. *Journal of Mathematical Behavior*, *3*, 3-27.

Rothman, R. W., & Cohen, J. (1989). The language of math needs to be taught. *Academic Therapy, 25*, 133-42.

Savin, H. B., & Perchonock, E. (1965). Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior*, *4*, 348-353.

Saxe, G. B. (1988). Linking language with mathematics achievement: Problems and prospects. In R. R. Cocking & J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics* (pp. 47-62). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schachter, P. (1983.) *On syntactic categories*. Bloomington: Indiana University Linguistics Club.

Shuard, H., & Rothery, A., (Eds.) (1984). *Children reading mathematics*. London: J. Murray.

Slobin, D. I. (1968). Recall of full and truncated passive sentences in connected discourse. *Journal of Verbal Learning and Verbal Behavior*, *7*, 876-881.

Spache, G. D. (1953). A new readability formula for primary-grade reading materials. *Elementary School Journal*, *53*, 410-413.

Spanos, G., Rhodes, N. C., Dale, T. C., & Crandall, J. (1988). Linguistic features of mathematical problem solving: Insights and applications. In R. R. Cocking & J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics* (pp. 221-240). Hillsdale, NJ: Lawrence Erlbaum Associates.

Spencer, P. L., & Russell, D. (1960). Reading in arithmetic. In F. E. Grossnickle (Ed.), *Instruction in arithmetic*. *Twenty-Fifth Yearbook of the National Council of Teachers of Mathematics* (pp. 202-223). Washington, DC: National Council of Teachers of Mathematics.

Sue, D. W., & Sue, D. (1990). *Counseling the culturally different*. *Theory and practice*. New York: John Wiley & Sons.

Taylor, W. L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, *30*, 415-533.

Travers, K. J. (1988). Opportunity to learn mathematics in the eighth-grade classrooms in the United States: Some findings from the Second International Mathematics Study. In R. R. Cocking & J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics* (pp. 187-199). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wang, L. L. (1988). *A causal analysis of achievement models for language minority students in the United States: A linear structural relations (LISREL) approach* (Doctoral dissertation, University of Illinois at Urbana-Champaign). *Dissertation Abstracts International*, *49,* 3210A.

Wang, M. D. (1970). The role of syntactic complexity as a determiner of comprehensibility. *Journal of Verbal Learning and Verbal Behavior*, *9*, 398-404.

Winefield, H. R., Teggemann, M., Goldney, R. D., & Winefield, A. H. (1992). Predictors and consequences of tertiary education: A nine-year follow-up of academically capable school children. *British Journal of Educational Psychology, 62*, 148-153.

Wolk, R. A., (Ed.). (1992, June 17). By all measures: The debate over standards and assessment [Special issue]. *Education Week*.

Yngve, V. H. (1960). A model and hypothesis for language structure. *Proceedings of the American Philosophical Association*, *404*, 444-466.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley.

Appendix A

Tables 1-25

Discriminant Analysis

Table 1

Background Variables Used in Discriminant Analyses for the 1990 Data

1.	DSEX	gender
2.	IEP	individualized education plan
3.	LEP	limited English proficiency
4.	b000901a	does your family get a newspaper regularly
5.	b000903a	is there an encyclopedia in your home
6.	b000904a	are there more than 25 books in your home
7.	b000905a	does your family get magazines regularly
8.	b005601a	does mother or stepmother live at home
9.	b005701	does father or stepfather live at home
10.	DRACE	race/ethnicity
11.	SCHTYPE	school type
12.	b003501a	mother's education level
13.	b003601a	father's education level
14.	PARED	parent's education level
15.	REGION	region of country
16.	b003201	how often other than English spoken in home
17.	HOMEEN2	reading materials
18.	SINGLEP	single parent
19.	b001801a	TV watch
20.	IDP	instruction dollars per pupil

Table 2

Background Variables Used in Discriminant Analyses for the 1992 Data

1.	DSEX	gender
2.	IEP	individualized education plan
3.	LEP	limited language proficiency
4.	b000901a	does your family get a newspaper
5.	b000903a	is there an encyclopedia
6.	b000904a	are there more than 25 books
7.	b000905a	does your family get magazines
8.	b005601a	does mother or stepmother live at home
9.	b005701	does father or stepfather live at home
10.	DRACE	derived ethnicity
11.	SDOC	sampling description of community
12.	SCHTYPE	school type
13.	b003501a	mother's education level
14.	b003601a	father's education level
15.	PARED	parent's education level
16.	REGION	region of country
17.	HOMEEN2	home environment, reading materials
18.	SINGLEP	how many parents
19.	b001801a	how much TV
20.	STOC	size and type of community
21.	b003201a	how often other than English
22.	RACOFTN	by race/ethnicity other than English
23.	LANGHOM	how often other than English

Table 3

Results of the DA, Mean of Long and Short Items by: Gender, Grade 8, Booklet 8, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	627	.52	607	.52	.082	.775	-1.21	-.196
Short items	627	.58	607	.57	.620	.431	1.41	.540
Canonical correlation							0.042	
Wilks's lambda							0.998	
Chi square							2.12	
df							2.00	
P							0.346	

Table 4

Results of the DA, Mean of Long and Short Items by: Newspaper Regularly, Grade 8, Booklet 8, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	942	.54	257	.47	17.62	.000	0.670	.959
Short items	942	.59	257	.53	14.94	.000	0.406	.883
Canonical correlation							0.126	
Wilks's lambda							0.984	
Chi square							19.00	
df							2.00	
P							0.001	

Table 5

Results of the DA, Mean of Long and Short Items by: Encyclopedia, Grade 8, Booklet 8, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	932	.53	267	.49	9.31	.002	0.457	.947
Short items	932	.59	267	.54	10.32	.001	0.622	.899
Canonical correlation							0.976	
Wilks's lambda							0.990	
Chi square							11.44	
df							2.00	
P							0.003	

Table 6

Results of the DA, Mean of Long and Short Items by: More than 25 Books, Grade 8, Booklet 8, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	1100	.54	75	.37	36.37	.000	0.347	.858
Short items	1100	.59	75	.42	46.43	.000	0.724	.969
Canonical correlation							0.201	
Wilks's lambda							0.960	
Chi square							48.38	
df							2.00	
P							0.000	

Table 7

Results of the DA, Mean of Long and Short Items by: Magazine Regularly, Grade 8, Booklet 8, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	923	.55	257	.44	46.77	.000	0.667	.957
Short items	923	.60	257	.51	39.69	.000	0.410	.881
Canonical correlation							0.204	
Wilks's lambda							0.958	
Chi square							49.96	
df							2.00	
P							0.000	

Table 8

Results of the DA, Mean of Long and Short Items by: How Often Other Than English, Grade 8, Booklet 8, 1990, Math

	Groups											
	1		2		3		Univariate		Function 1		Function 2	
Variables	No.	*M*	No.	*M*	No	*M*	*F*	*P*	Standardized Coefficients	Structure Coefficient	Standardized Coefficients	Structure Coefficient
Long items	789	.53	309	.51	125	.47	5.39	.005	1.39	0.850	-0.352	0.526
Short items	789	.58	309	.58	125	.55	0.91	.401	-0.756	0.244	1.22	0.969
Canonical correlation									0.108		0.029	
Wilks's lambda									0.987		0.999	
Chi square									15.43		1.02	
df									4.00		1.00	
P									0.004		0.313	

Table 9

Results of the DA, Mean of Long and Short Items by: Home Environment, Grade 8, Booklet 8, 1990, Math.

	Groups											
	1		2		3		Univariate		Function 1		Function 2	
Variables	No.	*M*	No.	*M*	No	*M*	*F*	*P*	Standardized Coefficients	Structure Coefficient	Standardized Coefficients	Structure Coefficient
Long items	268	.43	366	.50	594	.57	41.25	.000	0.563	0.927	-1.28	-0.374
Short items	268	.49	366	.56	594	.62	40.17	.000	0.523	0.915	1.29	0.403
Canonical correlation									0.270		0.005	
Wilks's lambda									0.927		1.00	
Chi square									92.37		0.25	
df									4.00		1.00	
P									0.000		0.403	

Table 10

Results of the DA, Mean of Long and Short Items by: Gender, Grade 8, Booklet 9, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	648	.40	596	.42	5.72	.017	1.25	0.65
Short items	648	.52	596	.51	0.528	.467	-0.98	-0.18
Canonical correlation							0.105	
Wilks's lambda							0.989	
Chi square							13.65	
df							2.00	
P							0.001	

Table 11

Results of the DA, Mean of Long and Short Items by: Newspaper Regularly, Grade 8, Booklet 9, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	906	.43	296	.37	16.64	.000	0.704	.949
Short items	906	.53	294	.48	12.76	.000	0.399	.831
Canonical correlation							0.123	
Wilks's lambda							0.985	
Chi square							18.32	
df							2.00	
P							0.001	

Table 12

Results of the DA, Mean of Long and Short Items by: Encyclopedia, Grade 8, Booklet 9, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	974	.42	233	.37	10.73	.001	0.142	.697
Short items	974	.53	233	.46	21.82	.000	0.907	.994
Canonical correlation							0.134	
Wilks's lambda							0.982	
Chi square							21.88	
df							2.00	
P							0.000	

Table 13

Results of the DA, Mean of Long and Short Items by: More than 25 books, Grade 8, Booklet 9, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	1137	.43	49	.25	32.12	.000	0.354	.804
Short items	1137	.53	49	.33	45.67	.000	0.745	.959
Canonical correlation							0.201	
Wilks's lambda							0.960	
Chi square							48.58	
df							2.00	
P							0.000	

Table 14

Results of the DA, Mean of Long and Short Items by: Magazine Regularly, Grade 8, Booklet 9, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	956	.44	227	.34	40.15	.000	0.873	.987
Short items	956	.54	227	.46	21.38	.000	0.189	.721
Canonical correlation							0.183	
Wilks's lambda							0.966	
Chi square							40.35	
df							2.00	
P							0.000	

Table 15

Results of the DA, Mean of Long and Short Items by: How often other than English, Grade 8, Booklet 9, 1990, Math.

	Groups											
	1		2		3		Univariate		Function 1		Function 2	
Variables	No.	*M*	No.	*M*	No	*M*	*F*	*P*	Standardized Coefficients	Structure Coefficient	Standardized Coefficients	Structure Coefficient
Long items	814	.43	317	.41	104	.32	10.50	.000	0.964	0.999	-0.823	-0.045
Short items	814	.53	317	.51	104	.46	4.44	.012	0.057	0.649	1.27	0.760
Canonical correlation									0.130		0.003	
Wilks's lambda									0.983		1.00	
Chi square									20.87		0.009	
df									4.00		1.00	
P									0.000		0.924	

Table 16

Results of the DA, Mean of Long and Short Items by: Home Environment, Grade 8, Booklet 9, 1990, Math.

	Groups											
	1		2		3		Univariate		Function 1		Function 2	
Variables	No.	*M*	No.	*M*	No	*M*	*F*	*P*	Standardized Coefficients	Structure Coefficient	Standardized Coefficients	Structure Coefficient
Long items	263	.32	371	.40	606	.46	41.86	.000	0.632	0.921	-1.07	-0.390
Short items	263	.44	371	.50	606	.56	36.63	.012	0.485	0.862	1.16	0.508
Canonical correlation									0.272		0.001	
Wilks's lambda									0.926		1.00	
Chi square									94.93		0.001	
df									4.00		1.00	
P									0.000		0.970	

Table 17

Results of the DA, Mean of Long and Short Items by: Gender, Grade 8, Booklet 10, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	644	.42	586	.43	0.742	.389	-1.39	-.249
Short items	644	.57	586	.55	2.28	.131	1.50	.436
Canonical correlation							0.098	
Wilks's lambda							0.990	
Chi square							11.93	
df							2.00	
P							0.003	

Table 18

Results of the DA, Mean of Long and Short Items by: Newspaper Regularly, Grade 8, Booklet 10, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	872	.45	320	.37	26.86	.000	0.791	.986
Short items	872	.58	320	.51	20.25	.000	0.258	.856
Canonical correlation							0.151	
Wilks's lambda							0.977	
Chi square							27.31	
df							2.00	
P							0.000	

Table 19

Results of the DA, Mean of Long and Short Items by: Encyclopedia, Grade 8, Booklet 10, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	951	.44	232	.37	17.69	.000	0.217	.839
Short items	951	.58	232	.50	24.60	.000	0.827	.990
Canonical correlation							0.144	
Wilks's lambda							0.979	
Chi square							24.82	
df							2.00	
P							0.000	

Table 20

Results of the DA, Mean of Long and Short Items by: More than 25 books, Grade 8, Booklet 10, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	1119	.44	68	.24	49.31	.000	0.459	.915
Short items	1119	.57	68	.37	53.44	.000	0.609	.952
Canonical correlation							0.218	
Wilks's lambda							0.953	
Chi square							57.44	
df							2.00	
P							0.000	

Table 21

Results of the DA, Mean of Long and Short Items by: Magazine Regularly, Grade 8, Booklet 10, 1990, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	916	.45	243	.34	40.08	.000	0.633	.959
Short items	916	.58	243	.49	36.06	.000	0.432	.910
Canonical correlation							0.191	
Wilks's lambda							0.964	
Chi square							42.76	
df							2.00	
P							0.000	

Table 22

Results of the DA, Mean of Long and Short Items by: How Often Other Than English, Grade 8, Booklet 10, 1990, Math.

	Groups											
	1		2		3		Univariate		Function 1		Function 2	
Variables	No.	*M*	No.	*M*	No	*M*	*F*	*P*	Standardized Coefficients	Structure Coefficient	Standardized Coefficients	Structure Coefficient
Long items	810	.43	303	.43	113	.36	5.28	.005	0.811	0.986	-1.31	-1.51
Short items	810	.56	303	.56	113	.50	3.92	.020	0.233	0.850	1.52	0.527
Canonical correlation									0.094		0.009	
Wilks's lambda									0.001		1.00	
Chi square									10.86		0.105	
df									4.00		1.00	
P									0.28		0.746	

Table 23

Results of the DA, Mean of Long and Short Items by: Home, Grade 10, Booklet 8, 1990, Math

	Groups											
	1		2		3		Univariate		Function 1		Function 2	
Variables	No.	*M*	No.	*M*	No	*M*	*F*	*P*	Standardized Coefficients	Structure Coefficient	Standardized Coefficients	Structure Coefficient
Long items	285	.32	362	.46	576	.48	45.05	.000	0.539	0.934	-1.39	-0.356
Short items	285	.42	362	.56	576	.61	44.87	.000	0.532	0.933	1.40	0.361
Canonical correlation									0.279		0.007	
Wilks's lambda									0.922		1.00	
Chi square									99.04		0.635	
df									4.00		1.00	
P									0.000		0.801	

Table 24

Results of the DA, Mean of Long and Short Items by: LANGHOM (1 = Never, 2 = Always), Grade 8, Booklets 1, 2, 15, 1992, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	156	.101	120	-.132	3.74	.05	0.70	.960
Short items	156	.090	120	-.117	3.96	.09	0.38	.855
Canonical correlation							0.121	
Wilks's lambda							0.985	
Chi square							4.01	
df							2.00	
P							0.135	

Table 25

Result of the DA, Composite Scores of Long and Short Items by: LANGHOM (1 = Never, 2 = Always), Grade 8, Booklets 10, 19, 24, 1992, Math

	Groups							
	1		2		Univariate		Function 1	
Variables	No.	Mean	No.	Mean	F	P	Standardized coefficients	Structure coefficient
Long items	155	.124	149	-.129	4.93	.03	1.23	.914
Short items	155	.034	149	-.035	0.363	.55	-0.517	.248
Canonical correlation							0.138	
Wilks's lambda							0.981	
Chi square							5.83	
df							2.00	
P							0.054	

Appendix B

Tables 26 and 27

Analysis of Variance Tables for Analyses Based on the Linguistic Characteristics of Items

(Tables 26 and 27 are repeated in text.)

Table 26

Analysis of Variance Summary Table, 1992, 8th Grade, Block 8

Source of variation	SS	*df*	*MS*	*F*	*P*
Between subjects					
A (sex)	.16	1	.16	1.43	.23
Subject within group					
Within subjects					
B (problems)	3.24	1	3.24	56.42	.00
AB (sex x problem)	.26	1	.26	4.44	.04
B x subject within group					

Table 27

Analysis of Variance Summary Table, 1992, 8th Grade, Block 15

Source of VARIATION	SS	*df*	*MS*	*F*	*P*
Between subjects					
A (sex)	2.65	1	2.65	23.71	.00
Subject within group					
Within subjects					
B (problems)	3.33	1	3.33	67.96	.00
B x subject within group					
AB (sex x problem)	.28	1	.28	5.73	.02
AB x subject within group					

Appendix C

Tables 28-33

Analysis for the Variable “LANGHOM”:

Omitted/Not-Reached NAEP Math Items for Grade 8, 1992

Table 28

Means for NAEP Items Omitted/Not Reached, Grouped on LANGHOM.

	Means		
Variable	LANHGOM = 1	LANGHOM = 2	LANGHOM = 3
M050261C	0.00091	0.0015	0.0018
M051101C	0.015	0.022	0.028
M051021C	0.011	0.013	0.017
M045801G	0.032	0.037	0.040
M045802G	0.033	0.034	0.040
M045803G	0.038	0.038	0.050
M045804G	0.034	0.034	0.051
M045861G	0.0090	0.010	0.011
M045901G	0.043	0.046	0.072
M045941G	0.012	0.011	0.024
M052421I	0.017	0.022	0.014
M053101I	0.026	0.028	0.027
M052821I	0.0096	0.013	0.0074
M054301L	0.040	0.048	0.066
M054341L	0.011	0.013	0.023
M052201M	0.026	0.028	0.051
M055501N	0.063	0.078	0.086
M055541N	0.017	0.018	0.031
M048721O	0.011	0.012	0.014
M049901C	0.0019	0.0019	0.00093
M050001C	0.0014	0.0011	0.0028
M050101C	0.0044	0.0034	0.0037
M050201C	0.013	0.016	0.015
M050202C	0.012	0.016	0.013
M050203C	0.014	0.019	0.019
M050204C	0.016	0.023	0.019
M050301C	0.0033	0.0031	0.0055
M050401C	0.0018	0.0034	0.00093
M050501C	0.0073	0.0095	0.012
M050601C	0.0059	0.0084	0.0093
M050701C	0.0029	0.0034	0.0018
M050801C	0.014	0.024	0.031
M050901C	0.050	0.052	0.071
M051001C	0.063	0.075	0.083
M017401D	0.0015	0.0011	0.00093
M017501D	0.00091	0.0	0.0018
M017601D	0.0019	0.0023	0.0037
M017701D	0.0019	0.0011	0.0027
M017801D	0.0018	0.0023	0.0037

Table 28 (continued)

Variable	Means		
	LANHGOM = 1	LANGHOM = 2	LANGHOM = 3
M017901D	0.0021	0.0015	0.0055
M018001D	0.0024	0.0023	0.0046
M018101D	0.0027	0.0026	0.0028
M018201D	0.0025	0.0023	0.0018
M018301D	0.0022	0.0026	0.0028
M018401D	0.0036	0.0023	0.0074
M018501D	0.0038	0.0046	0.0093
M018601D	0.0032	0.0038	0.0037
M018701D	0.0041	0.0061	0.012
M018801D	0.013	0.014	0.017
M018901D	0.0096	0.013	0.017
M019001D	0.0094	0.012	0.013
M019101D	0.014	0.016	0.025
M019201D	0.015	0.021	0.026
M019301D	0.026	0.035	0.033
M019601D	0.022	0.031	0.029
M021901E	0.0018	0.00038	0.0
M022001E	0.0029	0.0034	0.0037
M022101E	0.0014	0.0026	0.0018
M022201E	0.0052	0.0031	0.0083
M022301E	0.0017	0.0019	0.0
M022401E	0.0027	0.0027	0.00093
M022501E	0.0099	0.0099	0.015
M022601E	0.0021	0.0023	0.0027
M022701E	0.0033	0.0023	0.014
M022801E	0.0096	0.0080	0.015
M022802E	0.010	0.0073	0.016
M022901E	0.0074	0.0073	0.017
M023001E	0.0073	0.0069	0.012
M023101E	0.012	0.010	0.017
M023201E	0.015	0.012	0.019
M023301E	0.018	0.018	0.028
M023401E	0.022	0.024	0.033
M023501E	0.027	0.028	0.045
M023601E	0.032	0.035	0.053
M023701E	0.051	0.056	0.085
M023801E	0.038	0.041	0.059
M019701F	0.0053	0.0053	0.0093
M019801F	0.013	0.011	0.026
M019901F	0.0087	0.011	0.017
M020001F	0.010	0.018	0.033
M020101F	0.0036	0.0042	0.012
M020201F	0.014	0.019	0.031
M020301F	0.0029	0.0030	0.0083
M020401F	0.0064	0.0096	0.016

Table 28 (continued)

Variable	Means		
	LANHGOM = 1	LANGHOM = 2	LANGHOM = 3
M020501F	0.011	0.013	0.018
M020801F	0.035	0.037	0.061
M020901F	0.058	0.058	0.081
M021001F	0.011	0.010	0.021
M021101F	0.021	0.024	0.049
M021201F	0.025	0.024	0.051
M021301F	0.017	0.019	0.030
M021302F	0.021	0.023	0.039
M044501G	0.0023	0.0031	0.0065
M044601G	0.0068	0.0072	0.013
M044641G	0.0022	0.0019	0.0018
M044701G	0.0011	0.0011	0.00093
M044801G	0.0082	0.0053	0.0046
M044901G	0.0011	0.0011	0.0
M045001G	0.0015	0.0023	0.00093
M045101G	0.0091	0.0084	0.017
M045141G	0.0021	0.0038	0.0037
M045201G	0.0029	0.0026	0.0037
M045301G	0.030	0.032	0.043
M045341G	0.0082	0.0080	0.010
M045601G	0.017	0.018	0.023
M045641G	0.0054	0.0049	0.0037
M045701G	0.027	0.035	0.043
M045741G	0.0065	0.010	0.016
M012231H	0.0019	0.0019	0.0037
M012331H	0.0027	0.0046	0.0046
M012431H	0.010	0.0072	0.014
M012531H	0.0	0.0	0.0
M012631H	0.0068	0.0057	0.0055
M012731H	0.0067	0.0072	0.010
M012831H	0.0067	0.0084	0.0093
M012931H	0.011	0.015	0.016
M013031H	0.022	0.023	0.029
M013131H	0.049	0.048	0.063
M013231H	0.023	0.025	0.034
N202831H	0.015	0.016	0.024
M011131H	0.023	0.025	0.029
M013331H	0.026	0.027	0.038
M013431H	0.038	0.035	0.052
M013531H	0.053	0.053	0.073
M013631H	0.051	0.050	0.068
M013731H	0.074	0.068	0.091
M052301I	0.0018	0.0019	0.00093
M052401I	0.036	0.033	0.036
M052501I	0.013	0.015	0.010
M052601I	0.00061	0.0015	0.00093
M052701I	0.0015	0.0015	0.00093

Table 28 (continued)

Variable	Means		
	LANHGOM = 1	LANGHOM = 2	LANGHOM = 3
M052801I	0.0041	0.0042	0.0055
M052901I	0.017	0.016	0.019
M053001I	0.044	0.044	0.057
M061901J	0.010	0.013	0.015
M061903J	0.0054	0.0080	0.0083
M061904J	0.0081	0.0095	0.012
M061902J	0.013	0.016	0.024
M061907J	0.0081	0.014	0.015
M061908J	0.014	0.017	0.029
M061905J	0.017	0.023	0.029
M046001K	0.0090	0.0065	0.014
M046101K	0.00061	0.00038	0.00093
M046201K	0.00076	0.00038	0.00093
M046301K	0.0015	0.0019	0.0028
M046401K	0.00076	0.0015	0.0018
M046501K	0.0015	0.0023	0.0018
M046601K	0.0036	0.0038	0.010
M046701K	0.0011	0.0015	0.00285
M046801K	0.0029	0.0023	0.0065
M046901K	0.0033	0.0038	0.0093
M047001K	0.0021	0.00038	0.0018
M047101K	0.0021	0.0011	0.0028
M046201K	0.0033	0.0030	0.0046
M047301K	0.0079	0.0084	0.012
M047601K	0.0079	0.0095	0.017
M046701K	0.012	0.016	0.017
M046801K	0.010	0.011	0.021
M046901K	0.026	0.026	0.044
M048001K	0.021	0.023	0.037
M053501L	0.0024	0.0034	0.0028
M053601L	0.0032	0.0030	0.005+
M053701L	0.0064	0.0049	0.010
M053801L	0.0092	0.010	0.011
M053901L	0.0044	0.0038	0.0046
M054001L	0.018	0.021	0.022
M054041L	0.0053	0.0049	0.0037
M054101L	0.029	0.037	0.041
M054141L	0.0081	0.0088	0.011
M054201L	0.0070	0.0099	0.0056
M051201M	0.0064	0.0053	0.0084
M051301M	0.0019	0.0042	0.0028
M051401M	0.0010	0.0015	0.0028
M051501M	0.0071	0.0092	0.0093
M051601M	0.0084	0.0084	0.012
M051701M	0.0081	0.0053	0.010
M051801M	0.0077	0.0049	0.012
M051901M	0.0026	0.0034	0.0028

Table 28 (continued)

Variable	Means		
	LANHGOM = 1	LANGHOM = 2	LANGHOM = 3
M052001N	0.0058	0.0053	0.0056
M052101N	0.0079	0.010	0.011
M054701N	0.0072	0.011	0.0074
M054801N	0.0048	0.013	0.012
M054841N	0.0014	0.0034	0.0046
M054901N	0.0099	0.0084	0.011
M055101N	0.018	0.022	0.027
M055201N	0.0081	0.017	0.020
M055240N	0.0018	0.0042	0.012
M055301N	0.016	0.020	0.018
M055401N	0.019	0.026	0.021
M048101O	0.0017	0.0015	0.0056
M048201O	0.00061	0.00076	0.0037
M048301O	0.00091	0.00038	0.0018
M048401O	0.00046	0.00038	0.0018
M048501O	0.0021	0.0015	0.0028
M048601O	0.0014	0.0019	0.0028
M048701O	0.0099	0.016	0.024
M048741O	0.0023	0.0042	0.0093
M048801O	0.010	0.012	0.019
M048841O	0.0021	0.0019	0.0093
M048901O	0.0078	0.014	0.023
M048940O	0.0019	0.0030	0.0093
M049101O	0.0019	0.0019	0.0046
M049201O	0.0014	0.0026	0.0046
M049301O	0.012	0.012	0.011
M049401O	0.0030	0.0042	0.0065
M049501O	0.0091	0.011	0.010
M049601O	0.0085	0.010	0.014
M049701O	0.011	0.016	0.014
M049801O	0.037	0.047	0.063
M049841O	0.010	0.013	0.020

Note. LANGHOM = 1, “Never” other than English spoken in the home; LANGHOM = 2, “Sometimes” other than English spoken; LANGHOM = 3, “Always” other than English spoken.

Table 29

NAEP Items Omitted/Not Reached, Booklet 1, Grouped on LANGHOM

Variable	Means		
	LANHGOM=1	LANGHOM=2	LANGHOM=3
NUM2	0.015	0.021	0.0
NUM5	0.0078	0.010	0.023
NUM11	0.0039	0.020	0.0
NUM12	0.0	0.010	0.0
MEA1	0.0039	0.0	0.0
MEA4	0.0078	0.0	0.023
MEA5	0.023	0.020	0.023
GEO4	0.015	0.	0.069
GEO6	0.12	0.14	0.13
GEO7	0.015	0.031	0.069
STA6	0.047	0.072	0.069
ALG1	0.019	0.031	0.023
ALG3	0.015	0.041	0.046
ALG4	0.047	0.093	0.069
NUM4	0.011	0.020	0.0
NUM14	0.051	0.041	0.069
MEA2	0.035	0.020	0.046
MEA3	0.20	0.22	0.32
MEA7	0.051	0.11	0.11
MEA8	0.066	0.11	0.11
GEO1	0.043	0.072	0.16
GEO2	0.24	0.23	0.30
GEO3	0.011	0.0	0.023
GEO8	0.055	0.13	0.069
STA1	0.047	0.10	0.069
STA2	0.062	0.062	0.023
STA5	0.011	0.010	0.023
STA8	0.13	0.15	0.16
ALG6	0.0078	0.010	0.023

Note. LANGHOM = 1, "Never" other than English spoken in the home; LANGHOM = 2, "Sometimes" other than English spoken; LANGHOM = 3, "Always" other than English spoken.

Table 30

NAEP Items Omitted/Not Reached, Booklet 2, Grouped on LANGHOM

ITEM	Means		
	LANGHOM=1	LANGHOM=2	LANGHOM=3
NUM11	0.028	0.028	0.12
NUM14	0.177	0.17	0.27
NUM20	0.12	0.12	0.21
NUM22	0.19	0.21	0.24
MEA1	0.0041	0.0	0.030
MEA7	0.041	0.028	0.090
MEA10	0.24	0.20	0.36
MEA12	0.090	0.076	0.12
GEO2	0.012	0.0095	0.060
GEO6	0.045	0.066	0.030
GEO8	0.016	0.028	0.15
STA2	0.041	0.028	0.18
ALG2	0.033	0.0095	0.090
ALG3	0.061	0.038	0.15
ALG8	0.0	0.0095	0.030
ALG9	0.066	0.085	0.12
NUM8	0.016	0.019	0.030
NUM9	0.0082	0.019	0.0
NUM10	0.024	0.0095	0.030
NUM13	0.11	0.10	0.21
MEA4	0.066	0.028	0.27
MEA5	0.090	0.076	0.21
MEA6	0.012	0.0095	0.060
GEO1	0.012	0.0095	0.060
GEO7	0.045	0.057	0.15
STA1	0.0082	0.0	0.030
STA6	0.21	0.20	0.21
NUM2	0.012	0.0	0.030
NUM3	0.0082	0.0	0.060
NUM5	0.0082	0.0095	0.060
ALG5	0.016	0.028	0.0
ALG6	0.061	0.057	0.15
ALG7	0.13	0.12	0.242
ALG10	0.29	0.25	0.36

Note. LANGHOM = 1, “Never” other than English spoken in the home; LANGHOM = 2, “Sometimes” other than English spoken; LANGHOM = 3, “Always” other than English spoken.

Table 31

NAEP Items Omitted/Not Reached, Booklet 15, Grouped on LANGHOM

ITEM	Means LANGHOM=1	Means LANGHOM=2	Means LANGHOM=3
NUM8	0.016	0.0099	0.045
NUM14	0.016	0.039	0.045
MEA1	0.0	0.0	0.0
MEA10	0.016	0.0099	0.045
GEO1	0.0	0.0099	0.0
GEO7	0.053	0.039	0.090
GEO12	0.016	0.0099	0.022
STA2	0.037	0.049	0.090
STA3	0.057	0.079	0.15
NUM2	0.0041	0.0	0.022
NUM3	0.0082	0.0	0.022
NUM5	0.012	0.0099	0.022
NUM6	0.012	0.	0.022
NUM9	0.037	0.059	0.068
MEA2	0.0041	0.029	0.0
MEA5	0.041	0.049	0.13
MEA6	0.049	0.059	0.20
GEO13	0.033	0.029	0.090
GEO14	0.090	0.059	0.11
STA1	0.0	0.0099	0.022
STA4	0.11	0.17	0.18
STA5	0.074	0.10	0.25
ALG6	0.028	0.019	0.11

Note. LANGHOM = 1, “Never” other than English spoken in the home; LANGHOM = 2, “Sometimes” other than English spoken; LANGHOM = 3, “Always” other than English spoken.

Table 32

NAEP Items Omitted/Not Reached, Booklet 19, Grouped on LANGHOM

ITEM	Means		
	LANGHOM=1	LANGHOM=2	LANGHOM=3
S6	0.0040	0.011	0.
S7	0.016	0.011	0.
S8	0.036	0.045	0.047
S9	0.040	0.034	0.047
S10	0.10	0.12	0.19
S11	0.12	0.10	0.23
S12	0.16	0.12	0.33
S13	0.25	0.19	0.38
S17	0.020	0.0	0.023
S18	0.0040	0.0	0.023
S19	0.0040	0.0	0.0
S20	0.016	0.0	0.0
L11	0.088	0.10	0.21
L12	0.22	0.27	0.30
L13	0.35	0.31	0.50
L14	0.036	0.080	0.047
L20	0.048	0.11	0.095
L21	0.036	0.022	0.71
L22	0.052	0.034	0.71

Note. LANGHOM = 1, “Never” other than English spoken in the home; LANGHOM = 2, “Sometimes” other than English spoken; LANGHOM = 3, “Always” other than English spoken.

Tables 33

NAEP Items Omitted/Not Reached, Booklet 24, Grouped on LANGHOM

ITEM	Means		
	LANGHOM=1	LANGHOM=2	LANGHOM=3
S3	0.015	0.0	0.018
S4	0.023	0.018	0.056
S5	0.0079	0.0091	0.0
S15	0.027	0.0091	0.018
S16	0.015	0.018	0.037
S17	0.0	0.0091	0.037
S18	0.0039	0.0	0.037
S19	0.0039	0.0	0.037
S20	0.0079	0.0091	0.056
L6	0.043	0.055	0.13
L7	0.0	0.0091	0.037
L8	0.13	0.14	0.22
L9	0.051	0.10	0.094
L10	0.21	0.22	0.39
L19	0.027	0.045	0.094
L20	0.027	0.036	0.13
L21	0.031	0.045	0.094
L22	0.051	0.10	0.075

Note. LANGHOM = 1, "Never" other than English spoken in the home; LANGHOM = 2, "Sometimes" other than English spoken; LANGHOM = 3, "Always" other than English spoken.

Appendix D

Sample Protocol for Student Perceptions Study Interviews

Ask if it's okay to turn on the tape recorder.

Try to get the student to talk so that we can have a language sample. The following types of questions could be used:

What is your favorite subject in school? What's your best subject?

How do you take math tests: do you do all the hard problems first, the easy ones, or just take them one by one?

What kind of math problems would you rather do: numbers only, numbers and words, or words only? Which are easier for you?

Do you like problems that require a lot of thought, or problems that are easy to figure out?

Would you rather have problems about abstract situations, or problems about real life situations?

Present the original and the revised form of the first item and say:

Let me know when you've finished.

After the student has read the items silently, ask the following:

If you were really in a hurry on a test and you had to pick one of these problems to do, which one would you do?

Read it aloud to me.

Now read the other one aloud to me.

Are there words in either of them that might be confusing for some students or hard for them to understand?

What is it about the one you chose that seems easier?

Once the items have been read and discussed, say the following:

I have just one last question to ask you.

Do you speak any languages other than English? What language?

What language do you speak at home? to your friends at school? to your best friend? to your mother/father?

Appendix E

Tables 34 and 35

Interview Results, Stages 1 and 2

(Table 35 is repeated in text.)

Table 34

Stage 1 Interview Results: Students' Choices (N = 19)

Item #	Original item chosen	Revised item chosen
1	3	16
2	4	15
3	10	9
4	11	8

Table 35

Stage 2 Interview Results: Students' Choices (N = 17)

Item #	Original item chosen	Revised item chosen
4[a]	3	14
5	4.5[b]	12.5
6	2	15
7	2	15

[a] Modified (piloted for a second time) version of item 4.
[b] One student was ambivalent about his choice.

Appendix F

Original Test Items Plus Control Items

In item 6, the parenthetical statement “1 square yard = 9 square feet” was printed as “1 square yard = 9 feet,” and may have impacted interpretation of the question by some students. The item is printed correctly in this appendix.

1. A certain reference file contains approximately six billion facts. About how many millions is that?

   (A) 6,000,000

   (B) 600,000

   (C) 60,000

   (D) 6,000

   (E) 600

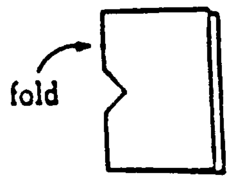
2. In a bag of marbles, 1/2 are red, 1/4 are blue, 1/6 are green, and 1/12 are yellow. If a marble is taken from the bag without looking, it is most likely to be

   (A) red

   (B) blue

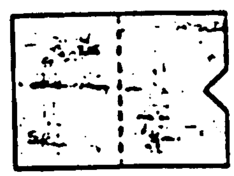
   (C) green

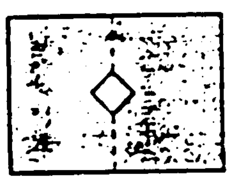
   (D) yellow

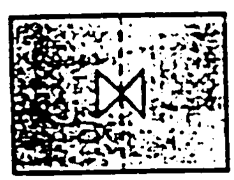


3. A sheet of paper is folded once and a piece is cut out as shown above. Which of the following looks like the unfolded paper?

A 

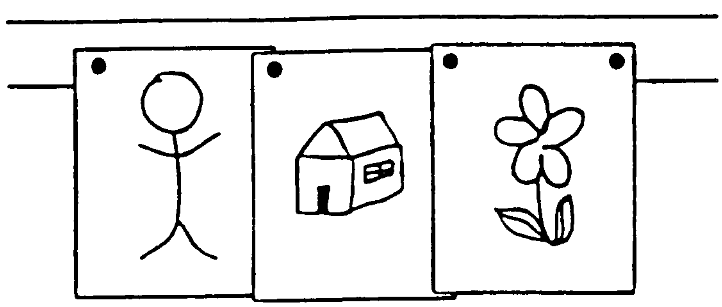
B 

C 

D 

4. Raymond must buy enough paper to print 28 copies of a report that contains 64 sheets of paper. Paper is only available in packages of 500 sheets. How many whole packages of paper will he need to buy to do the printing?

Answer: ______________________



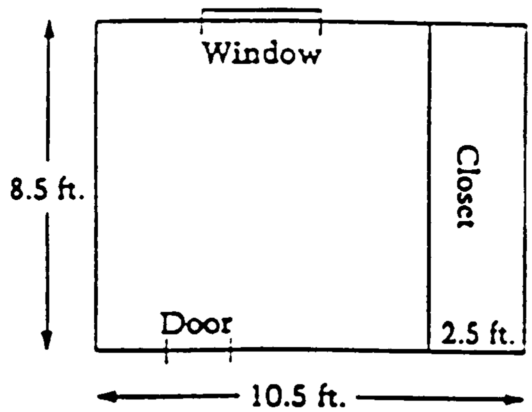
5. Children's pictures are to be hung in a line as shown in the figure above. Pictures that are hung next to each other share a tack. How many tacks are needed to hang 28 pictures in this way?

(A) 27

(B) 28

(C) 29

(D) 56



6. Chris wishes to carpet the rectangular room shown above. To the nearest square yard, how many square yards of carpet are needed to carpet the floor of the room if the closet floor will not be carpeted?
(1 square yard = 9 square feet)

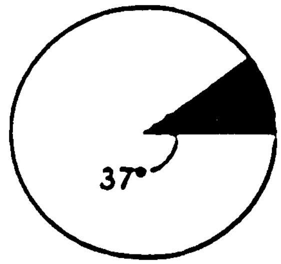
- (A) 8
- (B) 10
- (C) 11
- (D) 19
- (E) 22

7. Harriet, Jim, Roberto, Maria, and Willie are in the same eighth-grade class. One of them is this year's class president. Based on the following information, who is the class president?

1. The class president was last year's class vice president and lives on Vine Street.
2. Willie is this year's class vice president.
3. Jim and Maria live on Cypress Street.
4. Roberto was not last year's vice president.

- (A) Jim
- (B) Harriet
- (C) Roberto
- (D) Maria
- (E) Willie

RADIO SALES



8. The entire circle shown above represents a total of 2,675 radios sold. Of the following, which is the best approximation of the number of radios represented by the shaded sector of the circle?

Ⓐ 70

Ⓑ 275

Ⓒ 985

Ⓓ 25,880

Ⓔ 98,420

Puppy's Age	Puppy's Weight
1 month	10 lbs.
2 months	15 lbs.
3 months	19 lbs.
4 months	22 lbs.
5 months	?

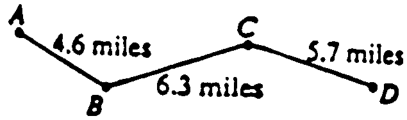
9. John records the weight of his puppy every month in a chart like the one shown above. If the pattern of the puppy's weight gain continues, how many pounds will the puppy weigh at 5 months?

Ⓐ 30

Ⓑ 27

Ⓒ 25

Ⓓ 24



10. Carol wanted to estimate the distance from A to D along the path shown on the map above. She correctly rounded each of the given distances to the nearest mile and then added them. Which of the following sums could be hers?

Ⓐ 4 + 6 + 5 = 15

Ⓑ 5 + 6 + 5 = 16

Ⓒ 5 + 6 + 6 = 17

Ⓓ 5 + 7 + 6 = 18

54 > 3 x □

11. Write two numbers that could be put in the □ to make the number sentence above true.

Answer: ______________________________

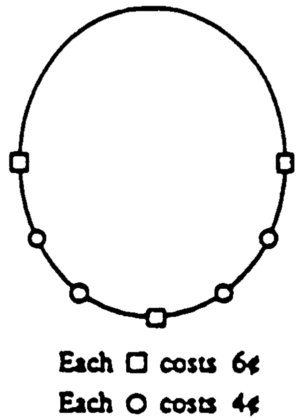
____

12. If □ represents the number of newspapers that Lee delivers each day, which of the following represents the total number of newspapers that Lee delivers in 5 days?

- (A) 5 + □
- (B) 5 x □
- (C) □ + 5
- (D) (□ + □) x 5

13. The length of a dinosaur was reported to have been 80 feet (rounded to the nearest 10 feet). What length other than 80 feet could have been the actual length of this dinosaur?

Answer: ________________ feet



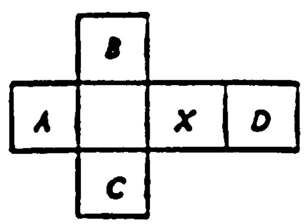
14. If the string does not cost anything, how much does the necklace above cost?

Ⓐ 10¢

Ⓑ 24¢

Ⓒ 28¢

Ⓓ 34¢



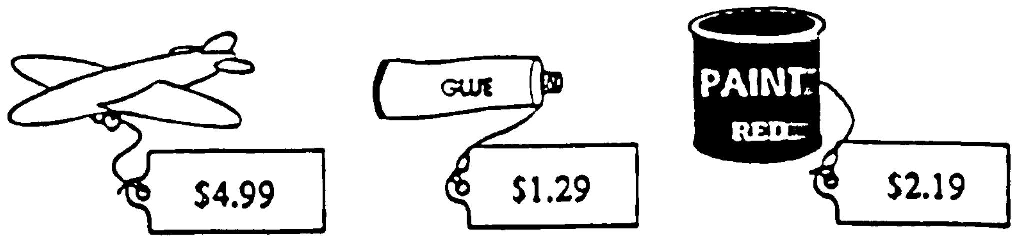
15. The squares in the figure above represent the faces of a cube which has been cut along some edges and flattened. When the original cube was resting on face X, which face was on top?

Ⓐ A

Ⓑ B

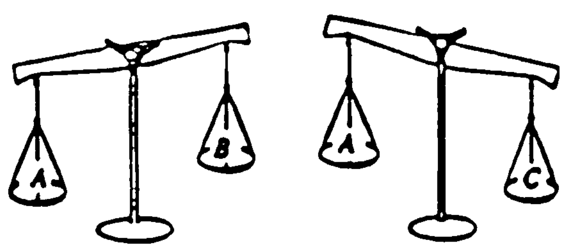
Ⓒ C

Ⓓ D



16. Chen had $10 to buy a model plane, glue, and paint as shown above. At which of the following times could an estimate have been used instead of exact numbers?

- (A) When Chen tried to decide whether or not he had enough money to buy the airplane, glue, and paint
- (B) When the clerk entered each amount into the cash register
- (C) When the clerk told Chen how much he owed
- (D) When Chen counted his change



17. The weights of three objects were compared using a pan balance. Two comparisons were made as shown in the figure above. Which object is the heaviest?

- (A) A
- (B) B
- (C) C
- (D) Not enough information is given.

18. From a shipment of 500 batteries, a sample of 25 was selected at random and tested. If 2 batteries in the sample were found to be dead, how many dead batteries would be expected in the entire shipment?

- (A) 10
- (B) 20
- (C) 30
- (D) 40
- (E) 50

19. The census showed that three hundred fifty-six thousand, ninety-seven people lived in Middletown. Written as a number, that is

- (A) 350,697
- (B) 356,097
- (C) 356,907
- (D) 356,970

20. Steve was asked to pick two marbles from a bag of yellow marbles and blue marbles. One possible result was one yellow marble first and one blue marble second. He wrote this result in the table below. List all of the other possible results that Steve could get.

y stands for one yellow marble

b stands for one blue marble

First Marble	Second Marble
y	b

42, 51, 49, 58, 56, . . .

If the pattern in the list above continues, what will be the next number after 56?

(A) 54

(B) 63

(C) 64

(D) 65

(E) 67

Jill needs to earn $45.00 for a class trip. She earns $2.00 each day on Mondays, Tuesdays, and Wednesdays, and $3.00 each day on Thursdays, Fridays, and Saturdays. She does not work on Sundays. How many weeks will it take her to earn $45.00?

Answer: ______________________________

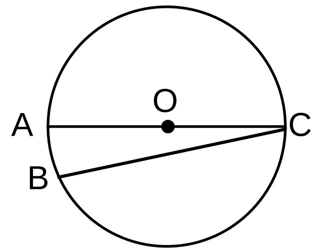
Christy has 88 photographs to put in her album. If 9 photographs will fit on each page, how many pages will she need?

Ⓐ 8

Ⓑ 9

Ⓒ 10

Ⓓ 11



Point O is the center of the circle above. Line segment AC is a diameter of the circle. Line segment BC does not pass through the center of the circle. Which of the following is true?

- (A) AC is longer than BC.
- (B) BC is longer than AC.
- (C) AC and BC are the same length.
- (D) BC is twice as long as OA.
- (E) The lengths of AC and BC change, depending on how this piece of paper is turned.

There are 50 hamburgers to serve 38 children. If each child is to have at least one hamburger, at most how many of the children can have more than one?

- (A) 6
- (B) 12
- (C) 26
- (D) 38

Appendix F (continued)

Revised Test Items

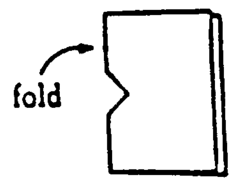
In item 6, the parenthetical statement “1 square year = 9 square feet” was printed as “1 square yard = 9 feet,” and may have impacted interpretation of the question by some students. The item is printed correctly in this appendix.

1. Mack's company sold six billion hamburgers. How many millions is that?

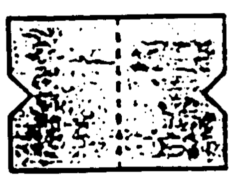
- (A) 6,000,000
- (B) 600,000
- (C) 60,000
- (D) 6,000
- (E) 600

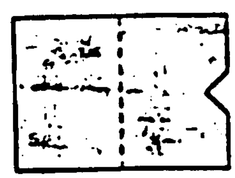
2. In a bag of marbles, 1/2 are red, 1/4 are blue, 1/6 are green, and 1/12 are yellow. If you take a marble from the bag without looking, it is most likely to be

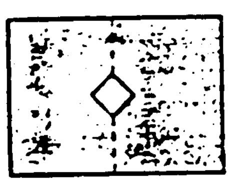
- (A) red
- (B) blue
- (C) green
- (D) yellow

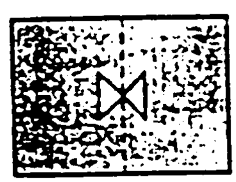


3. If you fold a sheet of paper once and cut out a piece as shown above, what will the unfolded sheet of paper look like?

(A) 

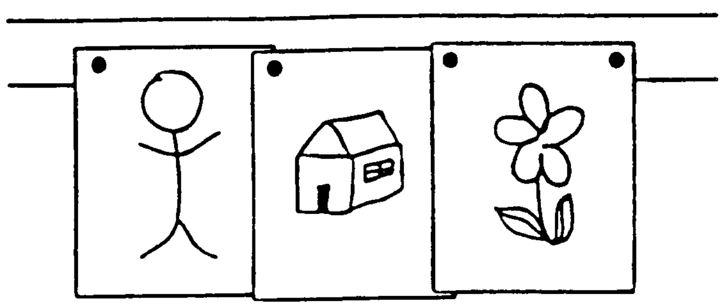
(B) 

(C) 

(D) 

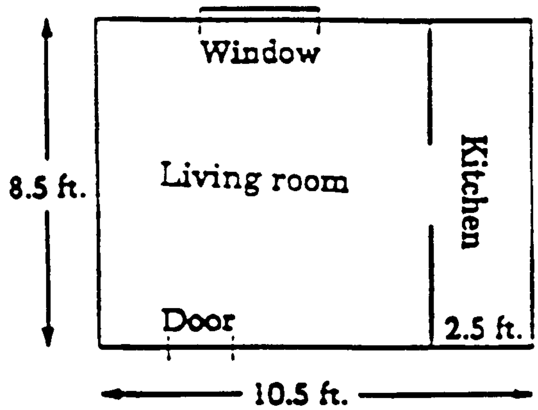
4. Raymond has to buy paper to print 28 copies of a report. He needs 64 sheets of paper for each report. There are 500 sheets of paper in each package. How many whole packages of paper must Raymond buy?

Answer: ______________________



5. The principal wants to hang student pictures in the auditorium. Pictures next to each other would share a tack as shown above. How many tacks does she need to hang 28 pictures?

- (A) 27
- (B) 28
- (C) 29
- (D) 56



6. Chris wants to put wall-to-wall carpet in the rectangular living room, as shown above. Without the kitchen, how much carpet will she need, to the nearest square yard? (1 square yard = 9 square feet)

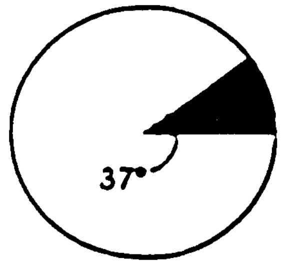
- (A) 8
- (B) 10
- (C) 11
- (D) 19
- (E) 22

7. Harriet, Jim, Roberto, Maria, and Willie ran for president of their 8th-grade class. One of them won. Who is president?

1. The president now was vice president last year and lives on Vine Street.
2. Willie is vice president now.
3. Jim and Maria live on Cypress Street.
4. Roberto was not vice president last year.

- (A) Jim
- (B) Harriet
- (C) Roberto
- (D) Maria
- (E) Willie

RADIO SALES



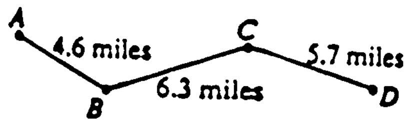
8. The circle represents the total number of 2,675 radios that Mrs. Jones sold. The shaded area represents the number of radios she sold to schools. Approximately how many radios did she sell to schools?

- (A) 70
- (B) 275
- (C) 985
- (D) 25,880
- (E) 98,420

Puppy's Age	Puppy's Weight
1 month	10 lbs.
2 months	15 lbs.
3 months	19 lbs.
4 months	22 lbs.
5 months	?

9. Mike weighs his puppy every month to see how much the puppy has gained. How much will the puppy weigh at five months if the pattern above continues?

- (A) 30 lbs.
- (B) 27 lbs.
- (C) 25 lbs.
- (D) 24 lbs.



10. Carol wants to travel from A to D on the map shown above. To estimate the total distance, she rounds off each of the given distances to the nearest mile. Which of the following shows her work?

- (A) 4 + 6 + 5 = 15
- (B) 5 + 6 + 5 = 16
- (C) 5 + 6 + 6 = 17
- (D) 5 + 7 + 6 = 18

$54 > 3 \times \square$

11. What number could you put in the □ to make the number sentence true?

Answer: ______________________________

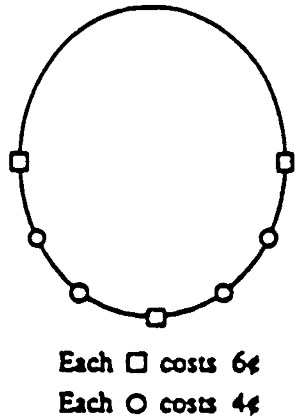
Write another number that could make it true.

Answer: ______________________________

12. Lee delivers □ newspapers each day. How many newspapers does he deliver in 5 days?

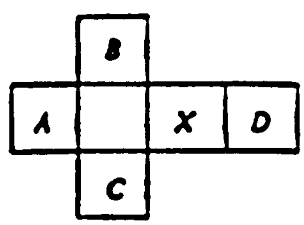
- (A) 5 + □
- (B) 5 x □
- (C) □ + 5
- (D) (□ + □) x 5

13. In a book about dinosaurs Pat read that a dinosaur was estimated to be 80 feet long, rounded to the nearest 10 feet. The dinosaur could have been 80 feet long, but it also could have been ______ feet long.



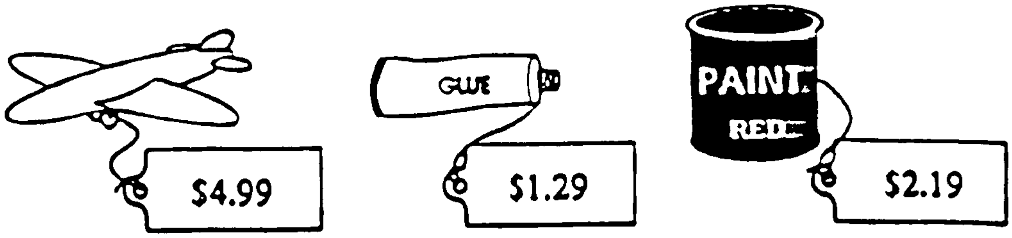
14. How much does the necklace above cost if the string does not cost anything?

- (A) 10¢
- (B) 24¢
- (C) 28¢
- (D) 34¢



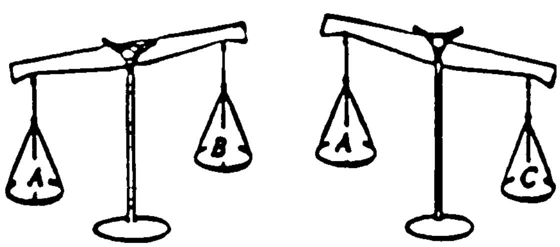
15. If you folded the squares above to form a cube with X on the bottom, which letter would be on top?

- (A) A
- (B) B
- (C) C
- (D) D



16. Chen had $10 to buy a model airplane, glue and paint as shown above. When could Chen have used an estimate instead of exact numbers?

- (A) When Chen tried to decide whether or not he had enough money to buy the airplane, glue, and paint
- (B) When the clerk entered each amount into the cash register
- (C) When the clerk told Chen how much he owed
- (D) When Chen counted his change



17. Sandra compared the weights of three objects using a pan balance. She made two comparisons, as shown above. Which object is the heaviest?

- (A) A
- (B) B
- (C) C
- (D) Not enough information is given.

18. Mr. Richards received a shipment of 500 skateboards. He selected a sample of 25 and checked them. He found 2 broken skateboards in the sample. How many broken skateboards should he expect to find in the entire shipment?

Ⓐ 10

Ⓑ 20

Ⓒ 30

Ⓓ 40

Ⓔ 50

19. Janet's video game score was three hundred fifty six thousand, ninety seven. Written as a number, that is

Ⓐ 350,697

Ⓑ 356,097

Ⓒ 356,907

Ⓓ 356,970

20. Steve had a bag with yellow and blue marbles in it. He took out two marbles. The first marble was yellow, and the second marble was blue. He wrote this result in the table below. List all of the other possible results that Steve could get.

y stands for one yellow marble

b stands for one blue marble

First Marble	Second Marble
y	b

Appendix G

Tables 36-42

Frequency Characteristics of Test Subjects

Table 36

Accuracy-Test and Speed-Test Samples: Gender

	Frequency	Percent
Accuracy-test sample		
Male	479	46.5
Female	552	53.5
Valid cases: 1031	Missing cases: 0	
Speed-test sample		
Male	68	47.6
Female	75	52.4
Valid cases: 143	Missing cases: 0	

Table 37

Accuracy-Test and Speed-Test Samples: Ethnicity

	Frequency	Percent
Accuracy-test sample		
Asian American	163	15.8
African American	192	18.6
Latino	365	35.4
White	265	25.7
Other	30	2.9
Declined to state	16	1.6
Total	1031	100.0
Valid cases: 1015	Missing cases: 16	
Speed-test sample		
Asian American	7	4.9
African American	28	19.6
Latino	97	67.8
White	9	6.3
Other[a]	—	—
Missing	2	1.4
Total	143	100.0
Valid cases: 141	Missing cases: 2	

Note. Different agencies use different categorical descriptors for ethnicity. The original descriptors from each agency have been retained in the table.

[a] For this sample, there were no students in this category.

Table 38

Accuracy-Test and Speed-Test Samples: Non-Native English Status

	Frequency	Percent
Accuracy-test sample		
Yes, that was first language	473	75.8
No, that was not first language	146	23.4
Missing	5	.8
Total	624	100.0
Valid cases: 619	Missing cases: 5	
Speed-test sample		
Yes, that was first language	103	93.6
No, that was not first language	7	6.4
Total	110	100.0
Valid cases: 110	Missing cases: 0	

Table 39

Accuracy-Test and Speed-Test Samples: Booklet

	Frequency	Percent
Accuracy-test sample		
Booklet A	525	50.9
Booklet B	506	49.1
Total	1031	100.0
Valid cases: 1031	Missing cases: 0	
Speed-test sample		
Booklet A	76	53.1
Booklet B	67	46.9
Total	143	100.0
Valid cases: 143	Missing cases: 0	

Note. Booklet A = all original items; Booklet B = all revised items.

Table 40

Accuracy-Test and Speed-Test Samples: Level of Math Class

	Frequency	Percent
Accuracy-test sample		
ESL	70	6.8
Low	53	5.1
Average	405	39.3
High	224	21.7
Algebra	157	15.2
Honors algebra	122	11.8
Total	1031	100.0
Valid cases: 1031	Missing cases: 0	
Speed-test sample		
ESL	97	67.8
Low[a]	—	—
Average[a]	—	—
High	25	17.5
Algebra	21	14.7
Honors algebra[a]	—	—
Total	143	100.0
Valid cases: 143	Missing cases: 0	

[a] For this sample, there were no students in this category.

Table 41

Accuracy-Test Sample: ESL Code Assigned by School

ESL code	Frequency	Percent
Initially fluent in English	49	4.8
Beginning bilingual	12	1.2
Intermediate bilingual	14	1.4
Advanced bilingual	5	.5
LEP	95	9.2
Awaiting redesignation	32	3.1
Preparing for redesignation	23	2.2
Redesignated fluent	90	8.7
No code assigned	711	68.8
Total	1031	100.0
Valid cases: 1031	Missing cases: 0	

Table 42

Accuracy-Test and Speed-Test Samples: Lunch Participation Code Assigned by School

	Frequency	Percent
Accuracy-test sample		
Free lunch	283	27.4
Reduced pay	13	1.3
Full payment	24	2.3
Nonparticipant	120	11.6
AFDC	75	7.3
Missing	516	50.0
Total	1031	100.0
Valid cases: 1031	Missing cases: 0	
Speed-test sample		
Free lunch	68	47.6
Reduced pay	4	2.8
Full payment[a]	—	—
Nonparticipant	35	24.5
AFDC	6	4.2
Missing	30	21.0
Total	143	100.0
Valid cases: 113	Missing cases: 30	

[a] For this sample, there were no students in this category.

Appendix H

Table 43

Design

Table 43

Design of Large-Scale Field Test

No. of items	Item type	Form A	Form B
10	Linguistically complex	Original	Revised
10	Linguistically complex	Revised	Original
5	Non-linguistically complex	Original	Original

Appendix I

Language Background Questionnaire

Language Questionnaire

1. Do you speak a language other than English? ❒ Yes ❒ No

2. If you speak a language other than English:

   a. What is that language? ______________________________

   b. Was that the first language you learned when you were a child? ❒ Yes ❒ No

   c. If not, how old were you when you began speaking that language? _______

   d. How often do you speak that language: (Check one for each item.)

		Always or most of the time	Sometimes	Never or hardly ever
i.	with your parents?	❒	❒	❒
ii.	with your grandparents?	❒	❒	❒
iii.	with your brothers and sisters?	❒	❒	❒
iv.	with your friends away from school?	❒	❒	❒
v.	with your friends at school?	❒	❒	❒

   e. How well do you: (Check one for each item.)

		Very well	Fairly well	Not well	Not at all
i.	understand that language?	❒	❒	❒	❒
ii.	speak that language?	❒	❒	❒	❒
iii.	read that language?	❒	❒	❒	❒
iv.	write that language?	❒	❒	❒	❒

   f. Do you prefer talking about school subjects:
   ❒ in English? ❒ in your other language?

3. For each of the subject areas below, how easy has it been for you in the past to understand your teacher's explanations? (Check one for each subject.)

	Very easy	Fairly easy	Fairly difficult	Very difficult
Math	❒	❒	❒	❒
Science	❒	❒	❒	❒
Social Studies/History	❒	❒	❒	❒

4. For each of the subject areas below, how easy has it been for you in the past to understand your textbooks? (Check one for each subject.)

	Very easy	Fairly easy	Fairly difficult	Very difficult
Math	❐	❐	❐	❐
Science	❐	❐	❐	❐
Social Studies/History	❐	❐	❐	❐

5. For each of the subject areas below, how easy has it been for you in the past to understand questions on tests? (Check one for each subject.)

	Very easy	Fairly easy	Fairly difficult	Very difficult
Math	❐	❐	❐	❐
Science	❐	❐	❐	❐
Social Studies/History	❐	❐	❐	❐

6. How well do you: (Check one for each item.)

		Very well	Fairly well	Not well	Not at all
i.	understand spoken English?	❐	❐	❐	❐
ii.	speak English?	❐	❐	❐	❐
iii.	read English?	❐	❐	❐	❐
iv.	write English?	❐	❐	❐	❐

Appendix J

Notes From Test Administrators

Observation Notes on Language Background Study
Protocol Interview
School Number 2

<u>1st Student</u>: Student A MAE Above Average Math

Start Time 9:23 a.m.
Picked
1. Revised
2. Original
3. Original
4. Revised

Very fast reader.
End Time 9:33 a.m.

<u>2nd Student</u>: Student B MAE Below Average Math

Stan Time 9:50 a.m.
Picked
1. Revised
2. Revised—listen to tape to confirm
3. Original
4. Original

Very willing to give comments. Very explanatory as to why she chose the item that she chose. Almost too willing. Wanted to be very much a representative of her age group. Said things like . . . "I just think kids our age could relate to this more." Referred to the parentheses in the last pair of items as "apostrophe's." Trying very hard to impress me.
End Time 10:03 a.m.

<u>3rd Student</u>: Student C AAL Below Average Math

Start Time 10:15 am.
Picked
1. Revised—listen to tape to confirm
2. Revised
3. Revised
4. Original

Very shy and overwhelmed by interview. Quiet and unwilling to speak (difficult to get a sample of her language skills). Very low SES. Had trouble pronouncing the words "reference" and "approximately," which appear in the first pair of items (in the original item). She picked all revised except for in the last pair of items. She chose the original because she said it was more challenging and it held her interest but if she were in a rush she would of picked the easier one. She was very overwhelmed when I asked her for her lunch order—very surprised that she would get a free lunch.
End Time 10:25 a.m.

4th Student: Student D MAE Average Math

Start Time 10:29 a.m.

Picked
1. Revised
2. Revised
3. Original
4. Revised

Very pensive but very willing to think hard before she made her choices. Gave each item a hard look over. She said that she picked the revised item on the last set because it did not have 1 square yard = 9 square feet like the original but had 1 square yard = 9 feet. The phrase square feet intimidated her, she said. I do believe this was not intentional on our part. When I asked if she spoke any other languages at home she said that she spoke Slovakian. She said she was born in Czechoslovakia and has been in the United States for about 8 years. More details on the tape.
End Time 10:40 a.m.

5th Student: Student E Asian Above Average

Start Time 10:53 a.m.

Picked
1. Original
2. Revised
3. Revised
4. Revised

Very self conscious of his limited English proficiency. Got very frustrated with the first pair of items. Did not understand what either of them were asking him to do. Sat for 3 to 5 minutes trying to figure it out even with me asking him to go on. Said he picked the original on the first pair because he felt that if he had a dictionary he could understand more what the item was asking him to do. Was very flustered so I had to encourage him to go on. He also was determined to solve each task and tell me the answer (hardly any of the other students did this). On the second pair of items he told me that he did not understand words like “shipment” and “selected” in the revised item and words like “selected” and “random” in the original item. However, for the third and fourth pair of items he said that he understood the questions and all the words perfectly. He was very defeated when we closed the interview despite my positive encouragement.
End Time 11:12 a.m.

6th Student: Student F AAL Above Average Math

Start Time 11:33 a.m.

Picked
1. Revised
2. Revised
3. Original
4. Original

Very outspoken. Had run for class president that day and given a speech in front of the entire student body. Very confident; very upper to middle class SES. On the fourth pair of items she was the first student that read the phrase in parentheses besides the girl who referred to them as apostrophes. However, she misread closet for closest.

7th Student: Student G Asian Above Average

Start Time 11:50 a.m.

Picked
1. Revised
2. Revised—listen to tape to confirm
3. Original
4. Original

He picked the revised item in round one because he said he liked hamburgers and could identify with that question more. He said that he picked the original in round four (the last pair of items—the carpet item—just because he read it first). He felt that the two items were very much the same, so he picked the one he read first.

8th Student: Student H AAL Average Math

Start Time 1:05 p.m.

Picked
1. Revised
2. Revised
3. Revised
4. Original

On the first round she picked the revised item because she said she related to it better. She also selected an answer for me. She also said that she related to the revised item more on the second round and third round as well. However, she chose the original item on the fourth and final round.
End Time 1:17 p.m.

9th Student: Student I Spanish Average Math

Start Time 1:35 p.m.

Picked
1. Revised
2. Revised
3. Revised
4. Original

She picked the revised item on the first round because she said it was easier. She picked the revised on the second round because she said she could understand what broken

skateboards were but she did not understand what a dead battery was or the concept of a battery being dead. On the fourth round (on the original item) she did not know how to pronounce the name Chris....she asked me if this word was a name and then continued and actually ended up picking hat item over the revised. She also mistook the word closet for closest.
End Time 1:45 p.m.

<u>10th Student</u>: Student J Spanish Above Average Math

Start Time 1:50 p.m.

Picked
1. Revised
2. Revised
3. Original
4. Original

She picked the revised item in the first round because it had Mack's company in it—she related to the concept of a company; this was a word/concept that she was familiar with. She picked the revised item in the second round because she said that "toys" (skateboards) were easier for [her] to understand than batteries. She picked the original item in the last and final round because she said it was more interesting to her.
End Time 2:01 p.m.

<u>11th Student</u>: Student K Spanish Below Average Math

Start Time 2:05 p.m.

Picked
1. Revised
2. Revised
3. Revised
4. Revised

Limited English skills. He told me right away in a very flippant manner that he "No hables Ingles!" In the first round he did not understand the word "approximately" in the original item but he did not want to admit it to me. I had to keep on asking him if there were any words he did not understand or that he found difficult. He did not understand what the original item was asking him at all because of the way it was worded and the language used in it. In the second round I asked him again if there were any words he did not understand. He said no but I could tell that he was trying to prove himself to me. For the third round he picked the revised because he said the picture (armchairs) was easier for him to understand (visualize) then the pictures with the tasks. He picked the revised on the fourth and final round because he said that it was just plain easier.
End Time 2:19 p.m.

3:00 p.m. Bought pizza and distributed it.

Appendix K

Tables 44-88

Results of Analyses

for the Language Background Questionnaire (LBQ)

Accuracy Test Discussion

Table 44 presents descriptive statistics for the LBQ questions 1, 2a, 2b, and 2c. As the data in this table indicate, out of the total 1031 students who participated in this study, 624 or 61% indicated that they use a language other than English at home and/or with friends. Of the 624 students who indicated the use of a language other than English, for 473 or 75.8% of them, that language was the first language they spoke and for only 146 students (23.4%) it was not the first language. A variety of languages were spoken by the students. However, many of these languages occurred only occasionally and were combined and placed into the "Other" category. Table 44 shows the five languages with the highest frequency of usage by students, plus the Other category which includes other languages spoken by small numbers of students. Among the languages listed in Table 44, Spanish has the highest level of usage. Of the 624 students who reported speaking a second language, 376 or 60% of them reported Spanish as their second language. The second most frequently reported language is Korean, which was used by 60 students (9.6%); next is Chinese with 32 (5.1%); then Farsi with 25 (4%); and Filipino, with a frequency of 19 (3%).

Questions 2d, 2e, 3, 4, 5, and 6 were Likert-type questions. In item 2d students were asked to rate their use of the "other language" with their "parents," "grandparents," "brothers and sisters," and "friends" on a 3-point scale ranging from *always or most of the time* to *never or hardly at all*. Table 45 presents frequencies and percentages of students' responses to each of the three options under different questions for item 2d. As Table 45 indicates, the frequency with which students use their native languages decreases dramatically as we move from "parents and grandparents" to "friends at school." Based on the results presented in this table, students speak their "other language" more often with their parents and grandparents than with their friends in school.

In item 2e, students rated their understanding, speaking, reading, and writing of the "other language" on a 4-point scale raging from *very well* to *not at all*. Table 46 shows frequencies and percentages of students' responses to questions under item 2e. Results in Table 46 indicate that students in this study understand the "other language" much better than they speak, read, or write that language. For example, 68.6% of the students indicated that they understand that language "very well" compared with only 32.1% who indicated that they also write that language "very well."

Table 47 shows frequencies and percentages of students' answers to item 3 of the LBQ, "How easy has it been for you in the past to understand your teacher's explanations" in the areas of math, science, and social sciences. As this table indicates, it has been easier for students to understand teachers' explanations in science and social sciences than math. For example, 422 (40.9%) of the students indicated that they easily understood their teacher's explanations in math as compared with 478 students (46.4%) who indicated that they understood their teachers' explanations in science very easily and 532 students (51.6%) who understood their teacher's explanations very easily in social studies.

Similarly, Table 48 depicts frequencies and percentages of students answering item 4 in the LBQ, "How easy has it been for you in the past to understand your textbooks?" For these questions also students indicated that textbooks for science and social sciences have been easier to understand.

The results of analyses for item 5 in the LBQ, "How easy has it been for you in the past to understand questions on tests?" indicate that students understand questions on tests for the three subject areas (math, science, and social sciences) at about the same level, but slightly better for science and social sciences than for math. Finally, Table 49 shows frequencies and percentages for item 6, "How well do you [use English]?" The results indicate that students in general believe they understand, speak, read, and write English "very well," with slightly higher responses in the "understanding" category.

As indicated earlier, items 2d, 2e, 3, 4, 5, and 6 were Likert-type items; thus, computing and reporting means and standard deviations for those items would be more appropriate than reporting simple frequencies and percentages. This is because means and standard deviations combine all response options and provide a more comprehensive measure for each item. To do this, we assigned numbers 1 to 4 to the different response options in these items.

For question 2d, we assigned 3 to *always or most of the time,* 2 to *sometimes,* and 1 to *never or hardly ever*. For question 2e, we assigned 4 to *very wel*l, number 3 to *fairly well,* 2 to *not well,* and 1 to *not at all*. Similarly, for questions 3 to 6, we assigned numbers 1 to 4 to very easy, fairly easy, fairly difficult, and *very difficult* respectively. Tables 50 to 54 summarize the results of our descriptive statistics for question 2d. Table 50 presents means, standard deviations and numbers of cases for students' answer to question 2di in the LBQ by students' background characteristics. The first part of this table shows means and standard deviations for all non-native English speakers, and the second part displays information by gender. As these data show, males (M = 2.72, SD = .52) and females (M = 2.70, SD = .52) reported about the same level of speaking the "other language" with their parents. The third section of the table presents the results by ethnicity. As this table indicates, Asian American students (with a mean of 2.63) speak more of the "other language" than either Latino (M = 2.74) or White students (M = 2.74). There were not enough students in the African American and Other categories to make meaningful comparisons. Following results by ethnicity, the results of analyses for question 2di by ESL codes are presented. Students in the No Code category indicated the least usage of the "other language" with their parents.

The results by type of math class are reported next. There were initially six categories of math classes: (a) ESL level, (b) Low, (c) Average, (d) High, (e) Algebra, and (f) Honors Algebra. Because of the small number of cases in some of these categories, we decided to combine categories ESL and Low into the new composite "Low"; Average and

High categories were not changed; and Algebra and Honors Algebra were combined into the new composite “Algebra.” The students in the four groups of math classes reported about the same level of usage of the “other” language with their parents, with the Low category using slightly more “other language” with their parents.

The last part of Table 50 presents the results of our analyses for question 2di of the LBQ by the free lunch program. In this categorization, the majority of students fell within the first and last categories, mainly, Free Lunch versus No Free Lunch. Some students in both categories reported no usage of “other language” with their parents; students in the Free Lunch program group used the “other language” with their parents slightly more than the No-Free-Lunch group did.

Table 51 presents results of the analyses for LBQ question 2dii, “How often do you speak that language with your grandparents.” The means and standard deviations of responses to this questions were categorized by ethnicity, ESL codes, type of math class, and school lunch program as in Table 50. Among the categories analyzed by different background variables, Beginning ESL within the categories by ESL codes indicated highest usage of the “other language” with the grandparents. However, few students indicated *always or most of the time* usage of the “other language” with their parents.

Similarly, Table 52 presents the results of descriptive analyses for LBQ question 2diii, “How often do you speak that language with your brothers and sisters?” A comparison of Table 52 with Tables 50 and 51 indicates the students in general tend to speak less of the “other language” with their brothers and sisters than with their friends. The mean for all non-native speakers for using “other language” with parents in Table 50 was 2.71 (SD = .52) and in Table 51 was 2.70 (SD = .65) as compared with a mean of 1.98 (SD = .69) for using “other language” with their brothers and sisters. Similar to the data reported in the previous tables, males and females reported about the same level of speaking the “other language” with their brothers and sisters. For ethnic differences on this variable, the means range from 1.87 for Asian students (SD = .71) to 2.67 for African American students (SD = .58). However, when ethnic groups with small numbers of subjects are removed, the difference becomes negligible. Some differences can also be seen in the means across categories of ESL, type of math class, and school lunch program. Again, when categories with small numbers of subjects are eliminated, the size of differences decreases.

Table 53 presents means, standard deviations, and number of subjects for students’ responses to LBQ item 2div, “How often do you speak that language with your friends away from school?” As this table indicates, the average for all non-native English speakers is 1.77 (SD = .70). The results in this table show no gender differences, but there are some differences at the levels of ethnicity, type of class, and school lunch program. However, as was seen in the previous tables, when categories with small numbers of subjects are deleted, the size of the differences decreases.

Table 54 presents results similar to those presented in Table 53, for speaking "that language with your friends at school." As this table indicates, the mean for the total non-native English speakers is 1.70, which is slightly lower than the corresponding mean in Table 53 (M = 1.77). However the trends of mean differences are very similar to those in Table 53. Comparison of Tables 50 to Table 54 reveals that students speak the "other language" more with their parents and grandparents than with their brothers and sisters and friends.

Table 55 and Table 56 report means and standard deviations for LBQ item 2ei, "How well do you understand your (native) language?" by gender, ethnicity, ESL codes, type of math class (combined categories), and school lunch program for the Accuracy Test and the Speed Test samples. As these tables indicate,[1] males expressed more understanding of their native language (M = 3.62, SD = .54) than females (M = 3.07, SD = .52). This difference is not statistically significant, $F_{1,467}$ = 3.21, p = .074. Different ethnic groups indicated relatively high but slightly different levels of understanding of their native language. White students had a mean of 3.74 (SD = .45) from the maximum possible score of 4.0. Latino students had the highest mean (M = 3.71, SD = .51); next were White students, followed by Asian American students. African American students had a lower mean than other groups (M = 3.33, SD = .58). The differences between means for the ethnic groups are significant, $F_{4,454}$ = 2.41, p = .05.

Table 55 and Table 56 present means and standard deviations for different ESL groups. These tables indicate, all groups of ESL students understand their native language; however, there are some small differences. The Initially Fluent students obtained the lowest mean among others (M = 3.44, SD = .56), and Awaiting/Redesignation obtained the highest mean (M = 3.77, SD = .43). The difference between means of ESL groups is statistically significant, $F_{5,463}$ = 3.92, p = .01.

The results by type of math class (Table 55 and Table 56) indicate that students in different math classes had equally good levels of understanding of their native language. The lowest mean is 3.57 (SD = .51) which belongs to the Algebra category, and the highest mean is 3.71 (SD = .48) which belongs to the Average category. These differences are not statistically significant, $F_{3,465}$ = 1.80, p = .15.

The last part of Tables 55 and 56 shows means and standard deviations for students by school lunch program. Again, students in all categories of lunch program indicated a good understanding of their native language, with a few minor differences. Students in the free lunch program obtained a slightly higher mean (M = 3.71) than other students. These differences are not statistically significant, $F_{5,463}$ = 2.12, p = .06.

Table 57 (Accuracy Test) and Table 58 (Speed Test) summarize the results of descriptive analyses for the second question under item 2e, "How well do you speak your (native) language?" The results for this question are similar to those reported for

---

[1] Due to imbalances in cell sizes for the ANOVA tables for this and several following tables, the reliability of these statistical tests is questionable.

question 2ei. In general, all students reported high-level ability in speaking their native language. There are, however, a few minor differences. For example, unlike results for question 2ei, now males and females indicated the same level of ability in speaking their native language. Similarly, Table 59 (Accuracy Test) and Table 60 (Speed Test) report the results of our analyses for question 2eiii, "How well do you read your (native) language?" There are relatively major differences between the results presented in these tables and those reported in Tables 55 and 56 and Tables 57 and 58. In general, based on the students' self-reports, they are not as good in reading their native language as they are in understanding and speaking it. In Table 55 and Table 56 ("understanding") the lowest mean was 3.07, and in Table 57 and Table 58 ("speaking") the lowest mean was 3.12, whereas in Table 59 the lowest mean was 2.00. (Note, however, that this lowest mean was for a group of only three students.)

Table 61 and Table 62 present means and standard deviations for students' responses to item 2eiv of LBQ, "How well do you write your (native) language?" The means in these tables in general are smaller than the means in the previous tables and indicate that non-native English speakers are less capable of writing in their native language than of understanding, speaking and reading. However, the standard deviations are higher than those in the previous tables, which indicates more heterogeneity between subjects in their response to this question. A comparison of means across different groups of students based on their background variables reveals that in some cases very large and significant differences exist. For example, male students are less capable of writing their native language (M = 2.77, SD = .98) than female students (M = 3.01, SD = .98). The difference between the two group means is statistically significant, $F_{1,461} = 6.86$, $p = .01$. The difference in means by categories of ethnicity is also evident. The largest mean among ethnic groups is 3.06 (SD = .92) for Latino students and the smallest mean is 2.5 (SD = 1.09) for White students. The differences between means of ethnic groups are significant, $F_{4,448} = 5.79$, $p = .01$. The largest group difference in means can be seen for students categorized by ESL codes. Students in the Initially Fluent category have a mean of 2.33 (SD = .95) as compared with the mean of 3.26 (SD = .90) for students in the Other category (the Beginning ESL category had few students and could not be used for a valid comparison). The results of analysis of variance comparing means of students' responses by categories of ESL revealed a significant difference, $F_{5,457} = 7.07$, $p = .01$. Table 61 also shows differences in means for students grouped by type of math class and school lunch program. Students in "low" classes tend to have higher means (M = 3.12, SD = .98) than those in the higher classes (for Algebra, M = 2.69 and SD = .97) $F_{3,459} = 3.18$, $p = .02$. For categories of free lunch program, students seem to have about the same level of performance in the two main categories. For the Free category M = 2.95 (SD = .95) and for the No Lunch Code category, M = 2.95 (SD = 1.03). The mean differences among the lunch groups are not statistically significant, $F_{5,457} = 1.89$, p = .10.

Table 63 (Accuracy Test) and Table 64 (Speed Test) present means and standard deviations for the students' responses to item 3i in the LBQ, "In the subject area Math, how easy has it been for you in the past to understand your teacher's explanation?" As these tables indicate, the means are generally high for this question for the different subgroups that were formed based on the different background variables. The means range from 3.10 for African American to 3.41 for Asian American. These means are very similar, which indicates that students' background variables did not have much impact on their answers to this question.

Table 65 (Accuracy Test) and Table 66 (Speed Test) show means and standard deviations for item 3ii in the LBQ. In this question students were asked to report how easy it has been in the past to understand their teacher's explanation of science lessons. Like the results presented in Tables 63 and 64, means are relatively high, indicating that students have a good understanding of their teachers in the area of science. Furthermore, there is not much difference between the subgroups that were formed based on different background variables. Similarly, Table 67 (Accuracy Test) and Table 68 (Speed Test) summarize the results of descriptive analyses for LBQ item 3iii. These results are very similar to those reported in Tables 63 and 64 and Tables 65 and 66: Students reported a high level of understanding of their teacher's explanation of social studies/history.

Table 69 (Accuracy Test) and Table 70 (Speed Test) depict means and standard deviations for the responses to item 4i: "In the subject area Math, how easy has it been for you in the past to understand your textbooks?" All of the means reported in Table 69 are above 3.0, indicating students reported good understanding of their textbook in math. There are no major differences among the subgroups of students that were formed based on different background variables. Similarly, Table 71 and Table 72 report the results of analyses for science, and Table 73 and Table 74 report similar findings for social studies. These results are very similar with those reported in Table 69 and Table 70 for math.

Tables 75 (Accuracy Test) and 76 (Speed Test), Tables 77 (Accuracy Test) and 78 (Speed Test), and Tables 79 (Accuracy Test) and 80 (Speed Test) report means and standard deviations for students' responses to LBQ item 5 for math, science and social studies, respectively.

Table 81 (Accuracy Test) and Table 82 (Speed Test) present means and standard deviations for answers to LBQ item 6i, "How well do you understand spoken English?" The means are well above 3.5 (and in some cases they are very close to the maximum of 4.0) for most of the categories, although not for some of the ESL categories. For example, the mean for Beginning ESL is 2.43 and for Intermediate/Advanced ESL is 3.19, but the mean for the Initially Fluent and No Code categories is 3.9. Analysis of variance showed no significant results except for the analysis by ESL codes.

Tables 83 (Accuracy Test) and 84 (Speed Test), Tables 85 (Accuracy Test) and 86 (Speed Test), and Tables 87 (Accuracy Test) and 88 (Speed Test) present similar results for proficiency in speaking, reading, and writing English respectively. All different subgroups of students reported high levels of proficiency in speaking, reading, and writing, except for some subgroups of students in the ESL section. Students in the Beginning ESL category always obtained the lowest possible mean score. The results of analyses of variance generally indicate that the subgroups under each of the background variables performed about the same except for the ESL groups.

Speed Test Discussion

Tables that were introduced earlier report frequencies and percentages for the various background questions in the LBQ for the Accuracy Test Study. As Table 39 indicates, 76 or 53.1% of the students answered questions in Booklet A, and 67 or 46.9% of the students answered questions in Booklet B. Table 45 presents frequencies and percentages of students' response to LBQ items 2d, "How often do you speak your (native) language?" The results of our descriptive analyses for this item for the speed-test sample, presented in Table 45, are very similar to the results for the performance-test sample presented in the same table. Based on these results, students in this group speak their native language more with parents and grandparents than with brothers and sisters and friends.

Table 46 shows frequencies and percentages for LBQ item 2e, "How well do you use your (native) language?" for both performance-test and speed-test groups. Most of the students in the speed-test group indicated that they understand, speak, read, and write that language well. However, in general, this group indicated lower proficiency in their native language than the performance-test group.

Table 47 presents the descriptive results for students' responses to LBQ item 3, "How easy has it been for you in the past to understand your teacher's explanations?" for both groups of subjects (performance and speed groups). The data in this table indicate that students have more difficulty understanding their teachers' explanations in math and science than social studies. But in general, students had a fair level of understanding of their teachers' explanations. The performance-test sample reported a slightly higher level of understanding of teacher's explanation in the three topics.

The frequencies and percentages of students' responses to item 4 in the LBQ are reported in Table 48 for both groups of subjects. The results reported in this table indicate that it has been relatively easy for the students to understand their textbooks in math, science and social studies, with math textbooks being slightly more difficult to understand than textbooks in science and social studies. On this question also, the speed-test group indicated a slightly lower level of understanding than the performance-test group.

Finally, Table 49 presents the results for LBQ item 6, "How well do you use English?" The results of descriptive analyses indicate that students believe that they are proficient in English (understanding, speaking, reading, and writing), and that they are more proficient in understanding English than reading or writing. The performance-test group reported a higher level of understanding, speaking, reading and writing in English than the speed-test group.

As mentioned in the performance-test discussion, items 2d, 2e, 3, 4, 5, and 6 in the LBQ were Likert-type items. To have a more comprehensive measure for these items, we computed their means and standard deviations. We assigned numbers 1 through 4 to the different response options. For question 2d, we assigned 3 to *always or most of the time,* 2 to *sometimes,* and 1 to *never or hardly ever.* For question 2e, we assigned 4 to *very well,* 3 to *fairly well,* 2 to *not well,* and 1 to *not at all.* Similarly, for questions 3 to 6, we assigned 1 to 4 to the *very easy, fairly easy, fairly difficult,* and *very difficult* categories of responses respectively. Table 56 presents means, standard deviations, and number of subjects for responses to LBQ items 2ei for the speed-test sample by different student background variables. In these questions students were asked "How well you understand your (native) language?" As the results in Table 56 indicate, students reported a high level of understanding of their native language. Males and females indicated about the same level of understanding (*M* = 3.53, *SD* = .63 for males; *M* = 3.59, *SD* = .64 for females). Different ethnic groups also reported about the same level of understanding of their native language (*M* = 3.40, *SD* = .55 for Asian students; *M* = 3.5, *SD* = .71 for African American students; *M* = 3.57, *SD* = .65 for Latino students; and *M* = 3.67, *SD* = .58 for White students).

There were larger differences in the means for students grouped by levels of ESL codes. The means ranged from 3.14 (*SD* = .95) for Beginning ESL to 3.74 (*SD* = .45) for No Code categories. As the data in Table 56 indicate, type of math class did not have much association with ESL classification. For type-of-math categories, means ranged from 3.56 (*SD* = .53) for "high" to 3.58 (*SD* = .66) for "low" categories. School lunch program seems to have an association with students' response to this item. Means ranged from 3.00 (*SD* = .97) to 4.00 (*SD* = 00). However, when we ignore categories with a small number of subjects, the means seem to be close, and no big differences are seen. Thus, if we ignore Reduced Payment and AFDC categories, the range is 3.40 (No Lunch Code category) to 3.63 (Free category).

Table 58 presents the results of descriptive analyses for item 2eii for the speed-test section of this study. The results of the analyses are reported by different background variables in different parts of the table. The results reported in this table are very similar to the results shown in Table 56, with one difference. For this group of students, the means are in general slightly lower than the means for item 2i, indicating that students expressed more efficiency in understanding their native language than speaking the language.

Similarly, Table 60 depicts means, standard deviations and number of subjects for students' response to LBQ item 2eiii. These results are also similar to those reported in Tables 56 and 58, with one difference. Means in this table are lower than those reported for item 2eii (Table 58) and even more so than those reported in Table 56. This difference indicates that non-native English speakers are far less comfortable reading their native language than understanding or speaking that language.

When students were asked “How well do you write your (native) language?” their response, as presented in Table 62, generally indicated that they feel comfortable in writing their native language. In other words, they feel that they are proficient in writing their native language. Means in Table 62 are considerably lower than those reported in Table 56 for understanding and even lower than those reported for speaking and reading (see Tables 56 to 62).

Table 64 presents the results of descriptive analyses for LBQ item 3i, “In the subject area Math, how easy has it been for you in the past to understand your teacher’s explanation?” It was indicated earlier that students’ responses to LBQ items 3, 4, and 5 ranged from *very easy* to *very difficult* on a 4-point Likert-type scale. The results of analyses shown in Table 64 indicate that students in general have indicated a good level of understanding of their teacher’s explanation in math. However, there are differences across some of the subgroups based on students’ background variables. Gender, ethnicity, English status and type of math class did not seem to have much impact on students’ responses on this question, but ESL codes and school lunch program showed some relationship to students’ responses. In the ESL code category, as Table 64 indicates, the means range from 2.00 (*SD* = .00) to 3.60 (*SD* = .89) with a difference of 1.60 (about 2 standard deviations). However, if we ignore categories with small frequencies, the mean differences decrease. For school lunch program categories, the means ranged from 2.75 (*SD* = .96) to 3.5 (*SD* = .55). Again, if we ignore categories with small frequencies, the differences decrease.

Table 66 presents results for the LBQ item 3ii involving science. These results are similar to those presented in Table 64 for math. In general, students indicated a high level of understanding of teacher’s explanation of science. No major differences were observed for gender, ethnicity, English status, and free lunch program. However, there are some differences across categories of ESL and type of math variables (see Table 66). Similarly, Table 68 presents results for social studies. In this table some ethnic differences can be seen. Asian students had a lower mean than students in other ethnic groups. There were also some differences by the school lunch program categories. As in other cases, excluding categories with small frequencies reduces the size of these differences (see Table 68).

Table 70 shows means, standard deviations and number of subjects for students’ responses to LBQ item 4i (understanding math textbook) by different categories of students. There is little variation among means across categories of gender, English status, and type of math class. For categories of ethnicity, means range from 2.57 (*SD* = .79) for Asian students to 3.24 (*SD* = .75) for Latino students. However, in the Asian category, there are too few students. For ESL, when categories with small frequencies are omitted, no major differences remain. For the categories of school lunch program, means range from 2.75 (*SD* = .96) to 3.33 (*SD* = .82). Ignoring those categories with a small number of subjects decreases the size of the differences (see Table 70).

Tables 72 and 74 present results for LBQ items 4ii and 4iii in the areas of science and social studies. Results for both tables indicate that students report having a high level of understanding of their textbooks in science and social science. No big differences were found between means of the different subgroups except for subgroups by ESL codes, and the category asking for information about understanding social science texts, in which native and non-native English speakers disagreed (3.64 and 2.99 respectively). Students in Beginning ESL classes had a lower mean for understanding science and social studies textbooks than for understanding their math textbooks (see Tables 72, and 74).

Table 76 shows summary statistics for responses to LBQ item 5i, "In the subject area Math, how easy has it been for you in the past to understand questions on tests?" Table 78 presents results from the same question as applied to science, and Table 80 depicts results for social studies. The results of the analyses reported in these three tables indicate in general that students in this part of the study understood their test questions well. However, for some subgroups of students, minor differences were observed. Tables 76, 78, and 80 report information for ESL code, type of math class and school lunch program.

Students' responses to LBQ item 6 have been summarized in Tables 82, 84, 86, and 88. The results in general indicate high-level proficiency in understanding, speaking, reading, and writing English; however, there are slight but systematic decreases in mean scores for reading and writing, which indicate that students in this group felt more comfortable in understanding and speaking than in reading and writing. Furthermore, within each table (i.e., subject area), there are some group differences. Major differences can be seen in subgroups of ESL. For example, in Table 82, the Beginning ESL category had a mean of 2.72 (SD = .53) as compared with a mean of 3.92 (SD = .40) for the No Code category. These trends can be seen in all four tables for LBQ item 6.

Table 44

Accuracy-Test Sample: Participants Who Speak Languages Other Than English

	Is this your first language?		
Language	Yes	No	Missing
Spanish	291	82	3
Korean	57	2	1
Chinese	27	5	
Farsi	18	7	
Filipino	13	6	
Other	67	44	1
Total	473	146	5

Note. 624 students reported speaking a second language. Among the 143 students in the Speed sample, 110 reported speaking a language other than English; of these, 103 reported that this was their first language. In 94 of these cases, the language was Spanish.

Table 45

Accuracy-Test and Speed-Test Samples: Responses from Non-Native Speakers of English to the Question "How Often Do You Speak Your (Native) Language?" (Item 2d)

	Always or most of the time	Sometimes	Never or hardly at all	Missing
Accuracy-test sample				
With your parents?	346	105	15	5
	73.5%	22.3%	3.2%	1.1%
With your grandparents?	363	39	47	22
	77.1%	8.3%	10.0%	4.7%
With your brothers and sisters?	101	232	112	26
	21.4%	49.3%	23.8%	5.5%
With your friends away from school?	70	211	173	17
	14.9%	44.8%	36.7%	3.6%
With your friends at school?	64	189	202	16
	13.6%	40.1%	42.9%	3.4%
Total: 471				
Speed-test sample				
With your parents?	66	16	3	1
	76.7%	18.6%	3.5%	1.2%
With your grandparents?	55	14	11	6
	64.0%	16.3%	12.8%	7.0%
With your brothers and sisters?	23	45	16	2
	26.7%	52.3%	18.6%	2.3%
With your friends away from school?	15	41	29	1
	17.4%	47.7%	33.7%	1.2%
With your friends at school?	20	38	25	3
	23.3%	44.2%	29.1%	3.5%
Total: 86				

Note. Only students whose native languages are not English are tabulated.

Table 46

Accuracy-Test and Speed-Test Samples: Responses From Non-Native Speakers of English to the Question "How Well Do You Use Your (Native) Language?" (Item 2e)

	Very well	Fairly well	Not well	Not at all	Missing
Accuracy-test sample					
Understand that language?	323	135	10	1	2
	68.6%	28.7%	2.1%	.2%	.4%
Speak that language?	276	168	22	a	5
	58.6%	35.7%	4.7%	a	1.1%
Read that language?	186	158	85	36	6
	39.5%	33.5%	18.0%	7.6%	1.3%
Write that language?	151	166	93	53	8
	32.1%	35.2%	19.7%	11.3%	1.7%
Total: 471					
Speed-test sample					
Understand that language?	51	27	3	1	4
	59.3%	31.4%	3.5%	1.2%	4.7%
Speak that language?	42	35	4	2	3
	48.8%	40.7%	4.7%	2.3%	3.5%
Read that language?	29	31	18	5	3
	33.7%	36.0%	20.9%	5.8%	3.5%
Write that language?	19	37	19	8	3
	22.1%	43.0%	22.1%	9.3%	3.5%
Total: 86					

Note. Only students whose native languages are not English are tabulated.

[a] For this sample, no students gave this response.

Table 47

Accuracy-Test and Speed-Test Samples: Responses to the Question "How Easy Has It Been for You in the Past to Understand Your Teacher's Explanations?" (Item 3)

	Very easy	Fairly easy	Fairly difficult	Very difficult	Missing
Accuracy-test sample					
Math	422	438	121	27	23
	40.9%	42.5%	11.7%	2.6%	2.2%
Science	478	405	97	17	34
	46.4%	39.3%	9.4%	1.6%	3.3%
Social studies/history	532	358	91	17	33
	51.6%	34.7%	8.8%	1.6%	3.2%
Total: 1031					
Speed-test sample					
Math	52	52	34	2	3
	36.4%	36.4%	23.8%	1.4%	2.1%
Science	51	68	18	3	3
	35.7%	47.6%	12.6%	2.1%	2.1%
Social studies/history	74	44	13	7	5
	51.7%	30.8%	9.1%	4.9%	3.5%
Total: 143					

Table 48

Accuracy-Test and Speed-Test Samples: Responses to the Question "How Easy Has It Been for You in the Past to Understand Your Textbooks?" (Item 4)

	Very easy	Fairly easy	Fairly difficult	Very difficult	Missing
Accuracy-test sample					
Math	443	411	122	24	31
	43.0%	39.9%	11.8%	2.3%	3.0%
Science	542	345	93	13	38
	52.6%	33.5%	9.0%	1.3%	3.7%
Social studies/history	518	360	93	21	39
	52.6%	33.5%	9.0%	1.3%	3.7%
Total: 1031					
Speed-test sample					
Math	57	52	32	a	2
	39.9%	36.4%	22.4%	a	1.4%
Science	67	54	15	5	2
	46.9%	37.8%	10.5%	3.5%	1.4%
Social studies/history	69	41	13	16	4
	48.3%	28.7%	9.1%	11.2%	2.8%
Total: 143					

Note. Only students whose native languages are not English are tabulated.

[a] For this sample, no students gave this response.

Table 49

Accuracy-Test and Speed-Test Samples: Responses to the Question "How Well Do You Use English?" (Item 6)

	Very well	Fairly well	Not well	Not at all	Missing
Accuracy-test sample					
Understand spoken English?	866	104	23	1	37
	84.0%	10.1%	2.2%	0.1%	3.6%
Speak English?	823	142	27	2	37
	79.8%	13.8%	2.6%	0.2%	3.6%
Read English?	796	173	25	0	37
	77.2%	16.8%	2.4%	0.0%	3.6%
Write English?	760	203	26	4	38
	73.7%	19.7%	2.5%	0.4%	3.6%
Total: 1031					
Speed-test sample					
Understand spoken English?	96	29	12	2	4
	67.1%	20.3%	8.4%	1.4%	2.8%
Speak English?	84	40	12	4	3
	58.7%	28.0%	8.4%	2.8%	2.1%
Read English?	79	45	16	0	3
	55.2%	31.5%	11.2%	0.0%	2.1%
Write English?	66	52	17	5	3
	46.2%	36.4%	11.9%	3.5%	2.1%
Total: 143					

Table 50

Model 2 Accuracy-Test Sample: Item 2di. How often do you speak that language with your parents?

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	2.7103	.5203	466
Gender			
Male	2.7182	.5170	220
Female	2.7033	.5243	246
Ethnicity			
Asian (American)	2.6271	.6104	118
African American	3.0000	.0000	3
Latino	2.7409	.4712	274
White	2.7358	.5244	53
Other	2.3750	.7440	8
Missing	2.9000	.3162	10
ESL code assigned by school			
Initially fluent (English)	2.6944	.5248	36
Beginning ESL	2.7143	.4880	7
Intermediate/advanced ESL	2.9412	.2425	17
(Awaiting) redesignation	2.7656	.4776	128
Other	2.7412	.4668	85
No code	2.6425	.5788	193
Type of math class			
Low	2.7882	.4110	85
Average	2.6983	.5177	179
High	2.6854	.5759	89
Algebra	2.6903	.5523	113
School lunch program			
Free	2.7461	.4924	193
Reduced payment	3.0000	.0000	3
Full payment	2.5000	.7559	8
Nonparticipant	2.6750	.4743	40
AFDC	2.8846	.3258	26
No lunch code	2.6633	.5628	196
Total valid cases: 466	Missing cases: 5		

Note. Only persons who are not native speakers of English are tabulated. Responses: 1 = never or hardly ever; 2 = sometimes; 3 = always or most of the time.

Table 51
Model 2 Accuracy-Test Sample: Item 2dii. How often do you speak that language with your grandparents?

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	2.7038	.6471	449
Gender			
Male	2.6588	.6881	211
Female	2.7437	.6071	238
Ethnicity			
Asian (American)	2.7179	.6137	117
African American	3.0000	.0000	3
Latino	2.6742	.6866	264
White	2.7755	.5502	49
Other	2.7500	.7071	8
Missing	2.8750	.3536	8
ESL code assigned by school			
Initially fluent (English)	2.8286	.4528	35
Beginning ESL	2.1667	.9832	6
Intermediate/advanced ESL	2.7500	.5774	16
(Awaiting) redesignation	2.7097	.6474	124
Other	2.5904	.7496	83
No code	2.7405	.6148	185
Type of math class			
Low	2.6024	.7315	83
Average	2.6647	.6875	170
High	2.6977	.6521	86
Algebra	2.8455	.4729	110
School lunch program			
Free	2.7081	.6521	185
Reduced payment	3.0000	.0000	3
Full payment	2.6250	.7440	8
Nonparticipant	2.6410	.6684	39
AFDC	2.8800	.4397	25
No lunch code	2.6878	.6628	189
Total valid cases: 449	Missing cases: 22		

Note. Only persons who are not native speakers of English are tabulated. Responses: 1 = never or hardly ever; 2 = sometimes; 3 = always or most of the time.

Table 52

Accuracy-Test Sample: Item 2diii. How often do you speak that language with your brothers and sisters?

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	1.9753	.6922	445
Gender			
Male	1.9571	.6869	210
Female	1.9915	.6979	235
Ethnicity			
Asian (American)	1.8673	.7134	113
African American	2.6667	.5774	3
Latino	2.0113	.6825	265
White	2.0000	.7223	47
Other	1.8750	.3536	8
Missing	2.0000	.7071	9
ESL code assigned by school			
Initially fluent (English)	1.8235	.6729	34
Beginning ESL	2.6667	.5164	6
Intermediate/advanced ESL	2.1333	.7432	15
(Awaiting) redesignation	1.9756	.6068	123
Other	2.1220	.7760	82
No code	1.9027	.6925	185
Type of math class			
Low	2.2169	.6993	83
Average	1.9337	.6618	166
High	1.9655	.6896	87
Algebra	1.8624	.7001	109
School lunch program			
Free	2.0161	.6696	186
Reduced payment	1.3333	.5774	3
Full payment	1.6250	.7440	8
Nonparticipant	1.7895	.6220	38
AFDC	2.1250	.6124	24
No lunch code	1.9785	.7275	186
Total valid cases: 445	Missing cases: 26		

Note. Only persons who are not native speakers of English are tabulated. Responses: 1 = never or hardly ever; 2 = sometimes; 3 = always or most of the time.

Table 53

Model 2 Accuracy-Test Sample: Item 2div. How often do you speak that language with your friends away from school?

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	1.7731	.6963	454
Gender			
Male	1.6901	.6853	213
Female	1.8465	.6992	241
Ethnicity			
Asian (American)	1.7179	.7525	117
African American	1.3333	.5774	3
Latino	1.7865	.6684	267
White	1.8824	.7388	51
Other	1.6250	.7440	8
Missing	1.7500	.4629	8
ESL code assigned by school			
Initially fluent (English)	1.4706	.7065	34
Beginning ESL	2.2857	.7559	7
Intermediate/advanced ESL	1.7647	.6642	17
(Awaiting) redesignation	1.7360	.6494	125
Other	1.9759	.7321	83
No code	1.7447	.6850	188
Type of math class			
Low	1.9881	.7027	84
Average	1.7647	.6905	170
High	1.7640	.7075	89
Algebra	1.6306	.6596	111
School lunch program			
Free	1.7566	.6716	189
Reduced payment	1.3333	.5774	3
Full payment	1.6250	.7440	8
Nonparticipant	1.8000	.6869	40
AFDC	1.7600	.7234	25
No lunch code	1.7989	.7233	189
Total valid cases: 454	Missing cases: 17		

Note. Only persons who are not native speakers of English are tabulated. Responses: 1 = never or hardly ever; 2 = sometimes; 3 = always or most of the time.

Table 54

Model 2 Accuracy-Test Sample: Item 2dv. How often do you speak that language with your friends at school?

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	1.6967	.7026	455
Gender			
Male	1.6526	.6598	213
Female	1.7355	.7375	242
Ethnicity			
Asian (American)	1.5726	.6863	117
African American	1.6667	.5774	3
Latino	1.8090	.7078	267
White	1.4706	.6435	51
Other	1.1250	.3536	8
Missing	1.7778	.6667	9
ESL code assigned by school			
Initially fluent (English)	1.4118	.6089	34
Beginning ESL	2.5000	.5477	6
Intermediate/advanced ESL	1.8824	.7812	17
(Awaiting) redesignation	1.7165	.6773	127
Other	1.9639	.7062	83
No code	1.5745	.6780	188
Type of math class			
Low	2.1084	.7159	83
Average	1.6590	.6857	173
High	1.7045	.6808	88
Algebra	1.4414	.5983	111
School lunch program			
Free	1.7579	.7305	190
Reduced payment	1.3333	.5774	3
Full payment	1.3750	.7440	8
Nonparticipant	1.5750	.6360	40
AFDC	1.7200	.6137	25
No lunch code	1.6772	.6969	189
Total valid cases: 445	Missing cases: 16		

Note. Only persons who are not native speakers of English are tabulated. Responses: 1 = never or hardly ever; 2 = sometimes; 3 = always or most of the time.

Table 55

Accuracy-Test Sample: Means and Standard Deviations of Responses From Non-Native Speakers of English to the Question "How Well Do You Understand Your (Native) Language?" (Item 2ei)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	3.6631	.5286	469
Gender			
Male	3.6171	.5400	222
Female	3.0745	.5157	247
Ethnicity			
Asian (American)	3.5630	.5769	119
African American	3.3333	.5774	3
Latino	3.7101	.5074	276
White	3.7358	.4451	53
Other	3.5000	.5345	8
Missing	3.4000	.6992	10
ESL code assigned by school			
Initially fluent (English)	3.4444	.5578	36
Beginning ESL	3.6250	.7440	8
Intermediate/advanced ESL	3.5556	.7838	18
(Awaiting) redesignation	3.7656	.4253	128
Other	3.7791	.4703	86
No code	3.5959	.5519	193
Type of math class			
Low	3.7045	.6095	88
Average	3.7079	.4802	178
High	3.6517	.5457	89
Algebra	3.5702	.5147	114
School lunch program			
Free	3.7077	.4887	195
Reduced payment	3.3333	.5774	3
Full payment	3.2500	1.0351	8
Nonparticipant	3.5250	.5986	40
AFDC	3.6154	.4961	26
No lunch code	3.6751	.5210	197
Total valid cases: 469	Missing cases: 2		

Note. Only students whose native languages are not English are tabulated. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

Table 56

Speed-Test Sample: Means and Standard Deviations of Responses from Non-Native Speakers of English to the Question "How Well Do You Understand Your (Native) Language?" (Item 2ei)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	3.5610	.6305	82
Gender			
Male	3.5333	.6252	45
Female	3.5946	.6438	37
Ethnicity			
Asian (American)	3.4000	.5477	5
African American	3.5000	.7071	2
Latino	3.5694	.6463	72
White	3.6667	.5774	3
Other	—	—	—
Missing	—	—	—
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	3.1429	.9493	14
Intermediate/advanced ESL	3.6923	.4804	13
(Awaiting) redesignation	3.4000	.5477	5
Other	3.5333	.6399	15
No code	3.7353	.4478	34
Type of math class			
Low	3.5758	.6577	66
Average	—	—	—
High	3.5556	.5270	9
Algebra	3.4286	.5345	7
School lunch program			
Free	3.6250	.5696	48
Reduced payment	4.0000	.0000	2
Full payment	—	—	—
Nonparticipant	3.5263	.6118	19
AFDC	3.0000	.0000	3
No lunch code	3.4000	.9661	10
Total valid cases: 82	Missing cases: 4		

Note. Only students whose native languages are not English are tabulated. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.
[a] For this subsample, there were no students in this category. [b] No data are missing.

Table 57

Accuracy-Test Sample: Means and Standard Deviations of Responses from Non-Native Speakers of English to the Question "How Well Do You Speak Your (Native) Language?" (Item 2eii)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	3.5451	.5858	466
Gender			
Male	3.5023	.5856	219
Female	3.5830	.5845	247
Ethnicity			
Asian (American)	3.4706	.6220	119
African American	3.3333	.5774	3
Latino	3.5730	.5838	274
White	3.6538	.4804	52
Other	3.2500	.4629	8
Missing	3.4000	.6992	10
ESL code assigned by school			
Initially fluent (English)	3.3333	.5345	36
Beginning ESL	3.4286	.9759	7
Intermediate/advanced ESL	3.6471	.6063	17
(Awaiting) redesignation	3.5984	.5234	127
Other	3.6897	.5769	87
No code	3.4792	.6050	192
Type of math class			
Low	3.5647	.6627	85
Average	3.5787	.5890	178
High	3.4831	.5663	89
Algebra	3.5263	.5356	114
School lunch program			
Free	3.5699	.5463	193
Reduced payment	3.3333	.5774	3
Full payment	3.1250	.8345	8
Nonparticipant	3.4000	.7089	40
AFDC	3.6538	.4852	26
No lunch code	3.5561	.5922	196
Total valid cases: 466	Missing cases: 5		

Note. Only students whose native languages are not English are tabulated. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

Table 58

Speed-Test Sample: Means and Standard Deviations of Responses from Non-Native Speakers of English to the Question "How Well Do You Speak Your (Native) Language?" (Item 2eii)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	3.4096	.6991	83
Gender			
Male	3.4783	.6579	46
Female	3.3243	.7474	37
Ethnicity			
Asian (American)	3.8000	.4472	5
African American	3.5000	.7071	2
Latino	3.4110	.7039	73
White	2.6667	.5774	3
Other[a]	—	—	—
Missing[b]	—	—	—
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	3.4000	.9856	15
Intermediate/advanced ESL	3.5385	.9674	13
(Awaiting) redesignation	3.2000	.4472	5
Other	3.2667	.4577	15
No code	3.4706	.5633	34
Type of math class			
Low	3.4030	.7190	67
Average[a]	—	—	—
High	3.4444	.7265	9
Algebra	3.4286	.5345	7
School lunch program			
Free	3.4286	.6455	49
Reduced payment	3.5000	.7071	2
Full payment[a]	—	—	—
Nonparticipant	3.2632	.8057	19
AFDC	3.0000	1.0000	3
No lunch code	3.7000	.6749	10
Total valid cases: 83	Missing cases: 3		

Note. Only students whose native languages are not English are tabulated. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

[a] For this subsample, there were no students in this category. [b] No data are missing.

Table 59

Accuracy-Test Sample: Means and Standard Deviations of Responses from Non-Native Speakers of English to the Question "How Well Do You Read Your (Native) Language?" (Item 2eiii)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	3.0624	.9437	465
Gender			
Male	2.9263	.9785	217
Female	3.1815	.8972	248
Ethnicity			
Asian (American)	2.9664	1.0079	119
African American	2.3333	.5774	3
Latino	3.2125	.8568	273
White	2.6538	1.0457	52
Other	2.2500	1.0351	8
Missing	3.1000	.8756	10
ESL code assigned by school			
Initially fluent (English)	2.6111	.9936	36
Beginning ESL	3.5714	.7868	7
Intermediate/advanced ESL	3.5294	.8745	17
(Awaiting) redesignation	3.1654	.8429	127
Other	3.3448	.8737	87
No code	2.8901	.9752	191
Type of math class			
Low	3.2235	.9683	85
Average	3.0618	.9337	178
High	3.0341	.8899	88
Algebra	2.9649	.9770	114
School lunch program			
Free	3.1192	.8787	193
Reduced payment	2.0000	1.0000	3
Full payment	3.1250	.9910	8
Nonparticipant	2.9500	.9594	40
AFDC	2.6923	.9703	26
No lunch code	3.0923	.9853	195
Total valid cases: 465	Missing cases: 6		

Note. Only subjects whose native languages are not English are tabulated. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

Table 60

Speed-Test Sample: Means and Standard Deviations of Responses from Non-Native Speakers of English to the Question "How Well Do You Read Your (Native) Language?" (Item 2eiii)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	3.0120	.9038	83
Gender			
Male	3.0217	.9543	46
Female	3.0000	.8498	37
Ethnicity			
Asian (American)	2.0000	1.2247	5
African American	3.0000	1.4142	2
Latino	3.1096	.8260	73
White	2.3333	1.1547	3
Other[a]	—	—	—
Missing[b]	—	—	—
ESL code assigned by school			
Initially fluent (English)	4.0000	.0000	1
Beginning ESL	3.2667	.8837	15
Intermediate/advanced ESL	3.0000	1.0801	13
(Awaiting) redesignation	2.6000	.8944	5
Other	2.8000	.9411	15
No code	3.0294	.8343	34
Type of math class			
Low	3.0000	.9045	67
Average[a]	—	—	—
High	2.8889	.9280	9
Algebra	3.2857	.9512	7
School lunch program			
Free	2.9796	.9012	49
Reduced payment	3.0000	.0000	2
Full payment[a]	—	—	—
Nonparticipant	2.9474	.9113	19
AFDC	2.6667	.5774	3
No lunch code	3.4000	1.0750	10
Total valid cases: 83	Missing cases: 3		

Note. Only students whose native languages are not English are tabulated. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

[a] For this subsample, there were no students in this category. [b] No data are missing.

Table 61

Accuracy-Test Sample: Means and Standard Deviations of Responses from Non-Native Speakers of English to the Question "How Well Do You Write Your (Native) Language?" (Item 2eiv)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	2.8963	.9880	463
Gender			
Male	2.7685	.9846	216
Female	3.0081	.9794	247
Ethnicity			
Asian (American)	2.7395	1.0207	119
African American	2.6667	.5774	3
Latino	3.0590	.9211	271
White	2.5000	1.0937	52
Other	2.2500	1.0351	8
Missing	3.0000	.9428	10
ESL code assigned by school			
Initially fluent (English)	2.3333	.9562	36
Beginning ESL	3.6667	.5164	6
Intermediate/advanced ESL	3.0000	1.1180	17
(Awaiting) redesignation	3.0000	.8997	127
Other	3.2588	.9018	85
No code	2.7396	1.0104	192
Type of math class			
Low	3.1205	.9803	83
Average	2.8927	1.0140	177
High	2.9551	.9282	89
Algebra	2.6930	.9697	114
School lunch program			
Free	2.9531	.9450	192
Reduced payment	2.0000	1.0000	3
Full payment	2.7500	.8864	8
Nonparticipant	2.6750	.9167	40
AFDC	2.5385	1.0288	26
No lunch code	2.9536	1.0296	194
Total valid cases: 463	Missing cases: 8		

Note. Only students whose native languages are not English are tabulated. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

Table 62

Speed-Test Sample: Means and Standard Deviations of Responses from Non-Native Speakers of English to the Question "How Well Do You Write Your (Native) Language?" (Item 2eiv)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	2.8072	.9034	83
Gender			
Male	2.7826	.9168	46
Female	2.8378	.8979	37
Ethnicity			
Asian (American)	1.4000	.5477	5
African American	3.0000	1.4142	2
Latino	2.9178	.8458	73
White	2.3333	.5774	3
Other[a]	—	—	—
Missing[b]	—	—	—
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	2.9333	1.1629	15
Intermediate/advanced ESL	2.8462	1.0682	13
(Awaiting) redesignation	2.6000	.5477	5
Other	2.4667	.9155	15
No code	2.9118	.7535	34
Type of math class			
Low	2.8060	.9085	67
Average[a]	—	—	—
High	2.5556	1.0138	9
Algebra	3.1429	.6901	7
School lunch program			
Free	2.8163	.7819	49
Reduced payment	3.0000	.0000	2
Full payment[a]	—	—	—
Nonparticipant	2.5263	1.0203	19
AFDC	2.6667	.5774	3
No lunch code	3.3000	1.2517	10
Total valid cases: 83	Missing cases: 3		

Note. Only students whose native languages are not English are tabulated. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.
[a] For this subsample, there were no students in this category. [b] No data are missing.

Table 63

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Math, How Easy Has It Been for You in the Past to Understand Your Teacher's Explanations?" (Item 3)

Background variables	Mean	*SD*	Cases
Full subsample	3.2450	.7657	1008
Gender			
Male	3.2430	.8088	465
Female	3.2468	.7276	543
Ethnicity			
Asian (American)	3.4125	.6949	160
African American	3.1027	.7411	185
Latino	3.2437	.8207	357
White	3.2077	.7373	260
Other	3.3333	.7581	30
Missing	3.6875	.4787	16
English status			
Native speaker	3.1837	.7585	528
Non-native speaker	3.3062	.7713	467
Missing	2.5385	.6602	13
ESL code assigned by school			
Initially fluent (English)	3.3265	.6888	49
Beginning ESL	3.2500	.8660	12
Intermediate/advanced ESL	3.1053	.8753	19
(Awaiting) redesignation	3.3931	.7388	145
Other	3.2340	.8351	94
No code	3.2134	.7602	689
Type of math class			
Low	3.1652	.9359	115
Average	3.2076	.7723	395
High	3.2624	.7651	221
Algebra	3.3177	.6704	277
School lunch program			
Free	3.2643	.7997	280
Reduced payment	3.0769	.8623	13
Full payment	3.1250	.8502	24
Nonparticipant	3.3000	.7735	120
AFDC	3.1831	.6614	71
No lunch code	3.2400	.7533	500
Total valid cases: 1008	Missing cases: 23		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.

Table 64

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Math, How Easy Has It Been for You in the Past to Understand Your Teacher's Explanations?" (Item 3i)

Background variables	Mean	*SD*	Cases
Full subsample			
Non-native speakers of English	3.1000	.8162	140
Gender			
Male	3.0152	.8681	66
Female	3.1757	.7650	74
Ethnicity			
Asian (American)	3.2857	.7559	7
African American	3.1481	.7698	27
Latino	3.0729	.8366	96
White	3.2222	.8333	9
Other[a]	—	—	—
Missing	2.0000	.0000	1
English status			
Native speaker	3.0909	.7721	44
Non-native speaker	3.0476	.8346	84
Missing	3.5000	.7977	12
ESL code assigned by school			
Initially fluent (English)	2.0000	.0000	1
Beginning ESL	3.1333	.9371	30
Intermediate/Advanced ESL	3.4000	.6325	15
(Awaiting) redesignation	3.6000	.8944	5
Other	3.0667	.7988	15
No code	3.0135	.7850	74
Type of math class			
Low	3.0737	.8283	95
Average[a]	—	—	—
High	3.2000	.7071	25
Algebra	3.1000	.9119	20
School lunch program			
Free	3.1045	.8190	67
Reduced payment	2.7500	.9574	4
Full payment[a]	—	—	—
Nonparticipant	3.0857	.7811	35
AFDC	3.5000	.5477	6
No lunch code	3.0714	.8997	28
Total valid cases: 140	Missing cases: 3		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 65

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Science, How Easy Has It Been for You in the Past to Understand Your Teacher's Explanations?" (Item 3ii)

Background variables	Mean	*SD*	Cases
Full subsample	3.3480	.7241	997
Gender			
Male	3.3457	.7440	460
Female	3.3501	.7073	537
Ethnicity			
Asian (American)	3.3375	.7922	160
African American	3.3989	.6949	183
Latino	3.2678	.7422	351
White	3.4109	.6846	258
Other	3.4828	.6336	29
Missing	3.3750	.6191	16
English status			
Native speaker	3.4119	.6793	522
Non-native speaker	3.2771	.7661	462
Missing	3.3077	.7511	13
ESL code assigned by school			
Initially fluent (English)	3.2653	.6701	49
Beginning ESL	2.4444	.7265	9
Intermediate/advanced ESL	2.8235	1.0146	17
(Awaiting) redesignation	3.2657	.6915	143
Other	3.2128	.8015	94
No code	3.4146	.6991	685
Type of math class			
Low	3.1651	.7995	109
Average	3.3265	.7011	392
High	3.2648	.7859	219
Algebra	3.5162	.6404	277
School lunch program			
Free	3.2445	.7526	274
Reduced payment	3.5000	.9045	12
Full payment	3.4167	.5836	24
Nonparticipant	3.3750	.6989	120
AFDC	3.1690	.7557	71
No lunch code	3.4173	.7030	496
Total valid cases: 997	Missing cases: 34		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.

Table 66

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Science, How Easy Has It Been for You in the Past to Understand Your Teacher's Explanations?" (Item 3ii)

Background variables	Mean	*SD*	Cases
Full subsample	3.1929	.7384	140
Gender			
Male	3.1364	.8017	66
Female	3.2432	.6787	74
Ethnicity			
Asian (American)	3.2857	.4880	7
African American	3.2963	.7240	27
Latino	3.1667	.7351	96
White	3.3333	.7071	9
Other[a]	—	—	—
Missing	1.0000	.0000	1
English status			
Native speaker	3.2500	.7813	44
Non-native speaker	3.1905	.7359	84
Missing	3.0000	.6030	12
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	3.0333	.5561	30
Intermediate/advanced ESL	2.9333	.7988	15
(Awaiting) redesignation	3.6000	.5477	5
Other	3.2000	.7746	15
No code	3.2838	.7855	74
Type of math class			
Low	3.1579	.7193	95
Average[a]	—	—	—
High	3.0000	.8660	25
Algebra	3.6000	.5026	20
School lunch program			
Free	3.0746	.7650	66
Reduced payment	3.5000	.5774	4
Full payment[a]	—	—	—
Nonparticipant	3.3429	.7648	35
AFDC	3.3333	.5164	6
No lunch code	3.2143	.6862	29
Total valid cases: 140	Missing cases: 3		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 67

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Social Studies/History, How Easy Has It Been for You in the Past to Understand Your Teacher's Explanations?" (Item 3iii)

Background variables	Mean	*SD*	Cases
Full subsample	3.4078	.7257	998
Gender			
Male	3.3696	.7715	460
Female	3.4405	.6831	538
Ethnicity			
Asian (American)	3.2688	.8066	160
African American	3.5191	.6617	183
Latino	3.3943	.7331	350
White	3.4672	.6719	259
Other	3.2667	.7397	30
Missing	3.1250	.9574	16
English status			
Native speaker	3.4579	.6781	522
Non-native speaker	3.3585	.7680	462
Missing	3.1538	.8987	13
ESL code assigned by school			
Initially fluent (English)	3.5306	.7101	49
Beginning ESL	2.3333	.7071	9
Intermediate/advanced ESL	3.0588	.8993	17
(Awaiting) redesignation	3.3986	.6934	143
Other	3.2872	.7846	94
No code	3.4402	.7077	686
Type of math class			
Low	3.2091	.8361	110
Average	3.4297	.7013	391
High	3.3000	.7711	220
Algebra	3.5415	.6449	277
School lunch program			
Free	3.4103	.7277	273
Reduced payment	3.2500	.9653	12
Full payment	3.5417	.6580	24
Nonparticipant	3.4417	.6456	120
AFDC	3.3803	.7244	71
No lunch code	3.3996	.7418	498
Total valid cases: 998	Missing cases: 33		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.

Table 68

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Social Studies/History, How Easy Has It Been for You in the Past to Understand Your Teacher's Explanations?" (Item 3iii)

Background variables	Mean	*SD*	Cases
Full subsample	3.3406	.8500	138
Gender			
Male	3.2344	.8682	64
Female	3.4324	.8289	74
Ethnicity			
Asian (American)	2.8571	1.0690	7
African American	3.7037	.4653	27
Latino	3.2128	.9025	94
White	3.8889	.3333	9
Other[a]	—	—	—
Missing	4.0000	.0000	1
English status			
Native speaker	3.4773	.7621	44
Non-native speaker	3.3253	.8569	83
Missing	2.9091	1.0445	11
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	2.6971	1.1001	28
Intermediate/advanced ESL	2.9333	.9612	15
(Awaiting) redesignation	3.8000	.4472	5
Other	3.6667	.6172	15
No code	3.6081	.5444	74
Type of math class			
Low	3.1720	.9397	93
Average[a]	—	—	—
High	3.6800	.4761	25
Algebra	3.7000	.4702	20
School lunch program			
Free	3.3182	.9308	66
Reduced payment	3.7500	.5000	4
Full payment[a]	—	—	—
Nonparticipant	3.4000	.6945	35
AFDC	4.0000	.0000	6
No lunch code	3.1111	.8916	27
Total valid cases: 138	Missing cases: 5		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 69

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Math, How Easy Has It Been for You in the Past to Understand Your Textbooks?" (Item 4i)

Background variables	Mean	*SD*	Cases
Full subsample	3.2730	.7662	1000
Gender			
Male	3.3030	.7704	462
Female	3.2472	.7624	538
Ethnicity			
Asian (American)	3.4534	.6418	161
African American	3.2043	.7719	186
Latino	3.2951	.7630	349
White	3.1418	.8128	261
Other	3.4643	.7445	28
Missing	3.6000	.7368	15
English status			
Native speaker	3.2042	.7977	529
Non-native speaker	3.3435	.7271	457
Missing	3.5714	.5136	14
ESL code assigned by school			
Initially fluent (English)	3.3265	.7184	49
Beginning ESL	3.1250	1.1260	8
Intermediate/advanced ESL	3.2222	.7321	18
(Awaiting) redesignation	3.3759	.7024	141
Other	3.2903	.7160	93
No code	3.2489	.7850	691
Type of math class			
Low	3.2973	.8485	111
Average	3.2672	.7905	393
High	3.3105	.7385	219
Algebra	3.2419	.7193	277
School lunch program			
Free	3.3199	.7469	272
Reduced payment	3.1538	.8006	13
Full payment	3.2917	.8065	24
Nonparticipant	3.2288	.7780	118
AFDC	3.1216	.8593	74
No lunch code	3.2826	.7567	499
Total valid cases: 1000	Missing cases: 31		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.

Table 70

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Math, How Easy Has It Been for You in the Past to Understand Your Textbooks?" (Item 4i)

Background variables	Mean	*SD*	Cases
Full subsample	3.1773	.7772	141
Gender			
Male	3.0746	.7846	67
Female	3.2703	.7639	74
Ethnicity			
Asian (American)	2.5714	.7868	7
African American	3.1852	.8338	27
Latino	3.2396	.7504	96
White	3.0000	.8660	9
Other[a]	—	—	—
Missing	3.0000	.0000	2
English status			
Native speaker	3.1136	.8413	44
Non-native speaker	3.1412	.7583	85
Missing	3.6667	.4924	12
ESL code assigned by school			
Initially fluent (English)	4.0000	.0000	1
Beginning ESL	3.1724	.7592	29
Intermediate/advanced ESL	3.3333	.7237	15
(Awaiting) redesignation	4.0000	.0000	5
Other	3.3333	.7237	15
No code	3.0526	.7982	76
Type of math class			
Low	3.1458	.7675	96
Average[a]	—	—	—
High	3.1200	.7257	25
Algebra	3.4000	.8826	20
School lunch program			
Free	3.2985	.7389	67
Reduced payment	2.7500	.9574	4
Full payment[a]	—	—	—
Nonparticipant	2.9429	.8023	35
AFDC	3.3333	.8165	6
No lunch code	3.2069	.7736	29
Total valid cases: 141	Missing cases: 2		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 71

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Science, How Easy Has It Been for You in the Past to Understand Your Textbooks?" (Item 4ii)

Background variables	Mean	*SD*	Cases
Full subsample	3.4260	.7148	993
Gender			
Male	3.4070	.7438	457
Female	3.4422	.6893	536
Ethnicity			
Asian (American)	3.3602	.7710	161
African American	3.4674	.7083	184
Latino	3.3681	.7117	345
White	3.5077	.6835	260
Other	3.5357	.6372	28
Missing	3.3333	.8165	15
English status			
Native speaker	3.5000	.6760	524
Non-native speaker	3.3407	.7424	455
Missing	3.4286	.9376	14
ESL code assigned by school			
Initially fluent (English)	3.4082	.5744	49
Beginning ESL	2.4286	.9759	7
Intermediate/advanced ESL	3.0625	.9287	16
(Awaiting) redesignation	3.4143	.6993	140
Other	3.1596	.8250	94
No code	3.4847	.6868	687
Type of math class			
Low	3.1121	.8504	107
Average	3.4092	.7135	391
High	3.3807	.7355	218
Algebra	3.6065	.5839	277
School lunch program			
Free	3.3717	.7093	269
Reduced payment	3.4615	.7763	13
Full payment	3.6667	.5647	24
Nonparticipant	3.4237	.7557	118
AFDC	3.2055	.7063	73
No lunch code	3.4758	.7074	496
Total valid cases: 993	Missing cases: 38		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.

Table 72

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Science, How Easy Has It Been for You in the Past to Understand Your Textbooks?" (Item 4ii)

Background variables	Mean	*SD*	Cases
Full subsample	3.2979	.7995	141
Gender			
Male	3.3088	.7776	68
Female	3.2877	.8246	73
Ethnicity			
Asian (American)	3.2857	.7559	7
African American	3.4815	.6427	27
Latino	3.2577	.8200	97
White	3.0000	1.0690	8
Other[a]	—	—	—
Missing	4.0000	.0000	2
English status			
Native speaker	3.3488	.7833	43
Non-native speaker	3.3140	.7556	86
Missing	3.0000	1.1282	12
ESL code assigned by school			
Initially fluent (English)	4.0000	.0000	1
Beginning ESL	2.9667	.9994	30
Intermediate/advanced ESL	3.0667	.7988	15
(Awaiting) redesignation	3.8000	.4472	5
Other	3.3333	.7237	15
No code	3.4267	.7008	75
Type of math class			
Low	3.2577	.8200	97
Average[a]	—	—	—
High	3.2083	.8330	24
Algebra	3.6000	.5982	20
School lunch program			
Free	3.2090	.8445	67
Reduced payment	3.7500	.5000	4
Full payment[a]	—	—	—
Nonparticipant	3.3429	.7648	35
AFDC	3.5000	.5477	6
No lunch code	3.3448	.8140	29
Total valid cases: 141	Missing cases: 2		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 73

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Social Studies/History, How Easy Has It Been for You in the Past to Understand Your Textbooks?" (Item 4iii)

Background variables	Mean	*SD*	Cases
Full subsample	3.3861	.7430	992
Gender			
Male	3.3736	.7516	455
Female	3.3966	.7362	537
Ethnicity			
Asian (American)	3.1988	.8202	161
African American	3.4570	.7431	186
Latino	3.3673	.7369	343
White	3.4865	.6550	159
Other	3.3571	.8262	29
Missing	3.2667	.9612	15
English status			
Native speaker	3.4685	.7214	525
Non-native speaker	3.2914	.7519	453
Missing	3.3571	.9288	14
ESL code assigned by school			
Initially fluent (English)	3.5714	.6455	49
Beginning ESL	2.3750	.7440	8
Intermediate/advanced ESL	3.0625	.7719	16
(Awaiting) redesignation	3.3714	.7131	140
Other	3.0968	.8652	92
No code	3.4344	.7191	687
Type of math class			
Low	3.1468	.8695	109
Average	3.3949	.7368	390
High	3.3287	.7885	216
Algebra	3.5126	.6290	277
School lunch program			
Free	3.3769	.7469	268
Reduced payment	3.3077	.7511	13
Full payment	3.5417	.5882	24
Nonparticipant	3.3866	.7144	119
AFDC	3.3699	.7361	73
No lunch code	3.3879	.7576	495
Total valid cases: 992	Missing cases: 39		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.

Table 74

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Social Studies/History, How Easy Has It Been for You in the Past to Understand Your Textbooks?" (Item 4iii)

Background variables	Mean	*SD*	Cases
Full subsample	3.1727	1.0139	139
Gender			
Male	3.0909	1.1194	66
Female	3.2466	.9095	73
Ethnicity			
Asian (American)	3.0000	1.1547	7
African American	3.6923	.5491	26
Latino	3.0104	1.0612	96
White	3.7500	.4629	8
Other[a]	—	—	—
Missing	2.5000	2.1213	2
English status			
Native speaker	3.4524	.8323	42
Non-native speaker	3.0588	1.0505	85
Missing	3.0000	1.2060	12
ESL code assigned by school			
Initially fluent (English)	4.0000	.0000	1
Beginning ESL	2.5517	1.2417	29
Intermediate/advanced ESL	2.8000	1.0142	15
(Awaiting) redesignation	3.8000	.4472	5
Other	3.6000	.6325	15
No code	3.3514	.8826	74
Type of math class			
Low	2.9375	1.0840	96
Average[a]	—	—	—
High	3.6667	.5647	24
Algebra	3.7368	.5620	19
School lunch program			
Free	3.0455	1.0732	66
Reduced payment	3.6667	.5774	3
Full payment[a]	—	—	—
Nonparticipant	3.2571	.8859	35
AFDC	4.0000	.0000	6
No lunch code	3.1379	1.0930	29
Total valid cases: 139	Missing cases: 4		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 75

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Math, How Easy Has It Been for You in the Past to Understand Questions on Tests?" (Item 5i)

Background variables	Mean	*SD*	Cases
Full subsample	3.2412	.7796	995
Gender			
Male	3.2729	.8033	458
Female	3.2142	.7585	537
Ethnicity			
Asian (American)	3.4136	.6654	162
African American	3.0939	.8544	181
Latino	3.2057	.8244	350
White	3.2385	.7118	260
Other	3.4286	.7902	28
Missing	3.7143	.4688	14
English status			
Native speaker	3.2023	.7853	524
Non-native speaker	3.2823	.7732	457
Missing	3.3571	.7449	14
ESL code assigned by school			
Initially fluent (English)	3.2041	.7354	49
Beginning ESL	3.0000	.9258	8
Intermediate/advanced ESL	3.2353	.7524	17
(Awaiting) redesignation	3.3147	.7998	143
Other	3.2447	.7576	94
No code	3.2310	.7818	684
Type of math class			
Low	3.1818	.9004	110
Average	3.1928	.8131	389
High	3.2329	.7696	219
Algebra	3.3394	.6759	277
School lunch program			
Free	3.2555	.7940	274
Reduced payment	3.3636	.6742	11
Full payment	3.3333	.7020	24
Nonparticipant	3.2137	.8390	117
AFDC	3.2329	.6774	73
No lunch code	3.2339	.7796	496
Total valid cases: 995	Missing cases: 36		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.

Table 76

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Math, How Easy Has It Been for You in the Past to Understand Questions on Tests?" (Item 5i)

Background variables	Mean	*SD*	Cases
Full subsample	2.9929	.8493	141
Gender			
Male	2.9552	.8245	67
Female	3.0270	.8754	74
Ethnicity			
Asian (American)	3.0000	.8165	7
African American	3.1481	.9074	27
Latino	2.9375	.8311	96
White	3.0000	1.0000	9
Other[a]	—	—	—
Missing	3.5000	.7071	2
English status			
Native speaker	3.0682	.9250	44
Non-native speaker	2.8941	.8021	85
Missing	3.4167	.7930	12
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	2.8966	.8170	29
Intermediate/advanced ESL	3.2000	.7746	15
(Awaiting) redesignation	3.4000	.8944	5
Other	3.0000	.9258	15
No code	2.9605	.8709	76
Type of math class			
Low	2.9375	.8311	96
Average[a]	—	—	—
High	3.1200	.8813	25
Algebra	3.1000	.9119	20
School lunch program			
Free	3.0597	.8327	67
Reduced payment	2.5000	1.2910	4
Full payment[a]	—	—	—
Nonparticipant	2.8000	.8677	35
AFDC	3.3333	.8165	6
No lunch code	3.0690	.7987	29
Total valid cases: 141	Missing cases: 2		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 77

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Science, How Easy Has It Been for You in the Past to Understand Questions on Tests?" (Item 5ii)

Background variables	Mean	*SD*	Cases
Full subsample	3.2722	.7810	992
Gender			
Male	3.2697	.8009	456
Female	3.2743	.7645	536
Ethnicity			
Asian (American)	3.2778	.8506	162
African American	3.2880	.7953	184
Latino	3.1983	.8030	343
White	3.3295	.6952	261
Other	3.4643	.6929	28
Missing	3.3571	.8419	14
English status			
Native speaker	3.3250	.7386	523
Non-native speaker	3.2132	.8187	455
Missing	3.2143	.9750	14
ESL code assigned by school			
Initially fluent (English)	2.0000	.8044	49
Beginning ESL	2.8235	.8165	7
Intermediate/advanced ESL	3.2357	.9510	17
(Awaiting) redesignation	3.1720	.7454	140
Other	3.2449	.8421	93
No code	3.3192	.7595	686
Type of math class			
Low	2.9813	.9005	107
Average	3.2442	.7458	389
High	3.2055	.8507	219
Algebra	3.4765	.6678	277
School lunch program			
Free	3.2082	.8066	269
Reduced payment	3.6364	.5045	11
Full payment	3.4583	.5882	24
Nonparticipant	3.2542	.7301	118
AFDC	3.0274	.7812	73
No lunch code	3.3300	.7827	497
Total valid cases: 992	Missing cases: 39		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.

Table 78

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Science, How Easy Has It Been for You in the Past to Understand Questions on Tests?" (Item 5ii)

Background variables	Mean	*SD*	Cases
Full subsample	3.0429	.7762	140
Gender			
Male	3.0455	.8121	66
Female	3.0405	.7483	74
Ethnicity			
Asian (American)	3.2857	.4880	7
African American	3.2222	.6980	27
Latino	3.0105	.7648	95
White	2.7778	1.2019	9
Other[a]	—	—	—
Missing	2.5000	.7071	2
English status			
Native speaker	3.0227	.7921	44
Non-native speaker	3.0545	.7660	84
Missing	3.0000	.8528	12
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	2.9310	.7036	29
Intermediate/advanced ESL	2.7333	1.0328	15
(Awaiting) redesignation	3.4000	.5477	5
Other	3.2667	.7037	15
No code	3.0800	.7669	75
Type of math class			
Low	3.0105	.8055	95
Average[a]	—	—	—
High	3.0000	.7638	25
Algebra	3.2500	.6387	20
School lunch program			
Free	3.0000	.8771	66
Reduced payment	3.2500	.5000	4
Full payment[a]	—	—	—
Nonparticipant	3.0571	.7648	35
AFDC	3.1667	.4082	6
No lunch code	3.0690	.6509	29
Total valid cases: 140	Missing cases: 3		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 79

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "In Social Studies/History, How Easy Has It Been for You in the Past to Understand Questions on Tests?" (Item 5iii)

Background variables	Mean	*SD*	Cases
Full subsample	3.2665	.7974	987
Gender			
Male	3.2583	.8103	453
Female	3.2734	.7870	534
Ethnicity			
Asian (American)	3.1852	.8577	162
African American	3.3204	.7938	181
Latino	3.2232	.8066	345
White	3.3346	.7321	257
Other	3.4286	.8357	28
Missing	3.0000	.8771	14
English status			
Native speaker	3.3243	.7586	518
Non-native speaker	3.2088	.8252	455
Missing[a]	—	—	—
ESL code assigned by school			
Initially fluent (English)	3.4898	.7107	49
Beginning ESL	2.0000	1.0000	7
Intermediate/advanced ESL	2.9375	.7719	16
(Awaiting) redesignation	3.1986	.8213	141
Other	3.1596	.7802	94
No code	3.3000	.7854	680
Type of math class			
Low	2.9725	.9473	109
Average	3.2526	.7858	384
High	3.1982	.8511	217
Algebra	3.4549	.6502	277
School lunch program			
Free	3.2379	.7793	269
Reduced payment	3.2500	.9653	12
Full payment	3.3333	.7614	24
Nonparticipant	3.2609	.8490	115
AFDC	3.2973	.7536	74
No lunch code	3.2759	.8020	493
Total valid cases: 987	Missing cases: 44		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] No data are missing.

Table 80

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "In Social Studies/History, How Easy Has It Been for You in the Past to Understand Questions on Tests?" (Item 5iii)

Background variables	Mean	*SD*	Cases
Full subsample	3.1071	.9345	140
Gender			
Male	3.0000	.9608	66
Female	3.2027	.9063	74
Ethnicity			
Asian (American)	3.1429	1.0690	7
African American	3.2963	.8234	27
Latino	3.0000	.9453	95
White	3.7778	.4410	9
Other[a]	—	—	—
Missing	2.5000	2.1213	2
English status			
Native speaker	3.2727	.8987	44
Non-native speaker	3.0833	.9595	84
Missing	2.6667	.7785	12
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	2.4483	1.0207	29
Intermediate/advanced ESL	2.7333	.9612	15
(Awaiting) redesignation	3.6000	.5477	5
Other	3.4667	.6399	15
No code	3.3333	.8275	75
Type of math class			
Low	2.9579	.9884	95
Average[a]	—	—	—
High	3.5200	.7703	25
Algebra	3.3000	.6569	20
School lunch program			
Free	3.0758	.9497	66
Reduced payment	3.0000	.8165	4
Full payment[a]	—	—	—
Nonparticipant	3.2000	.9641	35
AFDC	4.0000	.0000	6
No lunch code	2.8966	.9002	29
Total valid cases: 140	Missing cases: 3		

Note. Responses: 1 = Very difficult; 2 = Fairly difficult; 3 = Fairly easy; 4 = Very easy.
[a] For this subsample, there were no students in this category.

Table 81

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "How Well Do You Understand Spoken English?" (Item 6i)

Background variables	Mean	*SD*	Cases
Full subsample	3.8461	.4275	994
Gender			
Male	3.8410	.4368	459
Female	3.8505	.4196	535
Ethnicity			
Asian (American)	3.6852	.6153	162
African American	3.9558	.2061	181
Latino	3.7845	.4827	348
White	3.9423	.2496	260
Other	3.9655	.1857	29
Missing	3.7857	.5789	14
English status			
Native speaker	3.9579	.2102	523
Non-native speaker	3.7199	.5544	457
Missing	3.7857	.5789	14
ESL code assigned by school			
Initially fluent (English)	3.9592	.1999	49
Beginning ESL	2.4286	.5345	7
Intermediate/advanced ESL	3.1875	.5439	16
(Awaiting) redesignation	3.8380	.3884	142
Other	3.4468	.7422	94
No code	3.9242	.2961	686
Type of math class			
Low	3.5514	.6762	107
Average	3.8538	.4077	390
High	3.8914	.3778	221
Algebra	3.9130	.3070	276
School lunch program			
Free	3.8000	.4363	270
Reduced payment	3.9231	.2774	13
Full payment	3.9167	.4082	24
Nonparticipant	3.8974	.3568	117
AFDC	3.9452	.2292	73
No lunch code	3.8390	.4604	497
Total valid cases: 994	Missing cases: 37		

Note. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

Table 82

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "How Well Do You Understand Spoken English?" (Item 6i)

Background variables	Mean	*SD*	Cases
Full subsample	3.5755	.7120	139
Gender			
Male	3.5758	.6577	66
Female	3.5753	.7623	73
Ethnicity			
Asian (American)	3.4286	.7868	7
African American	3.9615	.1961	26
Latino	3.4421	.7816	95
White	3.8889	.3333	9
Other[a]	—	—	—
Missing	4.0000	.0000	2
English status			
Native speaker	3.7674	.6844	43
Non-native speaker	3.5952	.6423	84
Missing	2.7500	.7538	12
ESL code assigned by school			
Initially fluent (English)	4.0000	.0000	1
Beginning ESL	2.7241	.6490	29
Intermediate/advanced ESL	3.0000	.7559	15
(Awaiting) redesignation	4.0000	.0000	5
Other	3.9333	.2582	15
No code	3.9189	.3971	74
Type of math class			
Low	3.3895	.7895	95
Average[a]	—	—	—
High	3.9583	.2041	24
Algebra	4.0000	.0000	20
School lunch program			
Free	3.5385	.7088	65
Reduced payment	4.0000	.0000	4
Full payment[a]	—	—	—
Nonparticipant	3.7143	.7101	35
AFDC	3.8333	.4082	6
No lunch code	3.3793	.7752	29
Total valid cases: 139	Missing cases: 4		

Note. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.
[a] For this subsample, there were no students in this category.

Table 83

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "How Well Do You Speak English?" (Item 6ii)

Background variables	Mean	*SD*	Cases
Full subsample	3.7968	.4781	994
Gender			
Male	3.7817	.4954	458
Female	3.8097	.4628	536
Ethnicity			
Asian (American)	3.6111	.6617	162
African American	3.9171	.2959	181
Latino	3.7205	.5319	347
White	3.9157	.2918	261
Other	3.9310	.2579	29
Missing	3.7857	.5789	14
English status			
Native speaker	3.9293	.2916	523
Non-native speaker	2.6499	.5886	457
Missing	3.6429	.6333	14
ESL code assigned by school			
Initially fluent (English)	3.9184	.2766	49
Beginning ESL	2.1250	.6409	8
Intermediate/advanced ESL	3.2353	.7524	17
(Awaiting) redesignation	3.7518	.4497	141
Other	3.3298	.7675	94
No code	3.8949	.3254	685
Type of math class			
Low	3.4455	.7242	110
Average	3.8196	.4413	388
High	3.8500	.4168	220
Algebra	3.8623	.3850	176
School lunch program			
Free	3.7370	.5185	270
Reduced payment	3.9231	.2774	13
Full payment	3.9167	.2823	24
Nonparticipant	3.8448	.4492	116
AFDC	3.9315	.2543	73
No lunch code	3.7892	.4931	498
Total valid cases: 994	Missing cases: 37		

Note. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

Table 84

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "How Well Do You Speak English?" (Item 6ii)

Background variables	Mean	*SD*	Cases
Full subsample	3.4571	.7715	140
Gender			
Male	3.4776	.7253	67
Female	3.4384	.8163	73
Ethnicity			
Asian (American)	3.2857	.7559	7
African American	3.9615	.1961	26
Latino	3.2813	.8298	96
White	3.8889	.3333	9
Other[a]	—	—	—
Missing	4.0000	.0000	2
English status			
Native speaker	3.7442	.6933	43
Non-native speaker	3.4118	.6778	85
Missing	2.7500	1.1382	12
ESL code assigned by school			
Initially fluent (English)	4.0000	.0000	1
Beginning ESL	2.6667	.8023	30
Intermediate/advanced ESL	2.9333	.7988	15
(Awaiting) redesignation	3.8000	.4472	5
Other	3.7333	.4577	15
No code	3.7973	.4960	74
Type of math class			
Low	3.2604	.8366	96
Average[a]	—	—	—
High	3.8333	.3807	24
Algebra	3.9500	.2236	20
School lunch program			
Free	3.4091	.7641	66
Reduced payment	4.0000	.0000	4
Full payment[a]	—	—	—
Nonparticipant	3.6571	.6835	35
AFDC	3.8333	.4082	6
No lunch code	3.1724	.8892	29
Total valid cases: 140	Missing cases: 3		

Note. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.
[a] For this subsample, there were no students in this category.

Table 85

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "How Well Do You Read English?" (Item 6iii)

Background variables	Mean	*SD*	Cases
Full subsample	3.7757	.4739	994
Gender			
Male	3.7533	.4975	458
Female	3.7948	.4523	536
Ethnicity			
Asian (American)	3.6790	.5534	162
African American	3.9176	.2951	182
Latino	3.6484	.5564	347
White	3.8962	.3299	260
Other	3.9310	.2579	29
Missing	3.6429	.6333	14
English status			
Native speaker	3.8971	.3396	525
Non-native speaker	3.6462	.5554	455
Missing	3.4286	.6462	14
ESL code assigned by school			
Initially fluent (English)	3.8163	.4413	49
Beginning ESL	2.6250	.5175	8
Intermediate/advanced ESL	3.3750	.5000	16
(Awaiting) redesignation	3.7071	.4722	140
Other	3.2660	.6906	94
No code	3.8792	.3561	687
Type of math class			
Low	3.3364	.6535	110
Average	3.7933	.4363	387
High	3.8552	.4230	221
Algebra	3.8623	.3755	276
School lunch program			
Free	3.7222	.4960	270
Reduced payment	3.9231	.2774	13
Full payment	3.9583	.2041	24
Nonparticipant	3.8276	.3794	116
AFDC	3.8649	.3442	74
No lunch code	3.7666	.5059	497
Total valid cases: 994	Missing cases: 37		

Note. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

Table 86

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "How Well Do You Read English?" (Item 6iii)

Background variables	Mean	*SD*	Cases
Full subsample	3.4500	.6925	140
Gender			
Male	3.3731	.6927	67
Female	3.5205	.6894	73
Ethnicity			
Asian (American)	3.4286	.7868	7
African American	3.8462	.3679	26
Latino	3.3125	.7154	96
White	3.8889	.3333	9
Other[a]	—	—	—
Missing	3.0000	1.4142	2
English status			
Native speaker	3.6744	.6064	43
Non-native speaker	3.4000	.6761	85
Missing	3.0000	.8528	12
ESL code assigned by school			
Initially fluent (English)	4.0000	.0000	1
Beginning ESL	2.7667	.7279	30
Intermediate/advanced ESL	2.9333	.5936	15
(Awaiting) redesignation	3.8000	.4472	5
Other	3.6000	.5071	15
No code	3.7703	.4547	74
Type of math class			
Low	3.2917	.7387	96
Average[a]	—	—	—
High	3.7917	.4149	24
Algebra	3.8000	.4104	20
School lunch program			
Free	3.3485	.7124	66
Reduced payment	4.0000	.0000	4
Full payment[a]	—	—	—
Nonparticipant	3.6571	.5392	35
AFDC	4.0000	.0000	6
No lunch code	3.2414	.7863	29
Total valid cases: 140	Missing cases: 3		

Note. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

[a] For this subsample, there were no students in this category.

Table 87

Accuracy-Test Sample: Means and Standard Deviations of Responses to the Question: "How Well Do You Write English?" (Item 6iv)

Background variables	Mean	*SD*	Cases
Full subsample	3.7311	.5229	993
Gender			
Male	3.6908	.5575	456
Female	3.7654	.4895	537
Ethnicity			
Asian (American)	3.5802	.6755	162
African American	3.9000	.3189	180
Latino	3.6006	.5868	348
White	3.8692	.3490	260
Other	3.9310	.2579	29
Missing	3.5714	.6462	14
English status			
Native speaker	3.8757	.3738	523
Non-native speaker	3.5702	.6108	456
Missing	3.5714	.6462	14
ESL code assigned by school			
Initially fluent (English)	3.8980	.3058	49
Beginning ESL	2.5000	.5345	8
Intermediate/advanced ESL	3.0000	.6124	17
(Awaiting) redesignation	3.6429	.5099	140
Other	3.1702	.7425	94
No code	3.8467	.3990	685
Type of math class			
Low	3.3119	.7162	109
Average	3.7649	.4654	387
High	3.8145	.4833	221
Algebra	3.7826	.4630	276
School lunch program			
Free	3.6421	.5653	271
Reduced payment	3.9231	.2774	13
Full payment	3.8333	.4815	24
Nonparticipant	3.8190	.4293	116
AFDC	3.8082	.3964	73
No lunch code	3.7379	.5355	496
Total valid cases: 993	Missing cases: 38		

Note. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.

Table 88

Speed-Test Sample: Means and Standard Deviations of Responses to the Question: "How Well Do You Write English?" (Item 6iv)

Background variables	Mean	*SD*	Cases
Full subsample	3.2786	.7697	67
Gender			
Male	3.3433	.7697	67
Female	3.2192	.8539	73
Ethnicity			
Asian (American)	3.4286	1.1339	7
African American	3.8462	.3679	26
Latino	3.0625	.8054	96
White	3.8889	.3333	9
Other[a]	—	—	—
Missing	3.0000	1.4142	2
English status			
Native speaker	3.5349	.7668	43
Non-native speaker	3.2000	.7838	85
Missing	2.9167	.9962	12
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	2.4333	.8584	30
Intermediate/advanced ESL	2.8000	.7746	15
(Awaiting) redesignation	3.4000	.8944	5
Other	3.5333	.5164	15
No code	3.6622	.5041	74
Type of math class			
Low	3.0938	.8715	96
Average[a]	—	—	—
High	3.7083	.4643	24
Algebra	3.6500	.4894	20
School lunch program			
Free	3.1364	.8017	66
Reduced payment	4.0000	.0000	4
Full payment[a]	—	—	—
Nonparticipant	3.4571	.7005	35
AFDC	3.6667	.5164	6
No lunch code	3.2069	.9776	29
Total valid cases: 140	Missing cases: 3		

Note. Responses: 1 = not at all; 2 = not well; 3 = fairly well; 4 = very well.
[a] For this subsample, there were no students in this category.

Appendix L

Tables 89-106

Results of Analyses From Math Accuracy Tests

Table 89

Accuracy-Test Sample: Total Score by Ethnicity

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	15.4000	6.0025	525
Asian (American)	18.4878	4.8617	82
African American	12.9368	4.8527	95
Latino	12.3598	5.6151	189
White	19.0148	4.6311	135
Other	19.0667	5.4572	15
Missing	16.7778	7.4629	9
Booklet B			
Entire subsample	15.4664	5.8020	506
Asian (American)	18.4938	4.5638	81
African American	13.4845	5.4393	97
Latino	12.6818	5.1836	176
White	18.6923	4.8446	130
Other	19.6000	4.8961	15
Missing	9.1429	2.9114	7
All			
Entire sample	15.4326	5.9022	1031
Asian (American)	18.4908	4.7014	163
African American	13.2135	5.1512	192
Latino	12.5151	5.4063	365
White	18.8566	4.7308	265
Other	19.3333	5.1013	30
Missing	13.4375	6.9567	16
Total cases: 1031			

Note. Maximum possible score is 25.

Table 90

Accuracy-Test Sample: Total Score by ESL Code Assigned by School

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	15.4000	6.0025	525
Initially fluent (English)	16.6522	4.9691	23
Beginning ESL	5.8571	2.1157	7
Intermediate/advanced ESL	11.0000	5.6745	11
(Awaiting) redesignation	14.0294	4.8958	68
Other	10.0714	5.7780	42
No code	16.4786	5.7591	374
Booklet B			
Entire subsample	15.4664	5.8020	506
Initially fluent (English)	16.2692	5.0403	26
Beginning ESL	7.2000	3.1145	5
Intermediate/advanced ESL	8.1250	5.9387	8
(Awaiting) redesignation	13.6753	4.9403	77
Other	12.5926	5.7676	54
No code	16.5744	5.6161	336
All			
Entire sample	15.4326	5.9022	1031
Initially fluent (English)	16.4490	4.9584	49
Beginning ESL	6.4167	2.5391	12
Intermediate/advanced ESL	9.7895	5.8078	19
(Awaiting) redesignation	13.8414	4.9056	145
Other	11.4896	5.8777	96
No code	16.5239	5.6881	710
Total cases: 1031			

Note. Maximum possible score is 25.

Table 91

Accuracy-Test Sample: Total Score by Type of Math Class

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	15.4000	6.0025	525
ESL	7.5429	4.0099	35
Low	9.8889	4.4750	27
Average	12.9381	4.8726	210
High	17.5044	4.8239	113
Algebra	19.4615	3.6528	78
Honors algebra	21.6290	3.5540	62
Booklet B			
Entire subsample	15.4664	5.8020	506
ESL	8.1429	4.2018	35
Low	9.5769	4.9491	26
Average	13.3538	5.0902	195
High	17.5676	4.4469	111
Algebra	18.6962	3.4986	79
Honors algebra	21.0167	3.7257	60
All			
Entire subsample	15.4326	5.9022	1031
ESL	7.8429	4.0883	70
Low	9.7358	4.6705	53
Average	13.1383	4.9767	405
High	17.5357	4.6306	224
Algebra	19.0764	3.5851	157
Honors algebra	21.3279	3.6374	122
Total cases: 1031			

Note. Maximum possible score is 25.

Table 92

Accuracy-Test Sample: Total Score by School Lunch Program

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	15.4000	6.0025	525
Free lunch	14.3425	5.5850	146
Reduced pay	15.4000	4.3359	5
Full payment	19.7143	2.5246	14
Nonparticipant	14.4912	5.5617	57
AFDC	16.3889	5.9391	36
Missing	15.8127	6.3366	267
Booklet B			
Entire subsample	15.4664	5.8020	506
Free lunch	13.1825	5.0355	137
Reduced pay	14.1250	7.4726	8
Full payment	17.9000	4.8865	10
Nonparticipant	17.1429	4.8754	63
AFDC	16.2821	5.6613	39
Missing	16.1165	6.0767	249
All			
Entire subsample	15.4326	5.9022	1031
Free lunch	13.7809	5.3483	283
Reduced pay	14.6154	6.2655	13
Full payment	18.9583	3.7122	24
Nonparticipant	15.8833	5.3581	120
AFDC	16.3333	5.7571	75
Missing	15.9593	6.0284	516
Total cases: 1031			

Note. Maximum possible score is 25.

Table 93

Accuracy-Test Sample: Total Score by Native English Speaking Status

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	15.4000	6.0025	525
Native English speaking	16.2509	5.8526	283
Non-native English speaking	14.3713	6.0434	237
Missing	16.0000	6.0000	5
Booklet B			
Entire subsample	15.4664	5.8020	506
Native English speaking	16.4829	5.6561	263
Non-native English speaking	14.4615	5.7733	234
Missing	11.8889	5.3489	9
All			
Entire subsample	15.4326	5.9022	1031
Native English speaking	16.3626	5.7546	546
Non-native English speaking	14.4161	5.9046	471
Missing	13.3571	5.7326	14
Total cases: 1031			

Note. Maximum possible score is 25.

Table 94

Accuracy-Test Sample: Total Score by Gender

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	15.4000	6.0025	525
Male	15.9202	6.1963	238
Female	14.9686	5.8125	287
Booklet B			
Entire subsample	15.4664	5.8020	506
Male	15.4066	5.8517	241
Female	15.4208	5.7670	265
All			
Entire subsample	15.4326	5.9022	1031
Male	15.6618	6.0246	479
Female	15.2337	5.7920	552
Total cases: 1031			

Note. Maximum possible score is 25.

Table 95

Accuracy-Test Sample: Composite Scores on Items in Original Form by Ethnicity

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	5.7048	2.8045	525
Asian (American)	7.0976	2.1921	82
African American	4.6105	2.1699	95
Latino	4.2593	2.5768	189
White	7.3778	2.3527	135
Other	7.7333	2.5765	15
Missing	6.4444	3.4319	9
Booklet B			
Entire subsample	6.2292	2.5257	506
Asian (American)	7.5307	1.8513	81
African American	5.4227	2.5773	97
Latino	5.2159	2.3541	176
White	7.3462	2.2472	130
Other	7.7333	1.9809	15
Missing	3.8571	1.6762	7
All			
Entire subsample	5.9622	2.6828	1031
Asian (American)	7.3129	2.0352	163
African American	5.0208	2.4128	192
Latino	4.7205	2.5145	365
White	7.3623	2.2973	265
Other	7.7333	2.2581	30
Missing	5.3125	3.0270	16
Total cases: 1031			

Note. Maximum possible score is 10.

Table 96

Accuracy-Test Sample: Composite Scores on Items in Original Form by ESL Code Assigned by School

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	5.7048	2.8045	525
Initially fluent (English)	6.4348	2.5195	23
Beginning ESL	1.7143	.9512	7
Intermediate/advanced ESL	3.7273	2.4532	11
(Awaiting) redesignation	4.9265	2.2680	68
Other	3.4048	2.4798	42
No code	6.1925	2.7399	374
Booklet B			
Entire subsample	6.2292	2.5257	506
Initially fluent (English)	6.1923	2.2453	26
Beginning ESL	2.4000	1.8166	5
Intermediate/advanced ESL	3.5000	3.1168	8
(Awaiting) redesignation	5.7013	2.2247	77
Other	5.2407	2.5023	54
No code	6.6339	2.4701	336
All			
Entire subsample	5.9622	2.6828	1031
Initially fluent (English)	6.3061	2.3559	49
Beginning ESL	2.0000	1.3484	12
Intermediate/advanced ESL	3.3616	2.6710	19
(Awaiting) redesignation	5.3379	2.2706	145
Other	4.4375	2.6430	96
No code	6.4014	2.6232	710
Total cases: 1031			

Note. Maximum possible score is 10.

Table 97

Accuracy-Test Sample: Composite Scores on Items in Original Form by Type of Math Class

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	5.7048	2.8045	525
ESL	2.4857	1.7552	35
Low	3.2222	1.9081	27
Average	4.4952	2.2441	210
High	6.5310	2.4019	113
Algebra	7.6410	1.8303	78
Honors algebra	8.7581	1.7243	62
Booklet B			
Entire subsample	6.2292	2.5257	506
ESL	3.2571	2.1191	35
Low	3.7308	2.3926	26
Average	5.4154	2.3368	195
High	7.2523	1.9976	111
Algebra	7.3544	1.5855	79
Honors algebra	8.3167	1.6103	60
All			
Entire subsample	5.9622	2.6828	1031
ESL	2.8714	1.9702	70
Low	3.4717	2.1537	53
Average	4.9383	2.3322	405
High	6.8884	2.2353	224
Algebra	7.4968	1.7120	157
Honors algebra	8.5410	1.6770	122
Total Cases: 1031			

Note. Maximum possible score is 10.

Table 98

Accuracy-Test Sample: Composite Scores on Items in Original Form by School Lunch Program

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	5.7048	2.8045	525
Free lunch	5.1027	2.6671	146
Reduced pay	6.2000	1.6432	5
Full payment	8.0714	1.3848	14
Nonparticipant	5.3333	2.6682	57
AFDC	6.1944	1.7756	36
Missing	5.9139	2.8922	267
Booklet B			
Entire subsample	6.2292	2.5257	506
Free lunch	5.4453	2.3197	137
Reduced pay	5.5000	3.1623	8
Full payment	7.1000	1.5951	10
Nonparticipant	6.9841	2.2824	63
AFDC	6.2821	2.3614	39
Missing	6.4498	2.6317	249
All			
Entire subsample	5.9622	2.6828	1031
Free lunch	5.2686	2.5064	283
Reduced pay	5.7692	2.6190	13
Full payment	7.6667	1.5228	24
Nonparticipant	6.2000	2.5980	120
AFDC	6.2400	2.5513	75
Missing	6.1725	2.7798	516
Total cases: 1031			

Note. Maximum possible score is 10.

Table 99

Accuracy-Test Sample: Composite Scores on Items in Original Form by Native English Speaking Status

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	5.7048	2.8045	525
Native English speaking	6.1555	2.7322	283
Non-native English speaking	5.1603	2.8028	237
Missing	6.0000	2.9155	5
Booklet B			
Entire subsample	6.2292	2.5257	506
Native English speaking	6.5665	2.4670	263
Non-native English speaking	5.8889	2.5314	234
Missing	5.2222	2.9059	9
All			
Entire subsample	5.9622	2.6828	1031
Native English speaking	6.3535	2.6136	546
Non-native English speaking	5.5223	2.6934	471
Missing	5.5000	2.8216	14
Total cases: 1031			

Note. Maximum possible score is 10.

Table 100

Accuracy-Test Sample: Composite Scores on Items in Original Form by Gender

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	5.7048	2.8045	525
Male	5.9076	2.8640	238
Female	5.5366	2.7478	287
Booklet B			
Entire subsample	6.2292	2.5257	506
Male	6.0539	2.6175	241
Female	6.3887	2.4332	265
All			
Entire subsample	6.2292	2.5257	506
Male	6.0539	2.6175	241
Female	6.3887	2.4332	265
Total cases: 1031			

Note. Maximum possible score is 10.

Table 101

Accuracy-Test Sample: Composite Scores on Items in Revised Form by Ethnicity

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	6.3714	2.5317	525
Asian (American)	7.6341	2.1460	82
African American	5.3789	2.2700	95
Latino	5.2698	2.4875	189
White	7.6963	1.8739	135
Other	7.7333	2.2190	15
Missing	6.3333	3.1623	9
Booklet B			
Entire subsample	5.9051	2.6044	506
Asian (American)	7.1111	2.1966	81
African American	5.0412	2.2818	97
Latino	4.6080	2.2928	176
White	7.4846	2.1611	130
Other	7.8667	2.3258	15
Missing	3.0000	1.5275	7
All			
Entire subsample	6.1426	2.5770	1031
Asian (American)	7.3742	2.1804	163
African American	5.2083	2.2763	192
Latino	4.9507	2.4151	365
White	7.5925	2.0188	265
Other	7.8000	2.2345	30
Missing	4.8750	3.0304	16
Total cases: 1031			

Note. Maximum possible score is 10.

Table 102

Accuracy-Test Sample: Composite Scores on Items in Revised Form by ESL Code Assigned by School

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	6.3714	2.5317	525
Initially fluent (English)	6.5217	2.0861	23
Beginning ESL	2.8571	1.4639	7
Intermediate/advanced ESL	4.7273	2.7236	11
(Awaiting) redesignation	6.0147	2.3340	68
Other	4.1667	2.5654	42
No code	6.7888	2.3988	374
Booklet B			
Entire subsample	5.9051	2.6044	506
Initially fluent (English)	6.3462	2.5914	26
Beginning ESL	2.8000	1.3038	5
Intermediate/advanced ESL	3.0000	2.4495	8
(Awaiting) redesignation	5.1169	2.2940	77
Other	4.4815	2.3612	54
No code	6.3958	2.5359	336
All			
Entire subsample	6.1426	2.5770	1031
Initially fluent (English)	6.4286	2.3452	49
Beginning ESL	2.8333	1.3371	12
Intermediate/advanced ESL	4.0000	2.6874	19
(Awaiting) redesignation	5.5379	2.3482	145
Other	4.3438	2.4444	96
No code	6.6028	2.4707	710
Total cases: 1031			

Note. Maximum possible score is 10.

Table 103

Accuracy-Test Sample: Composite Scores on Items in Revised Form by Type of Math Class

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	6.3714	2.5317	525
ESL	3.1714	1.8389	35
Low	4.1852	2.4025	27
Average	5.5429	2.2964	210
High	7.2920	2.1451	113
Algebra	7.7564	1.6131	78
Honors algebra	8.5161	1.3640	62
Booklet B			
Entire subsample	5.9051	2.6044	506
ESL	2.8571	1.6114	35
Low	3.4615	2.0829	26
Average	5.0154	2.3209	195
High	6.5495	2.0791	111
Algebra	7.4430	1.8725	79
Honors algebra	8.4167	1.7878	60
All			
Entire subsample	6.1426	2.5770	1031
ESL	3.0143	1.7236	70
Low	3.8302	2.2595	53
Average	5.2889	2.3204	405
High	6.9241	2.1405	224
Algebra	7.5987	1.7499	157
Honors algebra	8.4672	1.5808	122
Total cases: 1031			

Note. Maximum possible score is 10.

Table 104

Accuracy-Test Sample: Composite Scores on Items in Revised Form by School Lunch Program

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	6.3714	2.5317	525
Free lunch	6.0205	2.3939	146
Reduced pay	6.0000	2.1213	5
Full payment	7.8571	1.1673	14
Nonparticipant	6.0185	2.3565	57
AFDC	6.6944	2.4357	36
Missing	6.5243	2.6772	267
Booklet B			
Entire subsample	5.9051	2.6044	506
Free lunch	4.7518	2.1753	137
Reduced pay	5.7500	2.9641	8
Full payment	7.4000	2.4129	10
Nonparticipant	6.6825	2.2420	63
AFDC	6.6667	2.7562	39
Missing	6.1687	2.6828	249
All			
Entire subsample	6.1426	2.5770	1031
Free lunch	5.4064	2.3732	283
Reduced pay	5.8462	2.5770	13
Full payment	7.6667	1.7611	24
Nonparticipant	6.3667	2.3116	120
AFDC	6.6800	2.5898	75
Missing	6.3527	2.6832	516
Total cases: 1031			

Note. Maximum possible score is 10.

Table 105

Accuracy-Test Sample: Composite Scores on Items in Revised Form by Native English Speaking Status

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	6.3714	2.5317	525
Native English speaking	6.6325	2.5094	283
Non-native English speaking	6.0633	2.5379	237
Missing	6.2000	2.2804	5
Booklet B			
Entire subsample	5.9051	2.6044	506
Native English speaking	6.4144	2.5321	263
Non-native English speaking	5.3889	2.5971	234
Missing	4.4444	1.7401	9
All			
Entire subsample	6.1426	2.5770	1031
Native English speaking	6.5275	2.5204	546
Non-native English speaking	5.7282	2.5869	471
Missing	5.0714	2.0555	14
Total cases: 1031			

Note. Maximum possible score is 10.

Table 106

Accuracy-Test Sample: Composite Scores on Items in Revised Form Gender

Background variables	Mean	*SD*	Cases
Booklet A			
Entire subsample	6.3714	2.5317	525
Male	6.5210	2.6715	238
Female	6.2474	2.4073	287
Booklet B			
Entire subsample	5.9051	2.6044	506
Male	5.9129	2.5958	241
Female	5.8981	2.6171	265
All			
Entire subsample	6.1426	2.5770	1031
Male	6.2150	2.6485	479
Female	6.0797	2.5140	552
Total cases: 1031			

Note. Maximum possible score is 10.

Appendix M

Tables 107-110

Results of Analyses From Speed Test

Table 107

Speed-Test Sample: Total Score Booklet A (All Original Items)

	Mean	*SD*	Cases
Full subsample	3.50000	2.6051	76
Gender			
Male	3.1714	1.8706	35
Female	3.7805	3.0944	41
Ethnicity			
Asian (American)	4.0000	3.3665	4
African-American	3.6875	3.0707	16
Latino	3.3542	2.5473	48
White	4.0000	2.0976	6
Other			
Missing	3.0000	1.4142	2
English status			
Native speaker	4.2500	2.8014	24
Non-native speaker	3.2391	2.5836	46
Missing	2.5000	1.0488	6
ESL code assigned by school			
Initially fluent (English)			
Beginning ESL	1.3333	1.0328	6
Intermediate/advanced ESL	2.1250	.8345	8
(Awaiting) redesignation	2.6667	2.0817	3
Other	3.5000	3.5857	8
No code	4.3571	2.7392	42
Type of math class			
Low	2.9600	2.2403	50
Average			
High	4.8571	2.8785	14
Algebra	4.1667	3.1575	12
School lunch program			
Free	3.0833	2.5453	36
Reduced payment	3.3333	1.5275	3
Full payment			
Nonparticipant	4.0526	2.8377	19
AFDC	5.5000	3.5355	2
No lunch code	3.5625	2.5812	16
Total cases: 76			

Note. Maximum possible score is 20.

Table 108

Speed-Test Sample: Total Score Booklet B (All Revised Items)

	Mean	*SD*	Cases
Full subsample	3.9851	2.4277	67
Gender			
Male	4.1818	2.6629	33
Female	3.7941	2.1989	34
Ethnicity			
Asian (American)	3.3333	.5774	3
African-American	4.5000	2.1950	12
Latino	3.5714	2.2174	49
White	9.3333	.5774	3
Other			
Missing			
English status			
Native speaker	4.1905	2.7680	21
Non-native speaker	4.0500	2.2753	40
Missing	2.8333	2.2286	6
ESL code assigned by school			
Initially fluent (English)	3.0000	.0000	1
Beginning ESL	1.3333	.5774	3
Intermediate/advanced ESL	2.7143	1.7043	7
(Awaiting) redesignation	4.0000	1.4142	2
Other	4.4286	2.0702	7
No code	5.0857	2.4659	35
Type of math class			
Low	3.6809	2.4769	47
Average			
High	5.1818	2.5620	11
Algebra	4.1111	1.6159	9
School lunch program			
Free	3.9063	2.6683	32
Reduced payment	4.0000	.0000	1
Full payment			
Nonparticipant	4.4375	2.4757	16
AFDC	5.7500	.5000	4
No lunch code	3.1429	1.9556	14
Total cases: 67			

Note. Maximum possible score is 20.

Table 109

Speed-Test Sample: Number of Items Attempted Booklet A (All Original Items)

	Mean	*SD*	Cases
Full subsample	9.2368	3.1363	76
Gender			
Male	8.5142	3.4246	35
Female	9.8537	2.7619	41
Ethnicity			
Asian (American)	10.2500	5.1235	4
African-American	9.3750	2.9411	16
Latino	8.9583	3.0524	48
White	11.3333	2.8752	6
Other			
Missing	6.5000	.7071	2
English status			
Native speaker	9.9583	3.1274	24
Non-native speaker	8.8478	3.0910	46
Missing	9.3333	3.5590	6
ESL code assigned by school			
Initially fluent (English)			
Beginning ESL	10.3333	3.9328	6
Intermediate/advanced ESL	6.8750	2.1002	8
(Awaiting) redesignation	8.6667	1.5275	3
Other	9.7500	2.8158	8
No code	9.5238	3.2176	42
Type of math class			
Low	9.0400	3.2637	50
Average			
High	10.2143	3.2148	14
Algebra	8.9167	2.4293	12
School lunch program			
Free	8.5278	2.9227	36
Reduced payment	8.6667	1.5275	3
Full payment			
Nonparticipant	9.8947	2.6852	19
AFDC	13.5000	4.9497	2
No lunch code	9.6250	3.7749	16
Total cases: 76			

Note. Maximum possible items is 20.

Table 110

Speed-Test Sample: Number of Items Attempted Booklet B (All Revised Items)

	Mean	*SD*	Cases
Full subsample	9.5373	2.8087	67
Gender			
Male	9.9697	3.2257	33
Female	9.1176	2.3063	34
Ethnicity			
Asian (American)	12.0000	6.0828	3
African-American	9.7500	3.0785	12
Latino	9.2449	2.5293	49
White	11.0000	1.7321	3
Other			
Missing			
English status			
Native speaker	9.0952	2.3644	21
Non-native speaker	9.8500	3.1095	40
Missing	9.0000	2.0976	6
ESL code assigned by school			
Initially fluent (English)	7.0000	.0000	1
Beginning ESL	7.3333	1.5275	3
Intermediate/advanced ESL	10.4286	4.3916	7
(Awaiting) redesignation	9.0000	.0000	2
Other	9.4286	2.8785	7
No code	10.1429	2.8195	35
Type of math class			
Low	9.5957	2.8942	47
Average			
High	9.6364	2.4606	11
Algebra	9.1111	3.0185	9
School lunch program			
Free	9.1875	2.4155	32
Reduced payment	12.0000	.0000	1
Full payment			
Nonparticipant	10.7500	3.6423	16
AFDC	11.5000	3.1091	4
No lunch code	8.2143	1.7177	14
Total cases: 67			

Note. Maximum possible items is 20.

Appendix N

Home Language Survey

____________________ DATE

____________________ SCHOOL

____________________ TEACHER

HOME LANGUAGE SURVEY

The California Education Code requires schools to determine the language(s) spoken at home by each student. This information is essential in order for schools to provide meaningful instruction for all students.

Name of student: ______________________________ ________ ________

Last Name First Name Middle Name Grade Age

1. Which language did your son or daughter learn when he or she first began to talk? ____________________
2. What language does your son or daughter most frequently use at home? ____________________
3. What language do you use most frequently to speak to your son or daughter? ____________________
4. Name the languages in the order most often spoken <u>by the adults at home</u>:
   a. ____________________
   b. ____________________
   c. ____________________

Signature of Parent or Guardian: ______________________________