

Statewide Mathematics Assessment in Texas

**A Group Including California State
University Math Professors & Statistician
Under the Organization of a Review Team
of Mathematically Correct Issued A
Devastating Analysis of the Integrity of
The Math Assessment Programs Including
End-of-Course Algebra in 1998**

THIS IS THE FULL REPORT

Statewide Mathematics Assessment in Texas

Paul Clopton, Wayne Bishop*, and David Klein
Mathematically Correct**

*** and Dept. of Mathematics and Computer Sciences, California State University, Los Angeles**

**** and Dept. of Mathematics, California State University, Northridge**

Permission to reproduce for non-commercial purposes without modification is granted.

Table of Contents

Abstract	ii
Introduction	1
Content Classification Schemes	2
Content Classification Schemes and the TAAS	3
Dimensions of the Essential Elements and Mathematics Objectives	4
Essential Elements Across Grade Levels	7
Grade Level Distributions of the TAAS Based on Essential Elements	9
Content Weakness Examples in the TAAS Exit Exam	11
Content “Slippage” due to TAAS Exit Exam Presentation Format	12
Examples of Low-Level Items in the TAAS Exit Exam	13
Sample Content Drawn from Japan	14
Judging the TAAS Exit Exam Items	15
Percent Correct by Item Grade Level Rating	17
Grade Level Ratings of TAAS Exit Exam Items by Objective	18
Technical Issues: Equating Tests for Difficulty	19
Technical Issues: Texas Learning Index Scores	19
Technical Issues: Curriculum Content Requirements	20
Minimum Expectations for “Passing” TAAS Exams	20
Distributions of Raw Scores on the TAAS Exit Exam	21
Algebra End-of-Course Exam Ratings	22
Content Weaknesses in the Algebra End-of-Course Exams	24
Potential Discontinuity in Texas Exam Objectives	26
Summary and Conclusions	27

Abstract

This report evaluates the mathematics assessments in use in Texas -- the Texas Assessment of Academic Skills (TAAS) mathematics exams and the end-of-course algebra 1 exams -- with respect to their mathematics content and grade level expectations as a way of understanding their impact on mathematics achievement in Texas. Several means of evaluating these assessment tools suggest that only low levels of achievement are being measured. Indeed, the high school exit exam seems more appropriate to sixth-grade achievement. Thus, the ability of the assessments to measure high levels of achievement is suspect. This suggests that Texas may not be effectively utilizing the power of a statewide assessment system to drive up achievement. Setting low expectations may also fail to keep students on track for later success in algebra and beyond. These effects seem to be a consequence of an assessment system that is too tightly focused on minimal achievement levels.

Introduction

The goal of the assessment program in Texas is to measure student progress toward achieving academic excellence. The primary purpose of the state student assessment program is to provide an accurate measure of student achievement . . .

Texas Student Assessment Program Technical Digest for the Academic Year 1996-1997, page 1

The statement of goal and purpose above does not make it explicit that the assessment program is intended to promote greater academic achievement statewide in Texas, or the desire to bring the achievement of Texas students up to “world-class standards” or similar lofty ambitions. Surely, promoting greater achievement is a central objective behind the implementation of the assessment system in Texas.

This report evaluates the mathematics assessments in use in Texas -- the Texas Assessment of Academic Skills (TAAS) mathematics exams and the end-of-course algebra 1 exams. Particular attention is paid to the mathematics content and grade level expectations of these exams as they relate to stimulating greater achievement in mathematics.

Statewide examinations in mathematics carry enormous potential as a tool to improve achievement at a relatively low cost. They can be useful to students to monitor their own progress and their relative strengths and weaknesses, and it is clear that the exams can provide a motivational influence. They are also useful in evaluating mathematics education. For example, they are helpful in identifying instructional methods and curriculum materials that are highly successful. Thus, these assessment devices are also motivational at the system level. Finally, they provide a degree of objective validity to the high school diploma and other indicators of academic progress – something that is needed after years of grade-inflation and social promotion have deteriorated the utility of these traditional indicators of academic success.

The introduction of assessment programs, particularly high-stakes assessments, can provide a strong source of motivation. This has become increasingly important as more and more teachers experience difficulty in motivating their students for academic success. The impact is widely evident in Texas where students spend extra time in test-preparation, textbooks address the assessments directly, and teachers strive to prepare their students for success on the examinations. Private industry even enters the picture by providing study materials and instruction specifically for the exams.

The potential benefits of a powerful statewide assessment system in mathematics are not guaranteed. Great care must be taken in designing these assessments, and the associated reporting mechanisms, so that the greatest possible benefits to mathematics achievement will result. With this in mind, this report evaluates the mathematics content and difficulty of these assessments. The goal is to clarify and interpret the nature of these assessments as a way of understanding their impacts on mathematics

achievement in Texas.

Content Classification Schemes

Various organizational schemes have been employed over the years to classify the content of mathematics education. Traditionally in the U.S. we have thought about mathematics education in terms of the subject areas of mathematics – things like arithmetic, algebra, geometry, and trigonometry. In more recent times, we have seen a variety of other sets of taxonomic categories proposed. These may be called strands, objectives, domains, or elements, but the underlying purpose is the same – to provide a structure for studying and thinking about the content of the mathematics curriculum.

In general, these systems provide a high-level structure of categories that runs across all grade levels. This has the obvious advantage of providing a unified vision of the curriculum across grades, and may help to clarify how particular content areas build over the years. Yet, the use of any such system must be approached with caution. It is clear, for example, that the meaning or interpretation of any one category changes considerably over the grade levels in terms of the kinds of material that it addresses. Also, none of the classification schemes is very successful in that particular mathematical ideas or procedures or problems are often difficult to assign to one and only one category.

In reality, all of these classification schemes are artificial and inexact. They should be thought of as nothing more than crude organizing methods that can serve as a helpful technique for the study of the curriculum. However, more often than not the system of classification takes on a life of its own and begins to influence both curriculum and assessment. There can be positive benefits that result. For example, discontinuities in the curriculum may be discovered and corrected. On the other hand, there can be negative consequences as well. One possibility is that there may be pressure to weight each content area equally across the grade levels. This seems certain to misrepresent what is important year to year, and gives far too much authority to a classification scheme that has arbitrary and inexact characteristics. For example, there is every reason to expect that instruction in the operations of arithmetic will be predominant over the use of algebraic manipulations in the early grades and that this will tend to reverse as students progress through the grade levels.

The dimensionality of the classification scheme may be used in attempts to change the face of the curriculum rather than to describe it. For example, as new categories are added, there will be a tendency to take emphasis away from “old” categories like arithmetic and algebra. While changes to the curriculum are not necessarily negative, each must be evaluated carefully on its own merit. There is no merit in curriculum changes introduced simply to satisfy the needs of the classification system itself.

Content Classification Schemes and the TAAS

The overall achievement goals in Texas have been divided into two, roughly parallel classification systems that are related to the development of the TAAS. These are illustrated below as a hierarchy of specifications that runs from overall achievement goals down to the actual test items.

Broad categories or strands are identified in the seven Essential Elements and in the thirteen TAAS Mathematics Objectives. The Objectives were generated subsequent to the Essential Elements, and are intended to encompass the material in the Essential Elements.

Grade level specifics are provided by the details of the Essential Elements and the Instructional Targets specified for the TAAS.

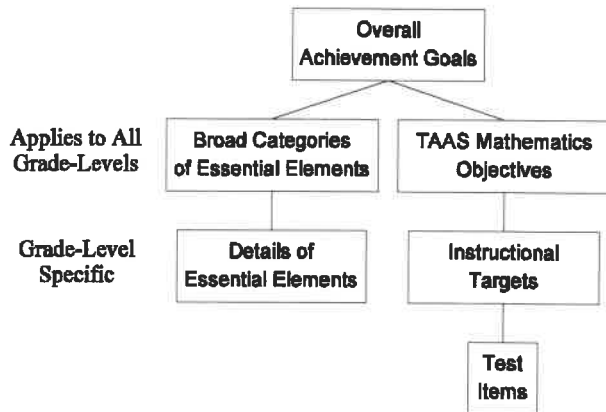
A cross-reference between the grade-level details of the Essential Elements and the Instructional Targets for the TAAS has been provided by the Texas Education Agency in the TAAS Mathematics Objectives and Measurement Specifications. This cross-reference provides a means to equate the two classification systems.

Test items are generated to match the TAAS Instructional Targets.

The number of test items in various content areas is based on weights for each Objective. Thus, relative emphases in the TAAS are determined by Objective.

The two classification schemes should be studied together to understand the content emphases.

Hierarchy of Achievement Specifications



Dimensions of the Essential Elements and Mathematics Objectives

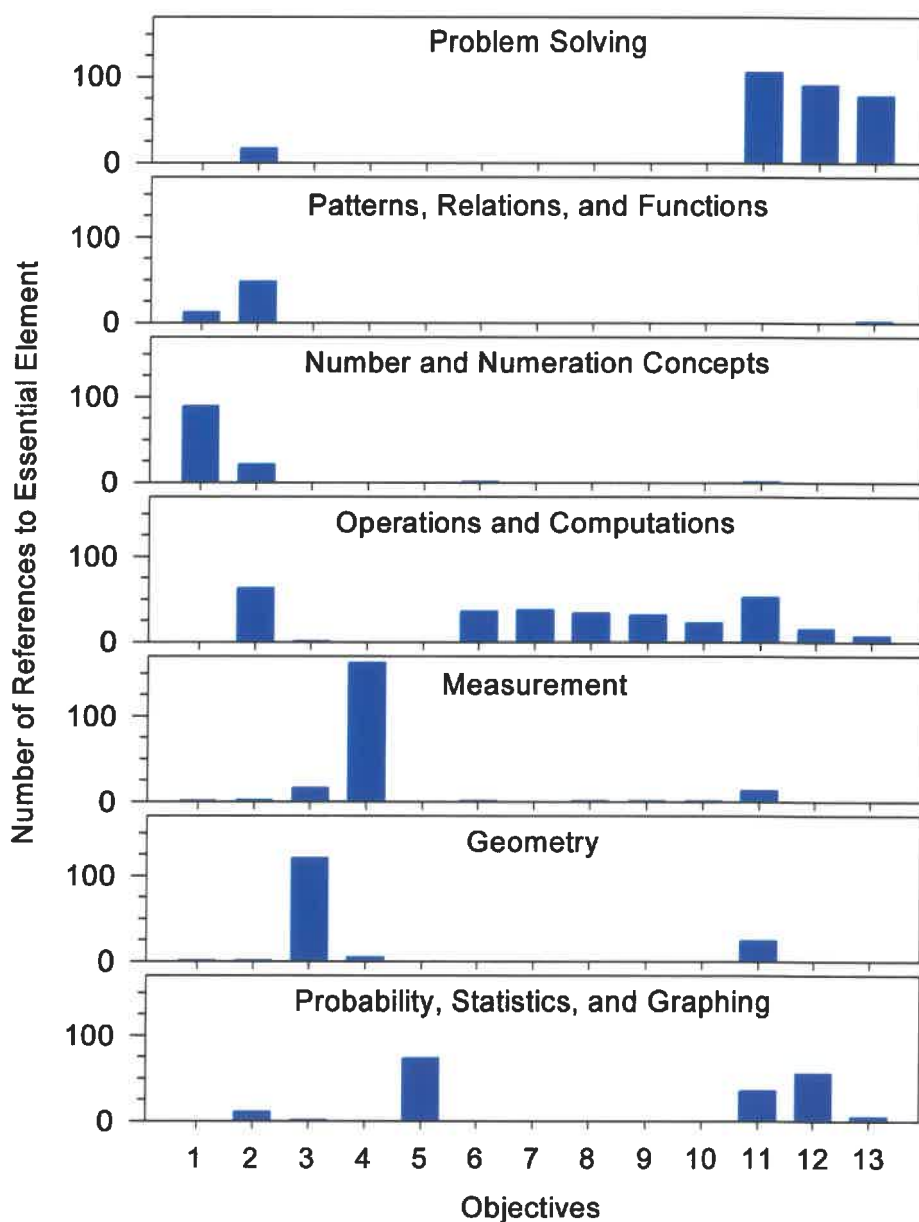
Both the Essential Elements and the Mathematics Objectives have strands that run across all grade levels. The dimensions of the two classification schemes can be roughly identified by the names in the table below:

Content Strands of the Essential Elements and TAAS Mathematics Objectives

Essential Elements	TAAS Mathematics Objectives
1 Problem solving	1 Number concepts
2 Patterns, relations, and functions	2 Relations, functions, and algebraic concepts
3 Number and numeration concepts	3 Geometric properties and relationships
4 Operations and computation	4 Measurement concepts
5 Measurement	5 Probability and statistics
6 Geometry	6 Addition
7 Probability, statistics, and graphing	7 Subtraction
	8 Multiplication
	9 Division
	10 Estimation
	11 Solution strategies
	12 Mathematical representation
	13 Evaluation of reasonableness

By counting the cross-references between the details of the Essential Elements and the Instructional Targets for each Mathematics Objective and summing these across grade levels, a rough map between the Essential Elements and the Mathematics Objectives can be generated. The results of this mapping are represented graphically below.

References to Essential Elements by Objective



By inspecting the figure up and down for each Mathematics Objective, the Essential Elements addressed by the Objective can be identified. There are no great surprises in this mapping.

By inspecting each row of the figure, the Mathematics Objectives that load on each Essential Element can be identified. An admittedly oversimplified summary is presented below.

Simplified Mapping of Mathematics Objectives onto Essential Elements

Essential Elements	TAAS Mathematics Objectives
1 Problem solving	11 Solution strategies 12 Mathematical representation 13 Evaluation of reasonableness
2 Patterns, relations, and functions	2 Relations, functions, and algebraic concepts
3 Number and numeration concepts	1 Number concepts
4 Operations and computation	6 Addition 7 Subtraction 8 Multiplication 9 Division 10 Estimation
5 Measurement	4 Measurement concepts
6 Geometry	3 Geometric properties and relationships
7 Probability, statistics, and graphing	5 Probability and statistics

There are a few obvious omissions from the table above. For example, Objective 2 - Relations, functions, and algebraic concepts - also taps Essential Element 4. Also, Essential Element 7 - Probability, statistics, and graphing - relates to the probability and statistics Objective, but shares some relationship to other Objectives.

Of primary importance in this mapping is the fact that the problem solving and the operations Essential Elements have been expanded to a more predominant position in the Objectives. This is important because the distribution of test items is set by the Objective, not by the Essential Element. This means that substantially more than 1/7th of the items will address Essential Element 4 - Operations and computation. This is altogether reasonable for the early grades where an emphasis on the operations of arithmetic is indicated. Whether it continues to be appropriate at higher grade levels is open to question and depends in part on the way the Objectives are elaborated. In particular, the transition to algebraic subject matter should appear in later grades.

There will also be an increased emphasis on Essential Element 1 - Problem solving. This is further augmented by the fact that it is related to Objectives 11 and 12 since these Objectives are given more items in the TAAS. This is reflective of the desire in Texas to provide more emphasis on “problem solving and complex thinking skills.” On the other hand, since these objectives are more difficult to define, they will also be related to the Essential Elements for operations and computation and probability, statistics, and graphing.

Essential Elements Across Grade Levels

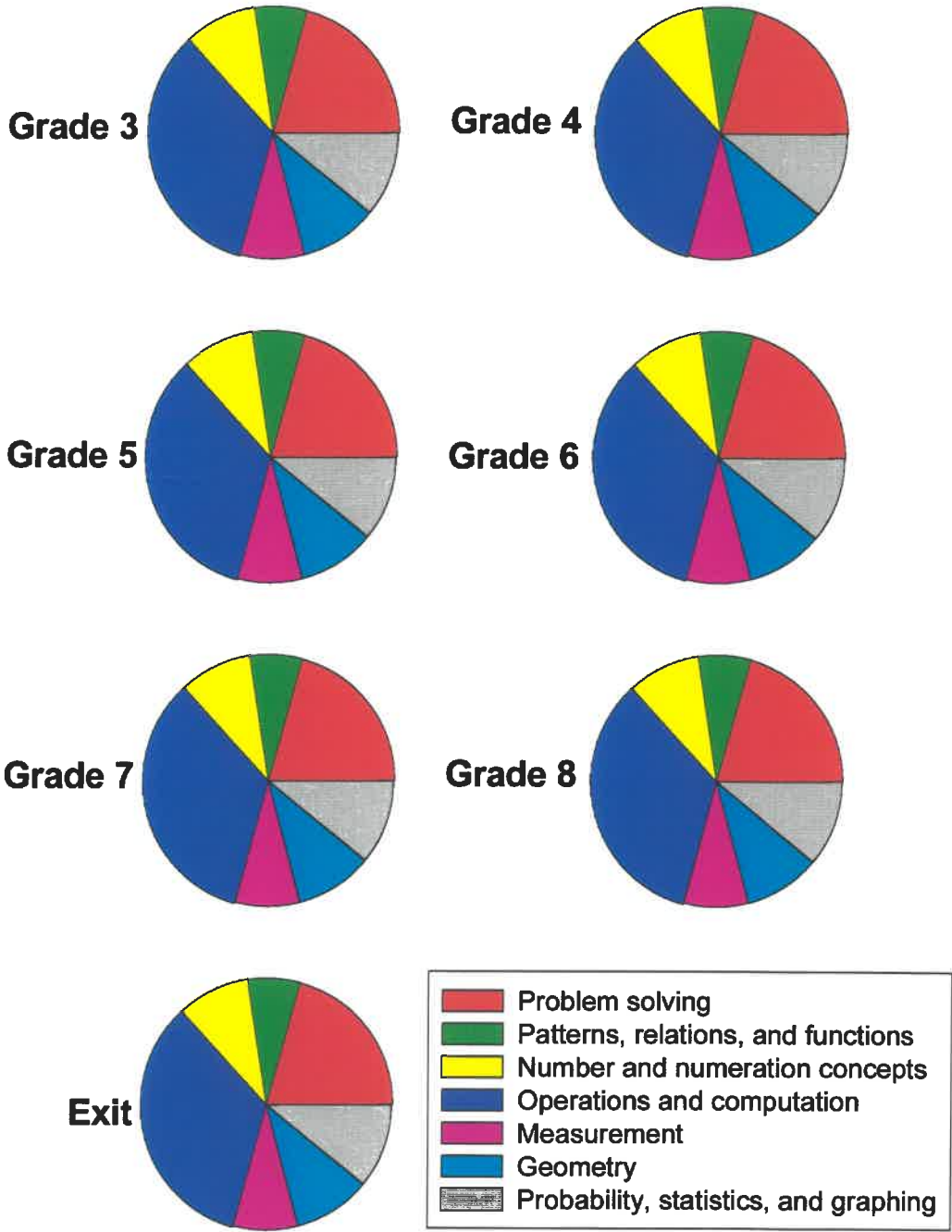
The mapping discussed above was collapsed across grade levels. Indeed, there is good reason to believe that relative emphases would vary across grade levels.

To address this issue, the references to Essential Elements were expressed as proportions within each Objective at each grade level. These were then weighted by the item counts for each Objective at each grade level so that weights for each Essential Element could be estimated. These were then expressed as a percentage of emphasis in the corresponding TAAS exam. The relative weights are illustrated in the figure below.

As anticipated, Essential Element 4 - Operations and computation - receives the largest emphasis at each grade level. Also as anticipated, Essential Element 1 - Problem solving - has the next largest weighting each year. The remaining Essential Elements have roughly equal weights. Moreover, this pattern is relatively consistent from grade 3 through the high school exit exam.

It is evident at this point that the system of specifications - Essential Elements, Objectives, Instructional Targets, and item counts per objective - has exerted an influence on the distribution of emphases in the TAAS. In part, this influence is entirely appropriate, such as the emphasis on operation and computation especially in the early test grades. On the other hand, the relatively consistent weights for the other Essential Elements are cause for concern. In particular, the lack of a demonstrated growth in the emphasis on algebraic content over the grades is worrisome. Similarly, one would expect a transition in the relative weights for measurement and geometry across the grades. The seriousness of these issues depends in part on the changing definitions of the strands across the grade levels in terms of the Instructional Targets and the actual test items that are generated. In any case, this information provides cause for concern that content emphases may be unduly influenced by the classification systems used to generate test specifications and test items.

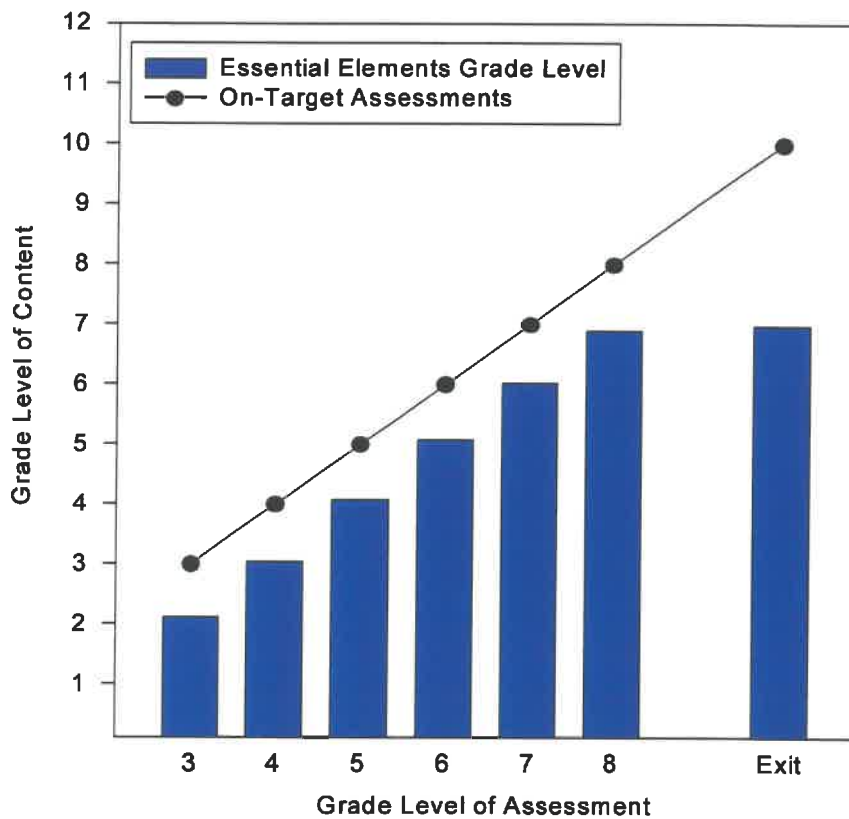
Distribution of Essential Element Weights by Grade Level



Grade Level Distributions on the TAAS Based on Essential Elements

Following the procedure of tracking references between Essential Elements and Instructional Targets, it is possible to relate item specifications for the TAAS to mean grade levels given in the Essential Elements. The results of this approach are illustrated below.

**TAAS Grade Level Estimates
Based on Essential Elements**



It is evident in the figure that mean item specifications in the TAAS lag a year behind the expected grade level in the Essential Elements, and that the expectations for the exit exam are nearly identical to the grade 8 exam. Both of these phenomena reflect processes intentional in the test design. Although legally mandated as a criterion-referenced assessment, item specifications generally address Essential Elements distributed across a range of grade levels - at grade level, one year behind, and two years behind. Thus, the average lag of about one grade level. In addition, special

consideration was made in designing the exit exam since high school students are not required to take higher level mathematics courses in order to graduate. Consequently, the exit exam appears in this analysis to be undifferentiated from the eighth grade exam.

Two points need to be clarified regarding the figure above. First, the minimum passing rates on these exams are approximately 70% correct. Thus, students will need to get at least some of the grade-level appropriate items correct to pass. This would have the effect of bringing the exam levels closer to the “on-target” line represented in the figure. However, these are also multiple-choice items with chance correct rates ranging from 20% to 25%. With chance taken into account, the grade-level estimates relative to the Essential Elements presented in the figure should be approximately correct.

Thus, relative to the Essential Elements, the TAAS exams appear to be behind grade level and the exit exam appears to be similar to the grade 8 exam.

However, these conclusions are based upon information provided by Texas Education Agency documents regarding item specifications and the Essential Elements. These conclusions do not speak to how the Essential Elements relate to improving achievement overall, and they do not address the relationship between the specifications and actual item content and difficulty levels. Rather than address each step in the hierarchy of achievement specifications that ultimately results in test items, content issues relating to the test items themselves are presented below.

Content Weakness Examples in the TAAS Exit Exam

An inspection of TAAS exit exam items reveals several achievement areas that are either missing or weakly represented in TAAS items. The following elements are provided to illustrate this finding, and are based upon the inspection of 240 TAAS items -- 60 per year for 4 years of exams.

Addition and subtraction of fractions with unlike denominators: Three addition items and three subtraction items were found. Their denominators are simple small integers in each case.

Multiplication and division of fractions or mixed numbers: There were no instances of the multiplication of two fractions. There was one instance of the division of a mixed number by a fraction.

Terminating and repeating decimals: There were no items related to this distinction.

Factors of numbers: There were no items found that directly addressed the factors of numbers, prime and composite numbers, greatest common factor or least common multiple.

Powers, roots, and exponents: There were two items found that called for the squares of integers (15 and 40). There was one item found that called for finding the two integers that bound the root of a number.

Properties of real numbers: There were two items found that dealt directly with the distributive property asking for the equivalence of two expressions.

Absolute value and negative numbers: There were no items found that dealt with absolute value. There was one item found that required sorting signed integers, one that asked about the distance between two altitudes one of which was below sea level, and one that required evaluating an expression containing a sum where one replacement value was negative.

Area and volume: There was one item found asking for the lateral surface area of a cylinder (although the formula is supplied). There was one item found asking for the volume of a rectangular prism.

Median and mode: There was one item found that asked for a median.

Solving equations: There were two items found that asked for the solutions of equations:
 $1.5x - 6 = 4.5$ and $c = \$15 + \$7.50p$ when c is $\$45$

The above content areas provide a flavor for elements of mathematics that are not well represented in the exit level TAAS.

Content “Slippage” due to TAAS Exit Exam Presentation Format

Multiple choice items are often difficult to write because students have the opportunity to work backwards from the available response choices. In other cases, item solutions can often be found by methods other than those intended to tap target objectives. The result is that what casual inspection may suggest to be the content addressed by an item is often different from the actual solution methods available to students. Three examples taken from the TAAS exit exam follow:

Students are asked for the ordered pair that represents the intersection of two lines given by linear equations. However, the lines are clearly graphed. This problem thus only requires being able to identify a point in the coordinate grid.

Students are told that two ladders are leaning against a building at the same angle. They are given the length of both ladders and the distance from the ladder base to the wall for the longer ladder (They are also informed in a figure that the ground forms a right-angle with the wall). Students evidently are to use reasoning about similar triangles and proportions to determine the distance from the base of the shorter ladder to the wall. Unfortunately, only one response choice is reasonable given the illustration that accompanies the problem. In fact, all incorrect response choices greatly exceed the entire length of the shorter ladder.

Three items appear to require the use of the Pythagorean theorem to solve for unknown lengths of right triangle sides, or at least the recognition and application of Pythagorean triples. However, the figures are drawn reasonably close to scale and only one response alternative for each item is reasonably possible given the figure.

Thus, some of the most difficult content areas addressed in the TAAS exit exam have simpler alternative solution strategies available that students are likely to employ.

Examples of Low-Level Items in the TAAS Exit Exam

Looking at actual TAAS exit exam items helps to clarify the nature of these tests. Some of the items discussed above are among the most advanced TAAS items in terms of the mathematics content. To illustrate the lower end of the spectrum, some of the items with the lowest level of mathematics content are listed below.

The total attendance recorded at the 1984 Summer Olympic Games in Los Angeles, California, was 5,797,923. What is this number rounded to the nearest thousand?

Mrs. Ramos has a plastic cube on her desk that holds photographs. There is a picture on every face of the cube except the bottom. How many pictures are displayed on the cube?

What is the approximate length of a new pencil before it is sharpened? (Response choices are 1.9 millimeters, 19 millimeters, 19 centimeters, and 1.9 meters)

Devon's house is on a rectangular block that is 330 yards long and 120 yards wide. What is the distance around his block?

Kenyon is 5 feet 6 inches tall. His sister Tenika is 7 inches taller than he is. How tall is Tenika?

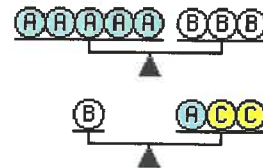
At a restaurant Steve ordered food totaling \$6.85. If he paid with a \$20 bill, how much change should he receive?

Certainly every high school graduate should be competent at solving problems of this nature. However, these items do not reflect the kinds of skills and knowledge that are grade level appropriate for high school students. There can be little question that these items are more appropriate to examinations used in much earlier grades.

Sample Content Drawn from Japan

In Japan, 12-year-olds are given a mathematics examination that consists of 225 story problems. These items show a depth of content for Japanese 12-year-olds that is striking in contrast to the TAAS exit examination items. Some of these items have been translated into English (Pacific Software Publishing, <http://www.japanese-online.com/math/index.htm>) and are reproduced here by permission.

How many 'C' balls does it take to balance one 'A' ball?



Jenny wanted to purchase 2 dozen pencils and a pen. Those items cost \$8.45 and she did not have enough money. So she decided to purchase 8 fewer pencils and paid \$6.05. How much was a pen?

Hose A takes 45 minutes to fill the bucket with water. Hose B can do the same in 30 minutes. If you use both hoses, how long will it take to fill the bucket?

A job takes 30 days to complete by 8 people. How long will the job take when it is done by 20 people?

Bob, Jim and Cathy each have some money. The sum of Bob's and Jim's money is \$18.00. The sum of Jim's and Cathy's money is \$21.00. The sum of Bob's and Cathy's money is \$23.00. How much money does each person have?

Tom's mother is 30 years old. The three children are 5, 3, and 0 years old. 12 years later, the total age of Tom's mother and father is twice as much as the total ages of all three children. How old is Tom's father?

Ellen baked cookies of the neighborhood children. She gave each child 6 cookies and she had 7 cookies remaining. So, she gave one more cookie to each child, but, was one cookie short. How many cookies did she bake in total?

It is 6 miles between Joe's house and Larry's house. Joe and Larry started to walk to each other's houses at noon, meeting at 12:30. Joe walked 2 miles per hour faster than Larry. How fast did Larry walk?

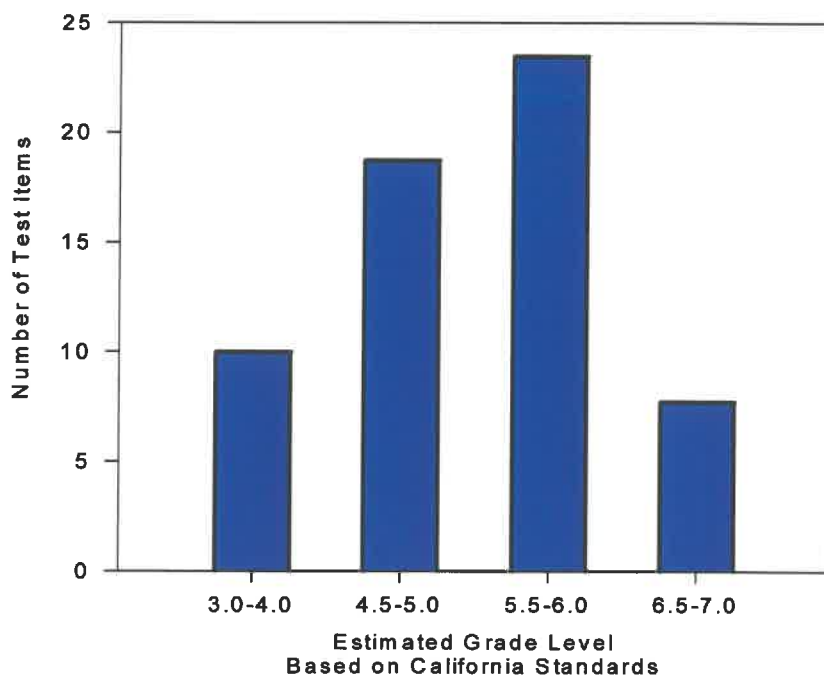
Judging the TAAS Exit Exam Items

To assess the target grade level of the TAAS exams against external criteria, individual exit exam items were evaluated as to grade level based on the newly established California Mathematics Standards. These standards provided a desirable benchmark for several reasons:

- They were designed carefully to be on-track with the best international competition, including Japan and Singapore.
- They are perhaps the most highly detailed of all sets of state mathematics standards, greatly facilitating item evaluation.
- They have been judged as the best available mathematics standards among all sets of state standards, even exceeding those from Japan (R. Raimi and L. Braden, State Mathematics Standards, Fordham Report, Vol 2 #3, <http://www.edexcellence.net/standards/math.html>).

Two of the authors (P.C. and D.K.) independently judged the grade level of every TAAS Mathematics exit exam item for the prior four years using the California standards as a guide. Although some degree of subjectivity was involved, these ratings proved to have a reasonably high level of rater reliability ($r=.813$). When the two ratings did not concur, they were averaged. The average distribution of item grade levels on a TAAS exit exam is illustrated below.

**Number of Items on TAAS Exit Exam
by Estimated Grade Level
Based on California Standards**



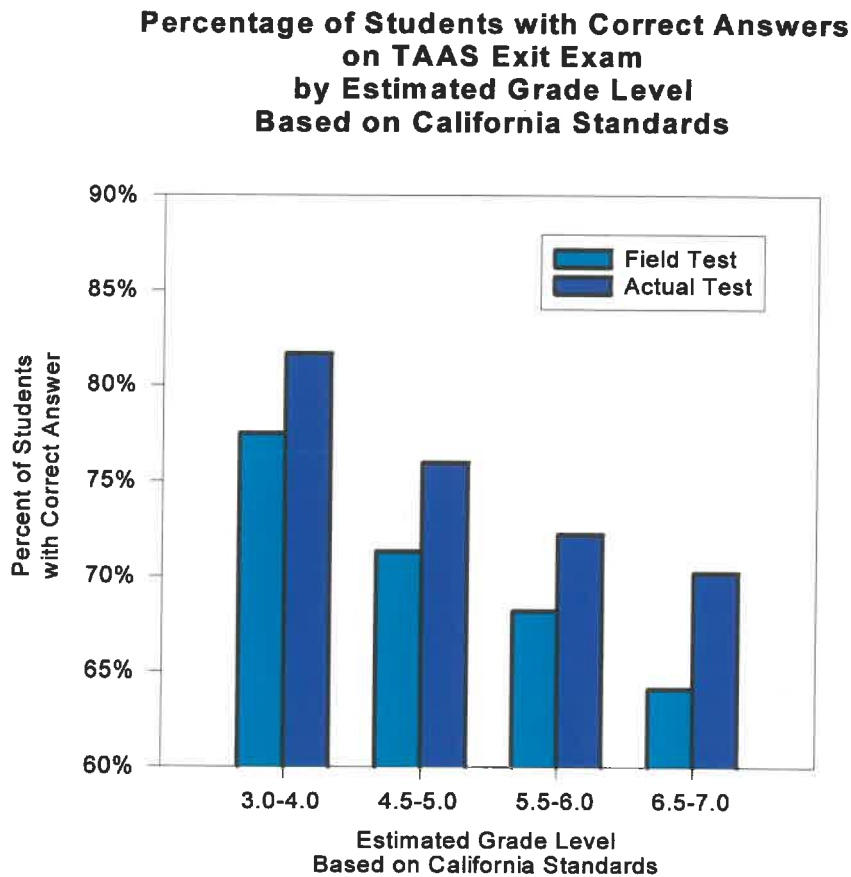
The ratings against the California Mathematics Standards yielded a mean grade level of 5.3 for the TAAS exit exam. The most advanced TAAS exit items were judged as equivalent to the California grade 7 standards.

Admittedly, the California Standards are set at a high level, being roughly equivalent to progress in Singapore and Japan. Nonetheless, the low estimated grade level is striking. Moreover, the California Standards are designed to complete the content of pre-algebra by grade 7 so that students will be ready to study algebra and geometry in grades 8 and above.

This finding raises the possibility that students could pass the TAAS exit level examination and still not be ready for the study of algebra. This possibility is consistent with the fact that Texas students enjoy greater success on the TAAS exams than on the algebra 1 end of course exam.

Percent Correct by Item Grade Level Rating

If the grade level ratings of TAAS exit exam items based on the California standards are a valid indicator, then TAAS exam scores would presumably be elevated by the inclusion of items below grade level. Likewise, the percentage of students with correct answers on individual items ought to vary as a function of item grade level ratings. Analysis of data for individual items from field tests and actual tests in Texas shows this is indeed the case as indicated below.

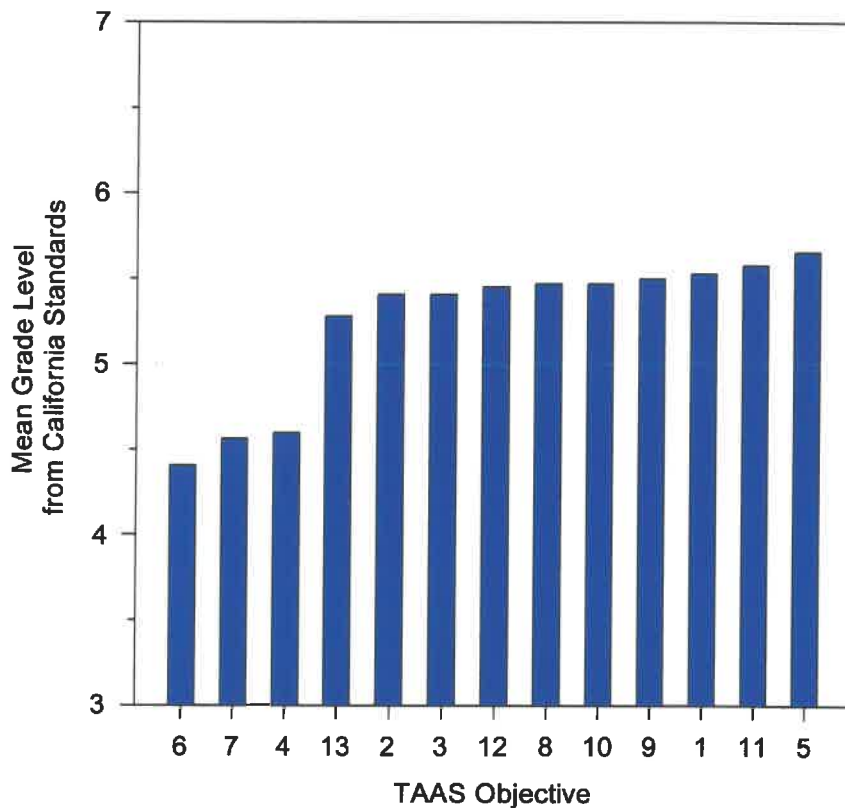


These results suggest that passing rates on the TAAS exit exam would be considerably lower if all items fell at the 7th grade (or higher) level based on the California standards.

Grade Level Ratings of TAAS Exit Exam Items by Objective

To study grade level expectations across content areas on the TAAS exit exam, the mean grade level rating was computed for each of the Mathematics Objectives indicated on the TAAS. This is possible because each TAAS item is tied to a Mathematics Objective. The grade level means are illustrated below.

**Mean Estimated Grade Levels
For TAAS Exit Exam
Based on California Standards
for Each TAAS Objective**



Mean grade level ratings were significantly lower for the Measurement, Addition, and Subtraction Objectives than for the other Objectives ($p < .05$ in each case). This may not be a surprising result as the topic areas of measurement, addition, and subtraction might be expected to enter the curriculum earlier than some of the other Objectives. Although more complex material could enter these content domains, this does not appear to be the case from the grade level ratings. Thus, it appears that the emphasis given by the weights in these three Objectives is grade level inappropriate in the exit exam. Likewise, Objectives 2 and 3, which should reflect algebraic and geometric material, do not appear to reflect more advanced content as might be expected.

Technical Issues: Equating Tests for Difficulty

The proportion of students meeting the minimum expectations on the TAAS exams has been rising. This is a positive indication, but the conclusion relies heavily on equating derived scores for test difficulty differences year to year. If the tests cannot be shown to be of equivalent difficulty (in derived scores), then the finding of improvement may be suspect.

TAAS exams are roughly equated for difficulty in the process of item selection. Furthermore, derived scores are further equated through statistical procedures that are mathematically sound. Since tests contain only a finite number of items, at least small fluctuations in the difficulty associated with a passing (minimum expectations) score are certain. Some further inaccuracy in the process of equating difficulty must exist. However, the degree of accuracy in the equating process is not documented in the Technical Digest. Furthermore, the tests are equated year to year, so that errors in this process may be compounding over time.

To address the equivalence of TAAS exit exams year to year, a comparison among the mean item grade-level ratings based on the California standards was computed for four years of exams. While this does not address the statistical equating process, it does provide an indication of relative item difficulties across years. The results of this test were not statistically significant ($p = .46$), with grade level means hovering around 5.3 for each of the test years. Thus, this method supported the notion that the tests were of roughly equal difficulty levels across years.

Technical Issues: Texas Learning Index Scores

Another design concern relates to the use of the Texas Learning Index (TLI). The design of the TLI is extremely useful in that it provides scores that students can use to gauge their relative achievement progress over the grades. Although a valuable resource, the TLI should be recognized as an essentially norm-based score for grades 3 to 8. This means that any unevenness in achievement across the grade levels in the norming year will be perpetuated in the TLI score distributions in succeeding years. For example, if the gain in mathematics achievement in the sixth grade was low as measured in the spring of 1994 due to some curriculum defects at the time, then a "passing" TLI score of 70 for grade 6 will continue to reflect this low level of achievement in future years. It is even possible that these future scores would be seen as endorsing achievement rates that would not be found if a criterion-referenced system was in place.

Technical Issues: Curriculum Content Requirements

During the screening of potential TAAS items, a review is conducted to insure that the item content is reasonably well addressed by the existing curriculum in Texas. Students are not to be tested on material they have not had the opportunity to study. However, this deviates from criterion-referenced assessment by adding the stipulation that criteria must be covered by the curriculum. The stipulation inhibits the ability of the assessments to drive advancements in the curriculum. There may be a risk of “curriculum stagnation” in which deficiencies in the curriculum tend to be perpetuated.

Minimum Expectations for “Passing” TAAS Exams

There are several reporting methods summarizing TAAS performance data. These include average scale scores, TLI scores, Texas Percentile Rank scores, Normal Curve Equivalents, and the percentage meeting minimum expectations. However, a great deal of emphasis is placed on the figures for the percentage of students meeting minimum expectations. The inherent risk in setting a “floor” or minimum requirement is that it will effectively become a “ceiling,” meaning that there may be a tendency to set a cap on achievement levels.

The basis of this thinking is that an emphasis on bringing up minimum achievement levels will tend to focus both the curriculum and instruction at levels that are too low to provide the greatest benefit to student achievement overall. Simply put, there is a risk that designating these targets will leave many students with lower achievement prospects than they are capable of.

This concern is amplified by the level of TAAS Instructional Targets relative to the Texas Essential Elements, the low achievement expectations evident in the items themselves and in comparison to the test for Japanese 12-year-olds, and the grade level ratings of TAAS items relative to the California Mathematics Standards. If the TAAS assessment levels are lower than optimal and there is a focus on minimum achievement relative to those assessments, there is an inherent risk that curriculum and instruction will be swayed toward sub-optimal levels as a result of the assessment process. Thus, it is possible that the TAAS exam system is not nearly as effective as it might be in promoting greater mathematics achievement statewide in Texas.

The presence of the TAAS examinations, and the minimum expectations used for graduation requirements, make this a high-stakes assessment system. As a consequence, a great deal of attention is given to preparation for TAAS exams. This includes material in school textbooks, classroom instruction time, and a sector of private industry that supplies materials and instruction. This illustrates the power of a high-stakes examination system. The findings given above suggest that Texas may not be making the greatest possible use of this power.

Distributions of Raw Scores on the TAAS Exit Exam

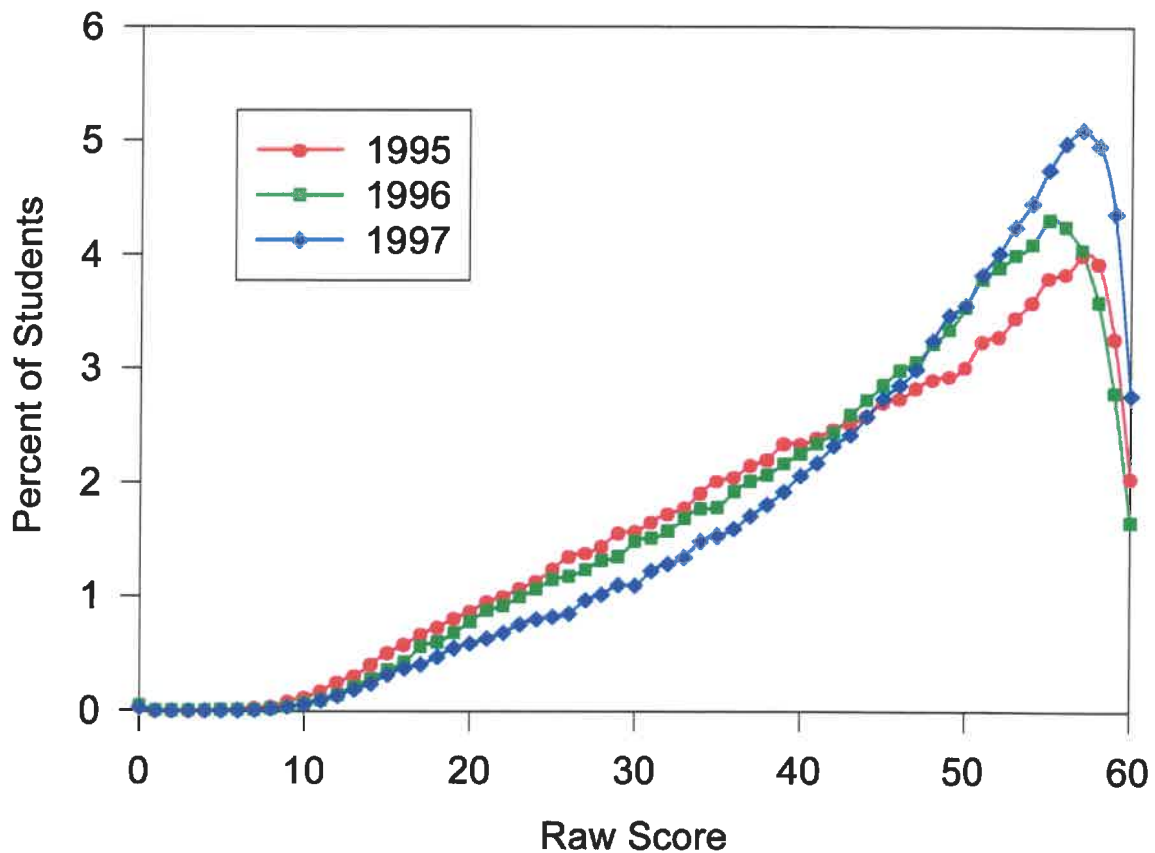
The distributions of raw scores on the TAAS exit exams are give below for three test years. These show strong negative skews. The presence of negative skew is not surprising given that the initial target of 70% correct is surpassed by a majority of students. However, the degree of skew is sufficient to suggest that the TAAS cannot function effectively in the identification of high achievement levels, and ceiling effects in the distribution are obvious. Since the exams do not differentiate well at higher achievement levels, we cannot tell whether or not the implementation of the assessment system is leading to similar ceiling effects in actual achievement. However, the lack of sensitivity to high achievement levels would suggest that the TAAS will not be effective at motivating achievement for a good proportion of students.

Algebra End-of-Course Exam Ratings

In addition to the TAAS exams, end-of-course examinations are also available in several content areas. In mathematics, students enrolled in introductory algebra take the algebra 1 end-of-course examination. They may use a passing score on this examination as part of their high school graduation requirement instead of the TAAS exit exam in mathematics.

In order to comment on both the content coverage and difficulty level of the Algebra 1 end-of-course exams, all of the exam items for four successive years were evaluated by one author (W.B.). To accomplish this, a list of common algebra concepts and skills was prepared by looking through several standard pre-algebra and algebra 1 textbooks and the examinations themselves. These were

Distributions of Raw Scores on TAAS Exit Exam



chosen so that an objective analysis of the content of each item could be completed.

Each item also received a 1-5 rating with the following identifiers:

- 1 Prior to Pre-Algebra
- 2 Pre-Algebra
- 3 Low Difficulty Algebra
- 4 Moderate Difficulty Algebra
- 5 High Difficulty Algebra

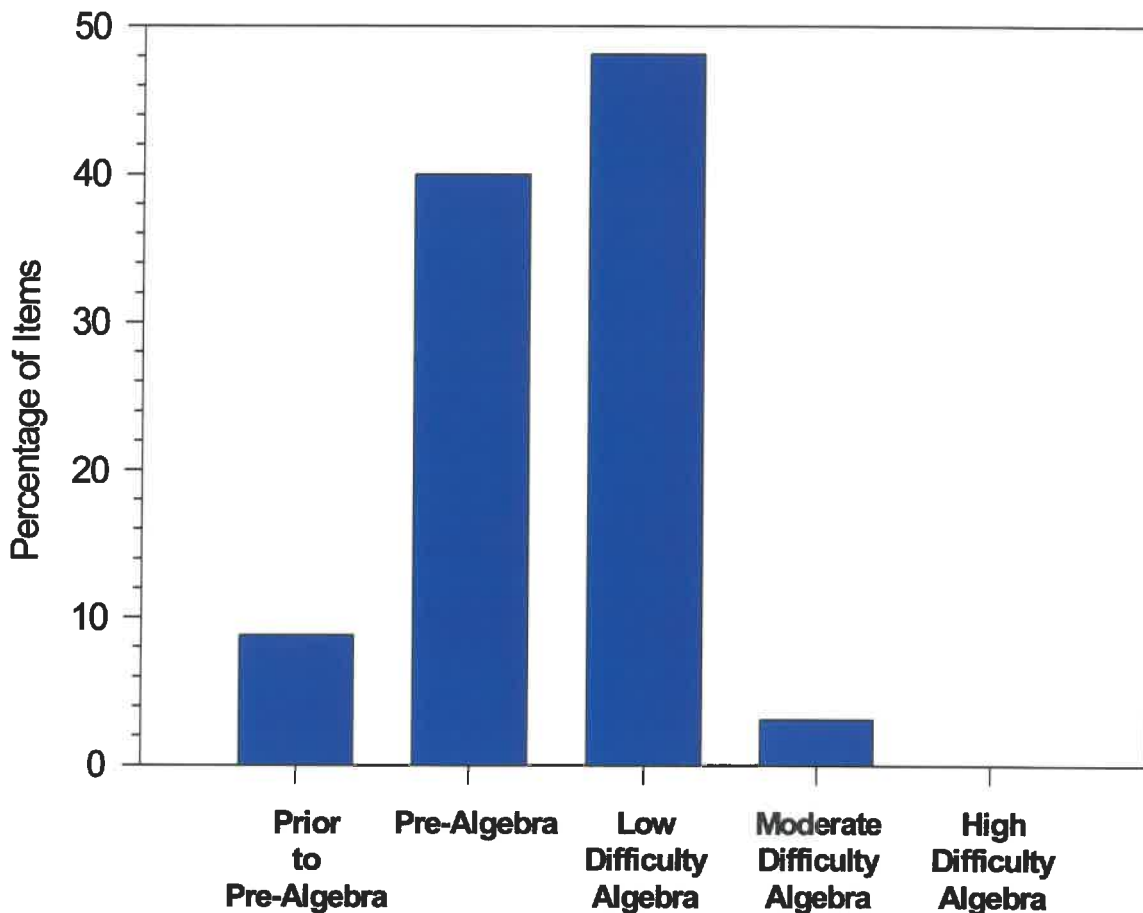
A rating of “3” represents the level of standard but easy algebra, the level of universal mastery of the content of algebra 1. A “2” represents standard pre-algebra, say at the level of Saxon Algebra 1/2 or Japanese Grade 7 Math (two of the sources used). A “1” is below that, roughly fourth or fifth grade math competence without even algebra readiness implied.

Going in the other direction, a “4” represents problems that require a more sophisticated level of algebra competence for solution. For example, clearing fractions and then solving a linear equation, or finding an obvious least common multiple of the denominators of two rational functions and adding them. A “5” level item is beyond that, though still appropriate for a broad screen, end-of-course algebra test. Examples would be something like clearing the fractions and then solving a resulting quadratic equation or solving an equation involving radicals that requires squaring both sides twice or that requires rejection of one or more of the “solutions” that were introduced in that process, or simplifying complex rational functions that involve quadratic (though easily factorable) expressions.

Generally speaking, if it appeared that an algebra-ready student who had not yet studied the subject should be able to solve the problem by inspection or by testing the given answer choices rather than using actual techniques of algebra, the item was rated at the pre-algebra level. That was especially true if that seemed to be the most likely approach to the problem, for example without clearing fractions, squaring both sides, or factoring as would be algebraically indicated. Some of the items would receive higher ratings if response grids rather than a multiple choice format were used.

The distribution of item ratings for the Algebra 1 end-of-course exams is indicated below.

Distribution of End of Course Algebra Exam Items



The mean rating across all items was 2.46 on the 5-point rating scale. As is evident in the figure, this means that the exams are primarily a combination of pre-algebra material and algebra at a low difficulty level. Mean ratings did not change significantly across test years ($p=.24$).

Approximately 14% of the items would have received higher ratings had alternative, less-advanced solution strategies not been available for the item in context (mostly if response grids could replace the multiple choice format). Thus, if these alternative solution strategies were not available, the mean rating would increase to 2.62 on the 5-point rating scale.

Content Weaknesses in the Algebra End-of-Course Exams

It is disturbing to see almost no factoring necessary on the tests. In fact, there is almost no algebraic simplification or algebraic arithmetic operation competency tested. Factoring appears to be intended in several items on each test, but it is always possible to get around it, sometimes trivially as with solution checking or with “factoring blocks” misused to give the actual factorization in this artificial form. Factoring simple polynomial expressions when possible is a very helpful tool in subsequent courses successive to this one so the idea needs to get more emphasis. The omission of factoring in this test sends a message to the state’s algebraic teaching community that symbolic manipulation skill does not matter.

Another area of concern is the need for better and more traditional word problems. The word problems that are presented mask a lack of algebraic depth. Direct and inverse variation (find the constant, for example, without a given model to follow) or, “How much water should be added to 20 quarts of a 30% solution to obtain a 14% solution,” have strong practical application in science. Less practical, perhaps, but excellent mathematical reading and algebra training are integer number problems such as, “The quotient of the successor of an integer number and one-third of the number is 4. What is the number?”

Finally, graphing linear equations in two variables is far too poorly done for the space that it consumes. Standard and useful items such as finding the slope-intercept form of a line that contains two given points, or a point and perpendicular to a given line, are almost nonexistent. Lots of pictures of graphs take up lots of pages but are not nearly as informative as a picture of one line with a couple of specified points for definiteness and then a response-grid question for the slope, the y-intercept, or the x-intercept. A good level “3” question here would be two lines given in standard form and graphed that appear to intersect at an integer pair, say (3,2). The question would then be something like, “These lines intersect at a point close to (3,2) but we cannot tell from the graph if that is precisely correct. Which of these is the correct value for the x-coordinate of the point of intersection?” A level “5” question along this line would be to not give equations of the lines but to give the coordinates of two points on each line so the first step is to find equations of the lines and the second is to use them to find the actual point of intersection. Questions like this can be multiple choice and still be testing the underlying algebra concepts. If they are not a bit more complicated, such as these suggestions are, the ability to just check the answers defeats confirmation of the intended algebra concepts.

The following points with respect to content should be noted.

1. There are too many formulas given. Most, such as the area of a rectangle or the slope of a line, should be assumed to be known.
2. There are no rational function reduction or arithmetic items.

3. There is almost no confirmation of algebraic arithmetic skills at all: products of binomials, reduction of fractions involving exponential monomials, etc.
4. There are no problems that require factoring polynomials, even “taking out” a common factor.
5. Standard word problems that lead to algebraic solution are inadequate in number and in depth.
6. Although there are enough items, linear graphing is inadequate; e.g., there are no items that involve the slopes of perpendicular lines.
7. Many items can be done by checking given answer choices or by inspection. Even just a “none of the above” would be helpful.
8. The distance formula is never used and all Pythagorean Theorem items can avoid it.
9. Radical equation exercises are trivial or don’t exist.
10. Scattergram questions are trivial and should not be included.
11. “Which BEST describes . . . ” language is used even when a perfect fit is among the choices.
12. The response-grid format is insufficiently used. It is down to one item in 1998.

In summary, this test is more of an algebra readiness test than it should be. There are no items that require more than the most trivial symbolic manipulation, a standard part of algebra, and many of the items only appear to require an algebraic solution. They can be done, and will be done, by inspection or by testing the given answer choices. The statistics and probability questions are not at algebra level under the most generous interpretation.

Potential Discontinuity in Texas Exam Objectives

By combining together much of the information given above, a preliminary look at the progression through algebra 1 in the Texas assessments is possible. This method is admittedly tentative, but may be useful for understanding the progression to algebra in Texas.

To compile this data:

California grade-level estimates for TAAS grades 3 to 8 were assumed to be discrepant from the Essential Elements reference mapping in a manner that was proportional to the difference seen in the exit exam.

Algebra exam items rated at 3 and above were assumed to reference California grade 8. Items rated at 2 were assumed to map to California grade 7. Items rated at 1 were assumed to map to California grade 5.

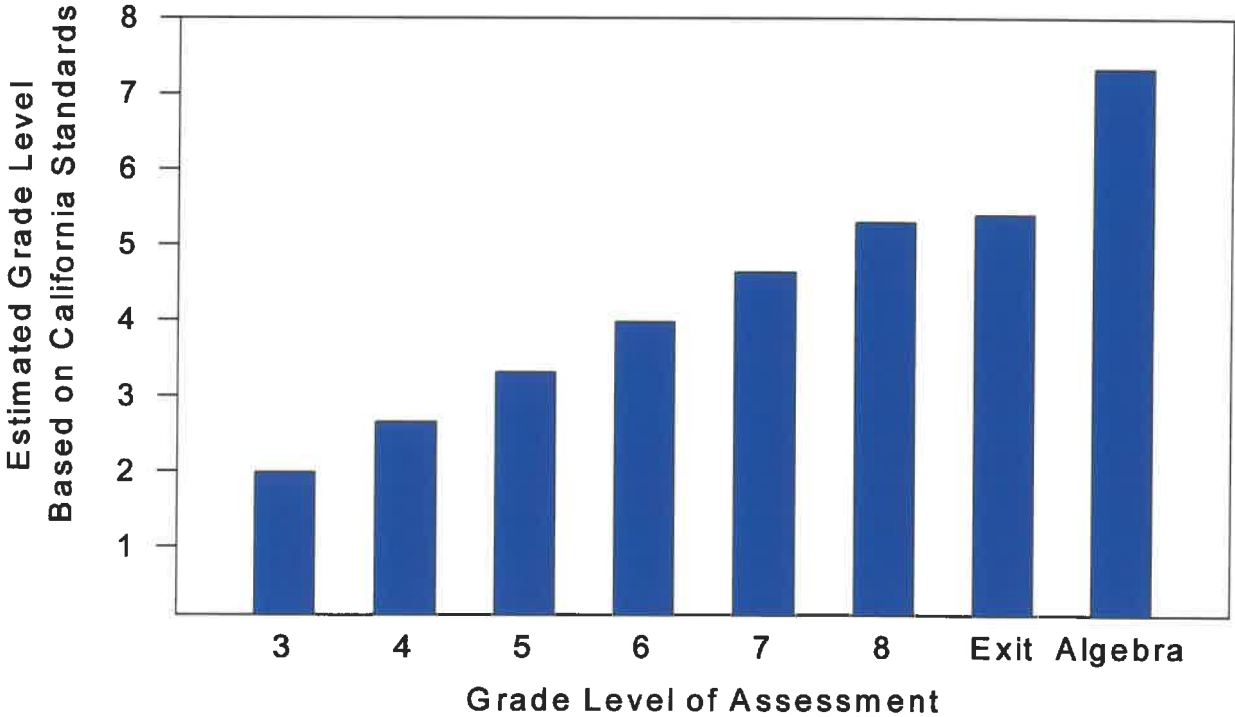
The results of the computations based on these assumptions are indicated below.

This preliminary model suggests that there is a discontinuity between the grade 8 (or even the exit level) TAAS exam and the algebra end-of-course exam. This is true even though the ratings placed the algebra end-of-course exam at a low level vis-a-vis California Standards (grade=7.34).

The implication of this finding is that Texas students who are minimally competent through grade 8 may have a difficult time mastering the content of algebra since this constitutes roughly a two-year jump in the model above. This is consistent with the fact that a smaller proportion of Texas students meet minimal expectations for the end-of-course algebra exam than do so for the TAAS mathematics exams.

While this finding should be considered preliminary based on the required assumptions, the implications are quite serious and demand further study. The possibility exists that the assessment system is not targeting the skills and understandings that are sufficient for success in algebra even at the modest levels addressed by the end-of-course exam. If this is true, then the assessment system is not contributing to the improvement in achievement that might otherwise be obtained.

Estimated Exam Grade Levels Based on California Standards



Summary and Conclusions

An Anecdote: It was a warm evening in Arizona, and some of the vacationers at the hotel were lounging by the Jacuzzi. A 14-year-old from Texas chanced to strike up a conversation with a math instructor from California. The instructor asked the teenager about the TAAS exam in Texas. *“Oh, that . . .”* replied the youth, *“well, the math is very easy and that’s all they teach us, so it gets pretty boring.”*

Texas has maintained a relatively stable statewide mathematics assessment system, and there is evidence to suggest that some improvements have been made in achievement. The percentage of students meeting minimum expectations on the TAAS has been going up, and the National Assessment of Educational Progress (NAEP) scores for Texas look promising relative to the rest of the country. But, judging by international comparisons like the Third International Mathematics and Science Study (TIMSS), doing better than other parts of the U.S. may not be saying much.

The evidence reviewed above is consistent in indicating that Texas assesses mathematics achievement at a low level. Indeed, the content of the high school exit exam is more appropriate as a target for the sixth grade. Students could pass this exam and yet have difficulty with the Japanese exam given to 12-year-olds.

Whether or not establishing low targets is a good way to stimulate achievement in mathematics statewide is a critical strategic question. The consequences for far too many students may be similar to the experience of the teenager in the anecdote above. Low-level objectives are not consistent with the high expectations for mathematics achievement that are being called for from one end of the country to the other. Low-level objectives are unlikely to bring student achievement up to the level of our international competition.

The “teaching to the test” phenomenon is often used in an undifferentiated attack on large scale assessment in general. But, this position only belittles the teaching and learning efforts that are motivated by assessment. The assessment system is an effective stimulus for learning to the extent that it promotes greater achievement. Students who acquire the knowledge and skills they need to do well on a challenging test have indeed been learning. Testing itself is not evil, but bad tests can be evil in terms of their consequences. They are a misuse of the power of assessments.

The review of the examinations used in Texas is suggestive of a system wherein the power of statewide assessments has focused on raising achievement only to a minimal level. The low expectations evidenced by the exam items themselves, and the fact that instruction is geared toward these exams, is cause for concern. This concern is amplified by the indications that the system design may be insufficient to promote greater success in algebra and higher level mathematics courses.

Without successful achievement in algebra and beyond, students are failing to reap the rewards of

mathematics education. Algebra is often referred to as a gateway course since it opens the door to the opportunities and benefits that are associated with achievement in mathematics. Mathematics Equals Opportunity, a recent report from the U.S. Dept. of Education, cited the need to:

- *Provide all students the opportunity to take algebra I or a similarly demanding course that includes fundamental algebraic concepts in the 8th grade and more advanced math and science courses in all four years of high school.*
- *Build the groundwork for success in algebra by providing a rigorous curriculum in grades K-7 that moves beyond arithmetic and prepares students for the transition to algebra.*
- *Ensure that all students, parents, teachers, and counselors understand the importance of students' early study of algebra as well as continued study of rigorous mathematics and science in high school.*

The findings reported above raise concern that the mathematics assessment system in Texas is not designed to meet these objectives.

There will always be advantaged students that find other routes to high achievement, but this is not true opportunity for all students. There is an incessant relationship between socio-economic status and achievement in mathematics that bodes poorly for the disadvantaged. However, it is a statistical given that the strength of this relationship will appear to decrease if the top of the achievement distribution is truncated by using a test that is too easy. The inherent risk is that data generated by the statewide assessment system would be misleading with respect to equity issues in education.

The power of a statewide assessment system can be exerted in many ways. In Texas, a sharp focus has been drawn by the minimum requirements for high school graduation. Even the score reports for the early grades are keyed to these minimum graduation requirements as the long-term objective. As a result, the public evaluations of schools and districts tend to focus on achievement rates for these minimal objectives as well. But such requirements necessarily target low achievement objectives since no system that fails most students would be socially acceptable.

Alternative systems of rewards and consequences should be investigated that can target higher achievement levels. For example, the public reports for the performance of schools and districts, even without mandated consequences, are powerful motivators at a system level. Greater emphasis in these reports should be given to achievement relative to higher objectives. Reward systems for increases in high-level achievement rates should also be considered.

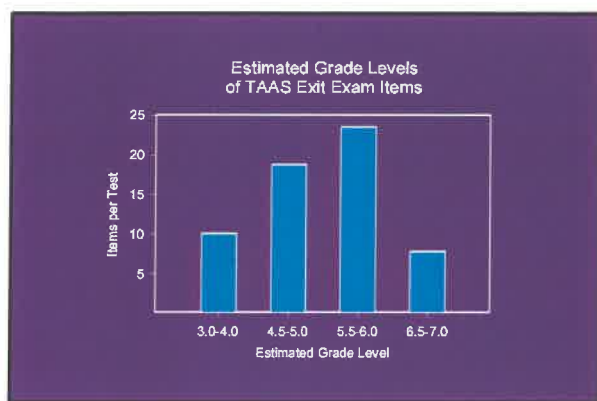
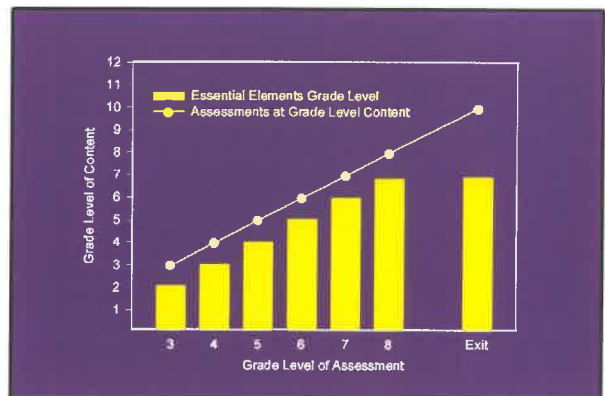
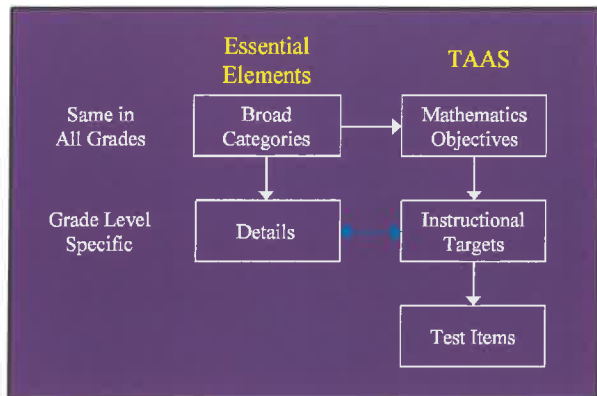
Any design that emphasizes high achievement levels needs to be carefully monitored to insure that the motivational factors do not lead toward artificial claims of success. This is exemplified by an algebra test that is, in reality, a combination of pre-algebra and low-difficulty algebra topics. High achievement goals must represent honest levels of high achievement, not just high achievement in name only.

In Texas, as in many other states, there has been a call for an emphasis on “problem solving and complex thinking skills.” But, also as in many other states, there is difficulty in operationalizing these goals in a way that is objective, can be measured, and shows progressive development across the grade levels. For this reason, these content areas require special scrutiny. There will be a tendency to think of objectives as high-level if they contain language about problem solving and complex thinking. However, this sort of material may not represent advanced achievement in practice. The development in these areas must be tied to more explicit objectives, such as “differentiating between relevant and irrelevant information in a problem situation requiring solution by multiplication or division.” In this way, the objectives in problem solving and complex thinking need to be tied to the mathematics content objectives at the appropriate level.

Much of the above discussion has focused on the need to employ the power of a statewide assessment system to promote high achievement levels. At more modest levels, there is evidence to suggest that the system of assessments in Texas has been effective. By comparison to California’s new standards, most Texas high school graduates can achieve at least at a fifth- or sixth-grade level. But, the possibility exists that even this benefit may have run its course and that any further gains will reflect diminishing returns.

In summary, the system of mathematics achievement assessment in Texas emerges as a powerful model but one that is too highly focused on minimal achievement. The incentives for improvement that accompany the statewide assessment system do not emphasize high achievement sufficiently. In fact, the design of the assessment devices themselves doesn’t even permit the measurement of high achievement levels with any degree of accuracy. Without a substantial adjustment to the objectives that are evidenced by the exam items themselves, it seems unlikely that the assessment system will effectively promote the kind of achievement necessary for students to realize the full benefit of a rigorous mathematics education.





TAAS Exit Exam Content Over 4 Years

Addition and subtraction of fractions with unlike denominators
 3 addition items
 3 subtraction items
 all denominators are small integers

Multiplication and division of fractions or mixed numbers
 0 multiplication items
 1 division item

Terminating and repeating decimals
 0 items

Greatest common factor, least common multiple
 0 items directly address this

TAAS Exit Exam Content Over 4 Years

Prime and composite numbers
 0 items directly address this

Powers, roots, and exponents
 1 square of 15
 1 square of 40
 1 find the integers that bound the root of a number

Properties of real numbers
 2 items

Absolute value
 0 items

TAAS Exit Exam Content Over 4 Years

Area
 1 item on lateral surface area, formula given

Volume
 1 item on volume of a rectangular prism

Median and mode
 1 item on median

Solving equations
 2 items
 $1.5x - 6 = 4.5$
 $c = \$15 + \$7.50p$ when c is \$45

TAAS Exit Exam Questions

The total attendance recorded at the 1984 Summer Olympic Games in Los Angeles, California, was 5,797,923. What is this number rounded to the nearest thousand?

Mrs. Ramos has a plastic cube on her desk that holds photographs. There is a picture on every face of the cube except the bottom. How many pictures are displayed on the cube?

TAAS Exit Exam Questions

What is the approximate length of a new pencil before it is sharpened?

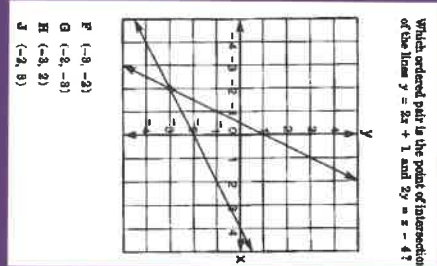
- 1.9 millimeters
- 19 millimeters
- 19 centimeters
- 1.9 meters

Devon's house is on a rectangular block that is 330 yards long and 120 yards wide. What is the distance around his block?

TAAS Exit Exam Questions

Kenyon is 5 feet 6 inches tall. His sister Tenika is 7 inches taller than he is. How tall is Tenika?

At a restaurant Steve ordered food totaling \$6.85. If he paid with a \$20 bill, how much change should he receive?



Which expression is equivalent to $5(3 + x)$?

- F $3 + 5x$
- G $15x$
- H $15 + x$
- J $15 + 5x$

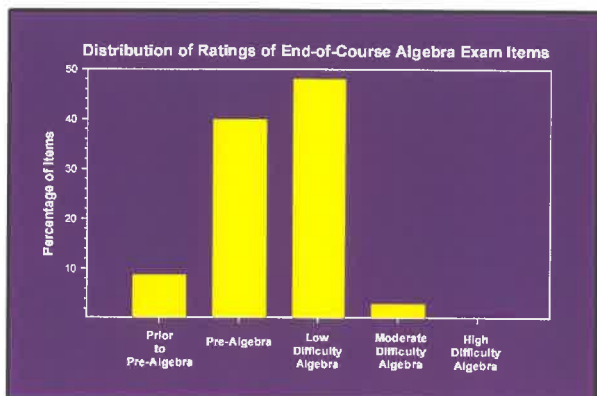
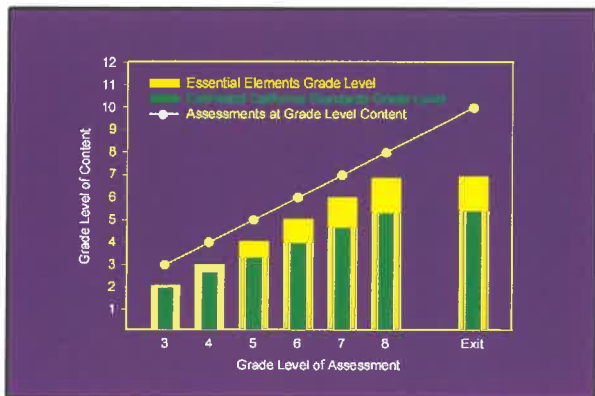
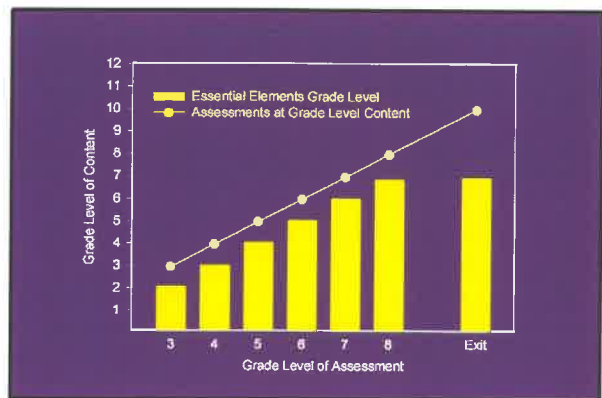
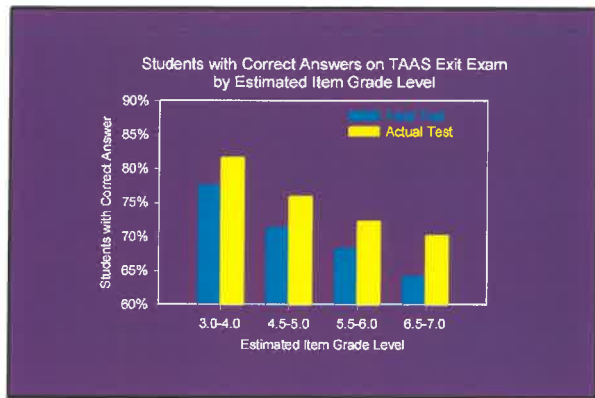
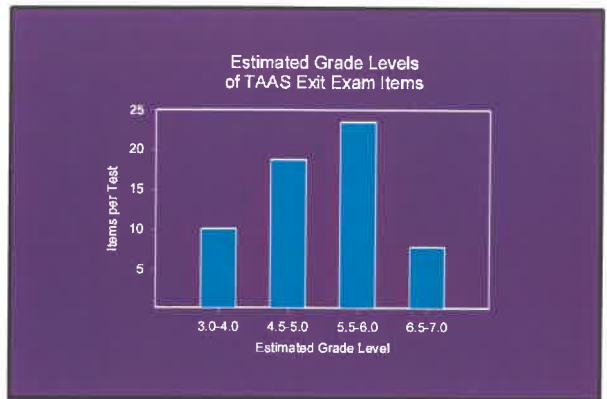
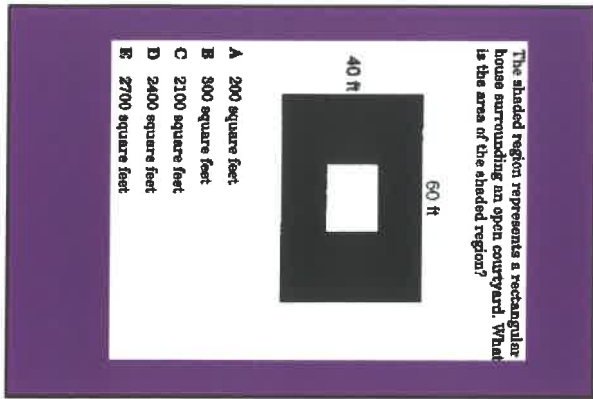
DEPARTMENT BUDGET

Year	Money Spent
1	\$8,050
2	\$8,850
3	\$9,750
4	\$10,725

A small department of a local government has the financial history shown in the table.

Which is the best prediction of the department's budget for Year 6?

- A \$12,700
- B \$12,550
- C \$11,800
- D \$11,050
- E \$11,000



End-of-Course Algebra Exam Weaknesses

- * There are too many formulas given. Most, such as the area of a rectangle or the slope of a line, should be assumed to be known.
- * There are no rational function reduction or arithmetic items.
- * There is almost no confirmation of algebraic arithmetic skills at all: products of binomials, reduction of fractions involving exponential monomials, etc.
- * There are no problems that require factoring polynomials, even "taking out" a common factor.

End-of-Course Algebra Exam Weaknesses

- * Standard word problems that lead to algebraic solution are inadequate in number and in depth.
- * Although there are enough items, linear graphing is inadequate; e.g., there are no items that involve the slopes of perpendicular lines.
- * Many items can be done by checking given answer choices or by inspection. Even just a "none of the above" would be helpful.
- * The distance formula is never used and all Pythagorean Theorem items can avoid it.

End-of-Course Algebra Exam Weaknesses

- * Radical equation exercises are trivial or don't exist.
- * Scattergram questions are trivial and should not be included.
- * "Which BEST describes ..." language is used even when a perfect fit is among the choices.
- * The response-grid format is insufficiently used. It is down to one item in 1998.

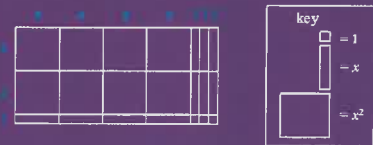
* ... too many formulas given. Most ... should be ... known

Formula Chart Algebra I End-of-Course Examination

Perimeter	square	$P = 4s$
	rectangle	$P = 2(l + w)$
Circumference	circle	$C = 2\pi r$
Area	square	$A = s^2$
	rectangle	$A = lw$ or $A = bh$
	triangle	$A = (b h) / 2$
	trapezoid	$A = 1/2 (b_1 + b_2) h$
	circle	$A = \pi r^2$

* ... no problems ... require factoring polynomials

A rectangle with an area of $8x^2 + 10x + 3$ is modeled below:



What are the dimensions of the rectangle in terms of x ?

$$(4x + 3)(2x + 1)$$

* ... word problems ... are inadequate ... in depth.

The perimeter of a square is 64 centimeters.
What is the area of the square?

- A 8 cm²
- B 16 cm²
- C 32 cm²
- D 256 cm²
- E 4096 cm²

* Many items can be done by checking given answer choices . . .

The length of time required by a high-speed printer to print a large set of documents is given by the equation

$$x^2 - 3x - 54 = 0$$

where x is the time in hours. How many hours are required to print the set of documents?

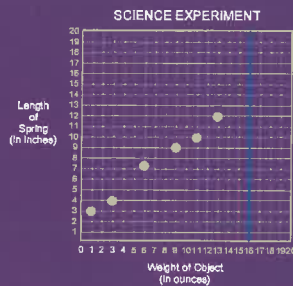
- F 2 hr
- G 3 hr
- H 6 hr
- J 9 hr
- K 18 hr

* . . . linear graphing is inadequate

Which is an equation for the line passing through the point (1,4) and having a slope of 2?

- A $y = 2x + 2$
- B $y = 2x + 5$
- C $y = 2x + 3$
- D $y = 2x - 2$
- E $y = 2x - 5$

* Scattergram questions are trivial . . .



. . . which is the best prediction of the length of the spring with a 16-ounce object attached?

- A 12 in.
- B 14 in.
- C 16 in.
- D 18 in.
- E 20 in.

End-of-Course Algebra Exam Weaknesses

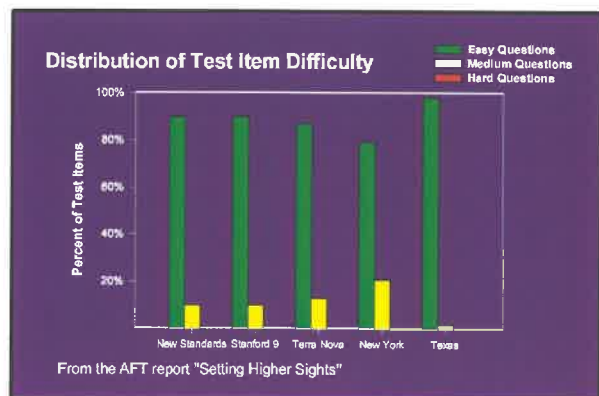
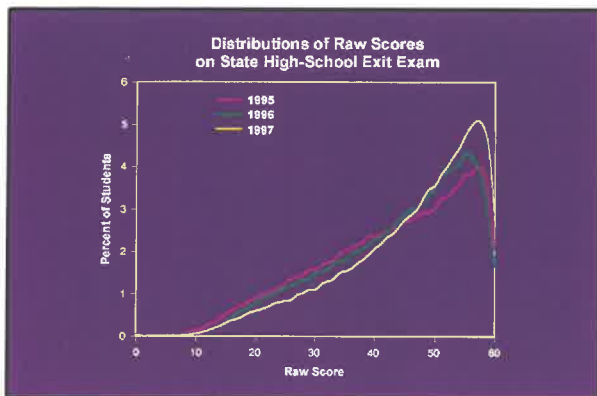
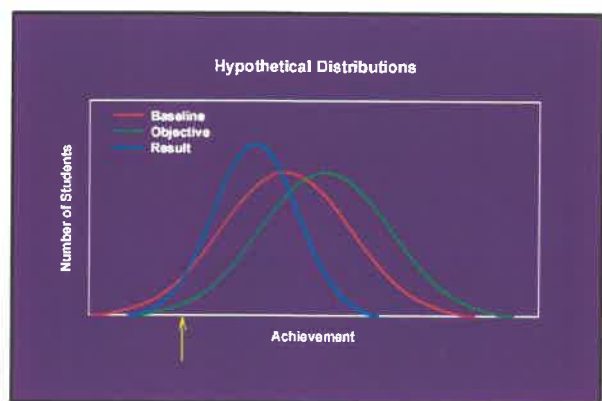
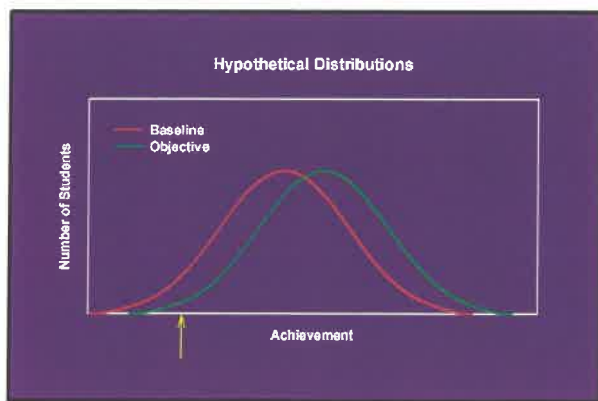
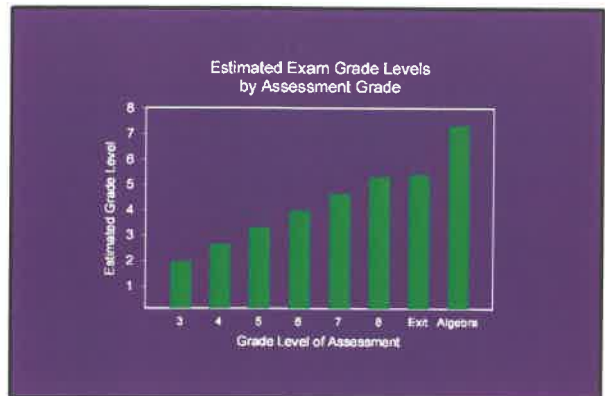
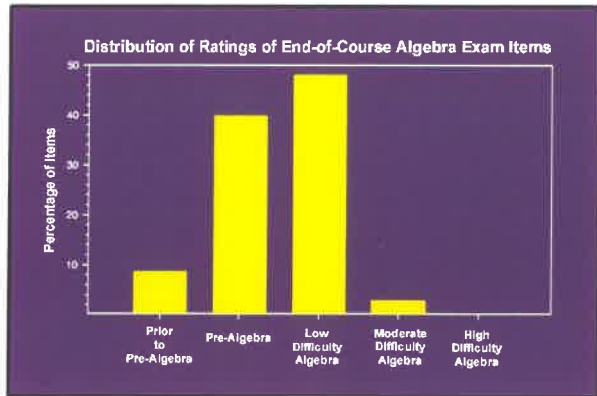
- * There are too many formulas given. Most, such as the area of a rectangle or the slope of a line, should be assumed to be known.
- * There are no rational function reduction or arithmetic items.
- * There is almost no confirmation of algebraic arithmetic skills at all: products of binomials, reduction of fractions involving exponential monomials, etc.
- * There are no problems that require factoring polynomials, even "taking out" a common factor.

End-of-Course Algebra Exam Weaknesses

- * Standard word problems that lead to algebraic solution are inadequate in number and in depth.
- * Although there are enough items, linear graphing is inadequate; e.g., there are no items that involve the slopes of perpendicular lines.
- * Many items can be done by checking given answer choices or by inspection. Even just a "none of the above" would be helpful.
- * The distance formula is never used and all Pythagorean Theorem items can avoid it.

End-of-Course Algebra Exam Weaknesses

- * Radical equation exercises are trivial or don't exist.
- * Scattergram questions are trivial and should not be included.
- * "Which BEST describes . . ." language is used even when a perfect fit is among the choices.
- * The response-grid format is insufficiently used. It is down to one item in 1998.



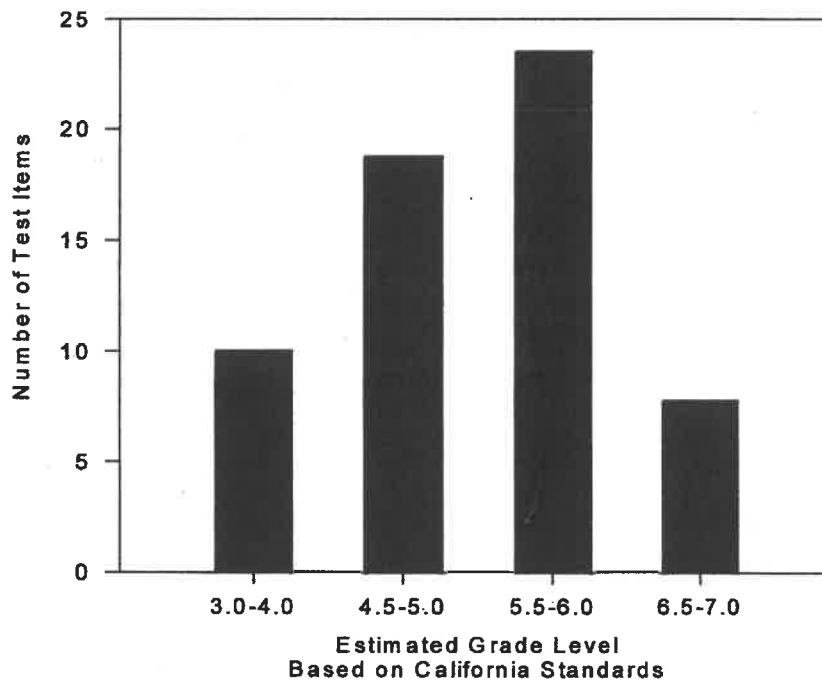
Judging the TAAS Exit Exam Items

To assess the target grade level of the TAAS exams against external criteria, individual exit exam items were evaluated as to grade level based on the newly established California Mathematics Standards. These standards provided a desirable benchmark for several reasons:

- They were designed carefully to be on-track with the best international competition, including Japan and Singapore.
- They are perhaps the most highly detailed of all sets of state mathematics standards, greatly facilitating item evaluation.
- They have been judged as the best available mathematics standards among all sets of state standards, even exceeding those from Japan (R. Raimi and L. Braden, State Mathematics Standards, Fordham Report, Vol 2 #3, <http://www.edexcellence.net/standards/math.html>).

Two of the authors (P.C. and D.K.) independently judged the grade level of every TAAS Mathematics exit exam item for the prior four years using the California standards as a guide. Although some degree of subjectivity was involved, these ratings proved to have a reasonably high level of rater reliability ($r=.813$). When the two ratings did not concur, they were averaged. The average distribution of item grade levels on a TAAS exit exam is illustrated below.

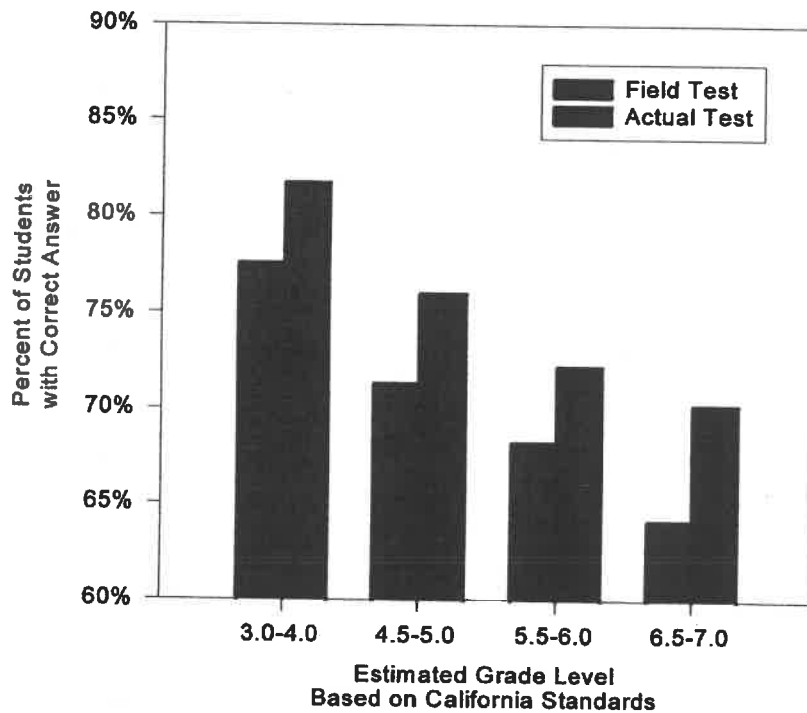
**Number of Items on TAAS Exit Exam
by Estimated Grade Level
Based on California Standards**



Percent Correct by Item Grade Level Rating

If the grade level ratings of TAAS exit exam items based on the California standards are a valid indicator, then TAAS exam scores would presumably be elevated by the inclusion of items below grade level. Likewise, the percentage of students with correct answers on individual items ought to vary as a function of item grade level ratings. Analysis of data for individual items from field tests and actual tests in Texas shows this is indeed the case as indicated below.

**Percentage of Students with Correct Answers
on TAAS Exit Exam
by Estimated Grade Level
Based on California Standards**

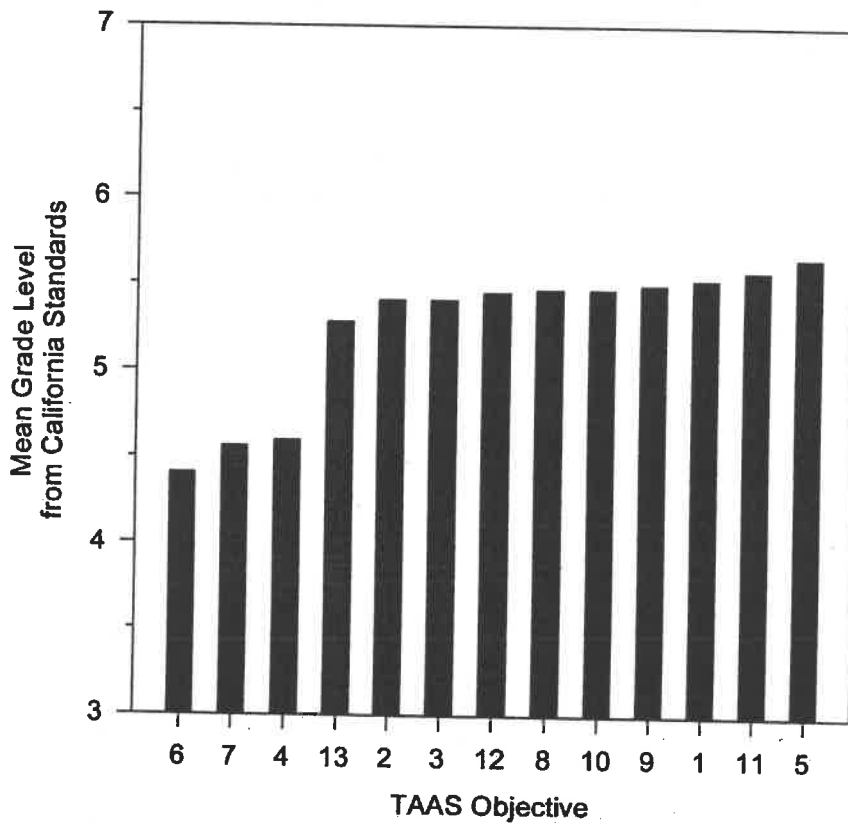


These results suggest that passing rates on the TAAS exit exam would be considerably lower if all items fell at the 7th grade (or higher) level based on the California standards.

Grade Level Ratings of TAAS Exit Exam Items by Objective

To study grade level expectations across content areas on the TAAS exit exam, the mean grade level rating was computed for each of the Mathematics Objectives indicated on the TAAS. This is possible because each TAAS item is tied to a Mathematics Objective. The grade level means are illustrated below.

**Mean Estimated Grade Levels
For TAAS Exit Exam
Based on California Standards
for Each TAAS Objective**



Mean grade level ratings were significantly lower for the Measurement, Addition, and Subtraction Objectives than for the other Objectives ($p < .05$ in each case). This may not be a surprising result as the topic areas of measurement, addition, and subtraction might be expected to enter the curriculum earlier than some of the other Objectives. Although more complex material could enter these content domains, this does not appear to be the case from the grade level ratings. Thus, it appears that the emphasis given by the weights in these three Objectives is grade level inappropriate in the exit exam. Likewise, Objectives 2 and 3, which should reflect algebraic and geometric material, do not appear to reflect more advanced content as might be expected.

Technical Issues: Equating Tests for Difficulty

The proportion of students meeting the minimum expectations on the TAAS exams has been rising. This is a positive indication, but the conclusion relies heavily on equating derived scores for test difficulty differences year to year. If the tests cannot be shown to be of equivalent difficulty (in derived scores), then the finding of improvement may be suspect.

TAAS exams are roughly equated for difficulty in the process of item selection. Furthermore, derived scores are further equated through statistical procedures that are mathematically sound. Since tests contain only a finite number of items, at least small fluctuations in the difficulty associated with a passing (minimum expectations) score are certain. Some further inaccuracy in the process of equating difficulty must exist. However, the degree of accuracy in the equating process is not documented in the Technical Digest. Furthermore, the tests are equated year to year, so that errors in this process may be compounding over time.

To address the equivalence of TAAS exit exams year to year, a comparison among the mean item grade-level ratings based on the California standards was computed for four years of exams. While this does not address the statistical equating process, it does provide an indication of relative item difficulties across years. The results of this test were not statistically significant ($p = .46$), with grade level means hovering around 5.3 for each of the test years. Thus, this method supported the notion that the tests were of roughly equal difficulty levels across years.

Technical Issues: Texas Learning Index Scores

Another design concern relates to the use of the Texas Learning Index (TLI). The design of the TLI is extremely useful in that it provides scores that students can use to gauge their relative achievement progress over the grades. Although a valuable resource, the TLI should be recognized as an essentially norm-based score for grades 3 to 8. This means that any unevenness in achievement across the grade levels in the norming year will be perpetuated in the TLI score distributions in succeeding years. For example, if the gain in mathematics achievement in the sixth grade was low as measured in the spring of 1994 due to some curriculum defects at the time, then a "passing" TLI score of 70 for grade 6 will continue to reflect this low level of achievement in future years. It is even possible that these future scores would be seen as endorsing achievement rates that would not be found if a criterion-referenced system was in place.

CURRICULUM VITAE

David Klein
Department of Mathematics
California State University
Northridge, California 91330-8313
818 677-2721

EDUCATION

Ph.D.	Cornell University	1981	Applied Mathematics
M.A.	U.C. Santa Barbara	1976	Mathematics
B.S.	U.C. Santa Barbara	1975	Physics
B.A.	U.C. Santa Barbara	1975	Mathematics

PROFESSIONAL EXPERIENCE

1993 to date	Professor, California State University, Northridge
1990 to 1993	Associate Professor, California State University, Northridge
1988 to 1990	Assistant Professor, California State University, Northridge
1987 to 1988	Visiting Assistant Professor, U.S.C.
1986 to 1987	Visiting Assistant Professor, U.C.L.A.
1981 to 1986	Assistant Professor, Louisiana State University, Baton Rouge

MEMBERSHIP

American Mathematical Society
International Association of Mathematical Physics
Mathematically Correct

FIELDS OF RESEARCH

Mathematical Physics, Statistical Mechanics, Probability

COURSES TAUGHT

Graduate courses: Functional Analysis, Theory of Differential Equations, Introduction to Analysis, Measure Theory, Topology, Dynamical Systems

Undergraduate courses: Calculus (all levels), Linear Algebra, Differential Equations, Probability, Statistics, Mathematical Modeling, Business Mathematics, Mathematics for Prospective Elementary School Teachers, Complex Variables, Differential Equations Partial Differential Equations, Real Analysis, Applied Mathematics, Mathematics Appreciation, Algebra and Trigonometry

SELECTED PROFESSIONAL ACTIVITIES

Cofounder and Codirector of the CSUN Mathematical Physics Institute
Referee for the Journal of Mathematical Physics
Referee for The College Mathematics Journal

MASTERS THESES DIRECTED

Jeanette Marquand	Metric Spaces and Fractals	1992
Keith Barker	Measure Theory and Its Applications	1993
Maria Coulson	The Mathematical Theory of Music: An Overview	1995
Lisa Whelan	The Mathematical Theory of Baire Spaces with...	1995

MATHEMATICS AND PHYSICS PUBLICATIONS

1. Dobrushin Uniqueness Techniques and the Decay of Correlations in Continuum Statistical Mechanics, *Communications in Mathematical Physics* 86, p. 227-246 (1982).
2. Convergence of Grand Canonical Gibbs Measures, *Communications in Mathematical Physics* 92, p. 295-308 (1984).
3. Uniqueness of One-Dimensional Continuum Gibbs States, *Communications in Mathematical Physics* 97, p. 595-604 (1985).
4. Uniqueness of Continuum One-dimensional Gibbs States for Slowly Decaying Interactions, *Journal of Statistical Physics* 43 no.s 1/2 p. 207-218 (1986).
5. A Class of Continuum Models with no Phase Transitions, *Journal of Mathematical Physics* 29, p. 2872-2875 (1986).
6. (with Wei-Shih Yang) Absence of Phase Transitions for Continuum Models of Dimension $d > 1$, *Journal of Mathematical Physics* 29, p. 1130-1133 (1988).
7. (with Wei-Shih Yang) A Note on the Critical Dimension for Weakly Self-Avoiding Walks, *Probability Theory and Related Fields* 79, p. 99-114 (1988).
8. (with Wei-Shih Yang) Phase Transition for a One-Dimensional Lattice Gas with Hard-Core, *Journal of Statistical Physics* 52, p. 735-745 (1988).
9. (with Wei-Shih Yang) Self-Repelling Processes Associated with Brownian Motions, *Journal of Functional Analysis* 84, p.322-342 (1989).
10. (with Wei-Shih Yang) Some Correlation Inequalities for Ising Antiferromagnets, *Journal of Statistical Physics* 57, p. 1049-1058 (1989).
11. (with Wei-Shih Yang) Low Temperature Smoothness of the Pressure for Antiferromagnets and Other Models, *Journal of Mathematical Physics* 32, p. 2181-2185 (1991).
12. (with Wei-Shih Yang) Classical Antiferromagnets: Correlation Inequalities and Remarks on the Phase Diagram, *Proceedings of the Conference on Probability Models in Mathematical Physics*, (Colorado Springs 1990) World Scientific Publishing Co. (1991).
13. (with Wei-Shih Yang) Absence of First Order Phase Transitions for Antiferromagnetic Systems, *Journal of Statistical Physics* 70 p.1391-1400 (1993).
14. (with Wei-Shih Yang) A Characterization of First-Order Phase Transitions for Superstable Interactions in Classical Statistical Mechanics. *Journal of Statistical Physics* 71 p. 1043-1061 (1993).
15. (with Lincoln Chayes) A Generalization of Poisson Convergence to 'Gibbs Convergence' with Applications to Statistical Mechanics." *Helvetica Physica Acta* 67 p.30-42 (1994).
16. (with Peter Collas) An Ergodic Adding Machine on the Cantor Set, *L'Enseignement Mathematique* 40, p.249-266 (1994).
17. (with Peter Collas and Hans-Peter Schwebler) Convergence of Hamiltonian Systems to Billiards, *Chaos* 8 p. 466-474 (1998).

VITA
Wayne W. Bishop, PhD

Personal Information:

Signature 

Address: 2125 Mar Vista Ave., Altadena, CA 91001
Telephone: (626) 794-1380

Marital Status: Married - Judith A. Bishop
Children: Jeffrey (31), Anthony (26), Melissa (9)

Employer: California State University, Los Angeles
Address: 5151 State University Drive, LA, CA
Telephone: (213)343-2159
FAX: (213)343-5071
Email: wbishop@calstatela.edu

Social Security No. 479-50-2750 Passport Number: B648551

Educational Preparation:

Ph.D. 1971 Mathematics, Western Michigan University, Kalamazoo, MI
M.A. 1967 Mathematics, Northwestern University, Evanston, IL
B.A. 1964 Mathematics, University of Northern Iowa, Cedar Falls, IA

Employment Experience:

1971-present Department of Mathematics and Computer Science, California State University, Los Angeles, CA 90032 Professor of Mathematics
Duties: Teaching a wide range of undergraduate and graduate Mathematics and Computer Science Courses, academic advisement, various university committees and administrative positions within department including department chair, 1983-1989.

1994-1995 Pasadena City College, Pasadena, CA
1980-1981 Duties: Teaching undergraduate Mathematics

1967-1971 Western Michigan University, Kalamazoo, MI
Duties: Graduate Assistant and Research Fellow

- 1981 Elementary Linear Algebra (with S. Venit); Prindle, Weber, Schmidt; textbook for sophomore level Linear Algebra.
- 1978 "How to Construct a Regular Polygon"; Am. Mathematical Monthly, March, Vol.85, No.3.
- 1977 "How to Construct a Regular Polygon"; to annual meeting of AMS/MAA; St. Louis, Mo.
- 1976 The Calculus Problem Book (with S. Venit); Kendall-Hunt; Calculus problem solving book.
- 1975 "Conversational Programming - An Instructional Plus", International Journal of Mathematics Education in Science and Technology.
- 1975 "A Theory of Multiplicity for Multiplicative Filtrations"; Journal fur die reine und angewante Mathematik (Crelle's Journal); Band 277, Seite 8, bis 26.

Other Professional Activities:

- 1996 Invited testimony to CA Assembly Education Committee, Chair Steve Baldwin
- 1995 Mathematics Task Force of CA State Supt. D. Easin
- 1992 External Reviewer for Program Review, California State Polytechnic University, Pomona
- 1992 External Reviewer for Program Review, California State University, Northridge
- 1990 External Reviewer for Program Review, California State University, Sacramento
- 1989 Project Director for CSU Resource Guide - Subject Matter Assessment for Prospective Teachers - Single Subject Mathematics
- 1989 Advisory Panel to Commission on Teacher Credentialing - Mathematics
- 1988 Member of Editorial Board, Journal of the Institute of Mathematics and Computer Sciences, Calcutta, India
- 1981 Fulbright Professor to Instituto Tecnológico y de Estudios Superiores de Monterrey, Mexico City - taught classes and led faculty seminars in Numerical Analysis and Computer Science for the academic year.

Paul L. Clopton
5224 Lodi Street
San Diego, CA 92117-1132
(619) 273-2819

Personal Data

Residence: 5224 Lodi St., San Diego, CA, 92117
Home Phone: (619) 273-2819
Office Phone: 858 (619) 552-8585x3534
Email: pclopton@ucsd.edu

Education

BA, UCLA Department of Psychology 1972
MS, SDSU Department of Psychology 1975
Postgraduate studies, UCLA Department of Education, 1975-1982

Professional Experience

1995-present	Cofounder and Member, Mathematically Correct Review, consultation, and direction roles for this mathematics education advocacy group
1975-present	VA Medical Center, La Jolla/UCSD School of Medicine Biomedical Research Statistician Statistical consultant to all medical departments
1998-present	Member, California Mathematics Content Panel Oversee development of statewide mathematics tests
1996-1997	Author, Western Pennsylvania Algebra Contest Exams
1997	Member, 1997 California Mathematics Curriculum Framework and Criteria Committee
1997-present	Member, San Diego Grade-Level Mathematics Standards Committee
1993-1996	Consultant, University of the World, La Jolla, CA Provide consultation regarding education and technology issues to this non-profit organization dedicated to the distribution of high-tech education worldwide
1977-present	Tutor, statistics and experimental design

Loredo, R., Longo, C., Salonen, D., Yu, J., Haghghi, P., Trudell, D., Clopton, P.L. & Resnick, D. (1995) Glenoid labrum: MR imaging with histologic correlation. *Radiology*, 196(1), pp. 33-41.

Yang, B.Y., Sartoris, D.J., Djukic, S., Resnick, D. and Clopton, P.L. (1995) Distribution of calcification in the triangular fibrocartilage region in 181 patients with calcium pyrophosphate dihydrate crystal deposition disease. *Radiology*, 196(2), pp.547-550.

Parry, B.L., Hauger, R., Lin, E., Le Veau, B., Mostofi, N., Clopton, P.L. & Gillin, J.C. (1994) Neuroendocrine effects of light therapy in late luteal phase dysphoric disorder. *Biological Psychiatry*, 36, pp. 356-364.

Klemfuss, H. & Clopton, P.L. (1993) Seeking tau: A comparison of six methods. *Journal of Interdisciplinary Cycle Research*, 24, pp. 1-16.

Brugge, K.L., Grove, G.L., Clopton, P.L., Grove, M.J., & Piacquadio, D.J. (1993) Evidence for accelerated skin wrinkling among developmentally delayed individuals with Down's syndrome. *Mechanisms of Ageing and Development*, 70, pp. 213-225.

Park, Y.H., Huang, G.S., Taylor, J.A., Marcelis, S., Kramer, J., Pathria, M.N., Clopton, P.L. & Resnick, D. (1993) Patterns of vertebral ossification and pelvic abnormalities in paralysis: A study of 200 patients. *Radiology*, 188, pp. 561-565.

Kin, K., Lee, J.H., Kushida, K., Sartoris, D.J., Ohmura, A., Clopton, P.L. & Inoue, T. (1993) Bone density and body composition on the Pacific rim: a comparison between Japan-born and U.S.-born Japanese-American women. *Journal of Bone and Mineral Research*, 8, pp. 861-869.

Chon, K.S., Sartoris, D.J., Brown, S.A. & Clopton, P.L. (1992) Alcoholism-associated spinal and femoral bone loss in abstinent male alcoholics as measured by dual X-ray absorptiometry. *Skeletal Radiology*, 21, pp. 431-436.

Patterson, T.L., Smith, L.W., Grant, I., Clopton, P., Josepho, S. & Yager, J. (1990) Internal vs external determinants of coping responses to stressful life events in the elderly. *British Journal of Medical Psychology*, 63, pp. 149-160.

Smith, L.W., Patterson, T.L., Grant, I. & Clopton, P.L. (1989) A shortened MMPI useful for psychiatric screening of the non-institutionalized elderly. *Journal of Clinical Psychology*, 45, pp 359-365.

Campbell, S.S., Gillin, J.C., Kripke, D.F., Erikson, P. and Clopton, P.L.: (1989) Gender differences in the circadian temperature rhythms of healthy elderly subjects: relationships to sleep quality. *Sleep*, 6, pp. 529-536.

Pineda, C.J., Martinex-Lavin, M, Goobar, J.E., Sartoris, D.J., Clopton, P.L. & Resnick, D. (1987) Periostitis in hypertrophic osteoarthropathy: relationships to disease duration. *American Journal of Roentgenology*, 148, pp 773-778.

of naloxone on methylphenidate-induced mood and behavioral changes: a negative study. *Psychopharmacology*, 67, pp. 125-130.

Janowsky, D.S., Huey, L.Y., Clopton, P.L. & Judd, L.L. (1980) Effects of methylphenidate on interpersonal perceptions. Presentation to the 12th CINP Congress, Goteborg, Sweden, pp. 189.

Janowsky, D.S., Clopton, P.L., Huey, L.Y. & Judd, L.L. (1980) Methylphenidate improves perceived empathy, warmth and genuiness. *Journal of Psychiatric Research* (abstract).

Janowsky, D.S., Clopton, P.L., Leichner, P., Abrams, A. & Judd, L.L. (1979) Interpersonal effects of marijuana: a model for the study of interpersonal psychopharmacology. *Archives of General Psychiatry*, 36, pp 781-785.

Clopton, P.L., Janowsky, D.S., Clopton, J.M., Judd, L.L. & Huey, L.Y. (1979) Marijuana and the perception of affect. *Psychopharmacology*, 61, pp 203-206.

Janowsky, D.S., Leichner, P., Parker, D., Judd, L.L., Huey, L.Y. & Clopton, P.L. (1978) Methylphenidate and serum prolactin in man. *Psychopharmacology*, 58, pp 43-47.

Janowsky, D.S., Leichner, P., Parker, D., Judd, L.L., Huey, L.Y. & Clopton, P.L. (1978) The effect of methylphenidate on serum growth hormone: influence of antipsychotic drugs and diagnosis. *Archives of General Psychiatry*, 35, pp 1384-1389.

Janowsky, D.S., Leichner, P., Clopton, P.L., Judd, L.L. & Parker, D. (1978) Comparison of oral and intravenous methylphenidate. *Psychopharmacology*, 59, pp 75-78.