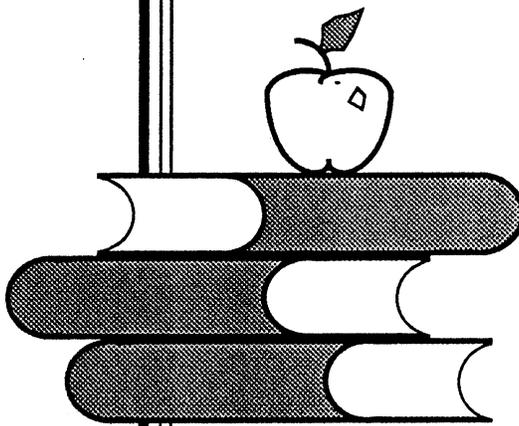


Missouri Mastery and Achievement Tests

*Steven J. Osterlind
Principal Author*

Technical Manual

Prepared by William R. Merz



*Center for Educational Assessment
University of Missouri-Columbia*

Missouri Mastery and Achievement Tests

Steven J. Osterlind, Ph. D.
Principal Author
Associate Professor of Education and
Director, Center for Educational Assessment
University of Missouri
403 South Sixth Street
Columbia, Missouri 65211
(314) 882-4694

Copyright © by Missouri Department of Elementary and
Secondary Education. All rights reserved. Printed in the
United States of America. No part of this publication may
be reproduced, stored in a retrieval system, or transmitted,
in any form or by any means, electronic, mechanical,
photocopying, recording, or otherwise, without the prior
written permission of the publisher.

Table of Contents

Background	1
Purpose of a Technical Manual	1
Introduction	1
Description of the Test Battery	2
Excellence in Education Act of 1985	3
Test Development	6
Purpose of the Tests	6
Chronology	6
Groups Involved	7
Development Process	8
Test Content Specifications	8
Item Writing	8
Item Review and Revision	9
Item Pool	10
Field Testing	10
Further Review and Revision	11
Preliminary Statistical Analysis	11
Additional Reviews	11
Final Version of <i>Missouri Mastery and Achievement Tests</i>	11
Administration of Forms A, B, and D	12
Score Reports	12
Norms, Scores, and Comparability	13
Norms	13
State Sample	13
Summary Statistics	13
Distribution of p -Values, Biserials, and Thresholds	14
p -Values and Thresholds	14
Biserial Correlations	16
Scores	17
Subject and Cluster Scaled Scores	17
Key Skill Mastery Information	18
Estimated Comparable National Percentile Ranks	18
Comparability	18
Linking	19
Concordance	19
Validity	21
Content-Related Evidence	21
Criterion-Related Evidence	22
Construct-Related Evidence	22

Reliability and Errors of Measurement	24
Raw and Scaled Score Subject Reliabilities	24
Scaled Score Cluster Reliabilities	24
Key Skill Information	24
Manuals and Reports	26
Manuals	26
<i>Missouri Mastery and Achievement Tests</i> Score Reports	26
Individual Student Report	27
Individual Student Score Label	28
Pupil List Report	28
Grade Level Key Skill Report	29
Grade Level Cluster Report	30
Summary Report	31
Chapter 1 Eligibility Report	31
Conclusions	32
References	33
Appendix A: Key Skill Abbreviations and Clusters	34
Appendix B: Samples of Test Content Specifications	47
Appendix C: Raw Score and Scaled Score Statistics for Subject Areas	63
Appendix D: Distribution of Scaled Score Statistics	71
Appendix E: Distribution of <i>p</i> -Values, Biserials, and Thresholds	85
Appendix F: Pearson Product Moment Coefficients for Corresponding Iowa Test of Basic Skills/Test of Achievement an Proficiency and <i>Missouri Mastery and Achievement Tests</i> Examinees	99
Appendix G: Common Factor Analysis, Grade 10 Form A	103
Appendix H: Mastery Classification Reliability Estimates	110
Appendix I: MMAT Manuals and Reports	114

Background

Purpose of a Technical Manual

The purpose of a technical manual is to provide the user and the reviewer with the necessary information to evaluate the appropriateness and the efficacy of a given test for specified uses. A technical manual describes how the test is to be used, what generalizations may be made from the data generated by the test, and the proper cautions which should be exercised in utilizing information gathered by the test. The *Missouri Mastery and Achievement Tests* Technical Manual presents this information, as well as other relevant material.

Introduction

This manual describes the inception of the *Missouri Mastery and Achievement Tests* (MMAT) battery after the "Excellence in Education Act of 1985," its development, and its technical characteristics. This manual further demonstrates that the developers of the MMAT made every effort to ensure that the battery would meet the *Standards for Educational and Psychological Testing* (AERA/APA/NCME, 1985). At each step, the guiding principle for development was that the MMAT must provide reliable and valid measures of student achievement on the Key Skills listed in *Core Competencies and Key Skills for Missouri Schools* (1986).

As testimony to the MMAT's adherence to these accepted standards and to the curriculum guidelines presented in *Core Competencies and Key Skills*, this manual offers information on a variety of psychometric properties and test development procedures:

The first section gives a general overview and background of the MMAT;

The second section provides descriptions of test development and revision procedures, including the statewide involvement of teachers, test content specifications, item writing, field testing, editing, reviewing, and final form composition;

The third section describes the procedures employed for sampling, scaling, equating, and norming;

The fourth section presents data on the evidence of construct, content, and criterion-related validity, as well as evidence that the test meets appropriate primary, conditional, and secondary standards;

The fifth section presents data on standard errors of measurement and reliability, both traditional and item response theories (IRT), as well as evidence that the test meets appropriate primary, conditional, and secondary standards; and

The sixth section describes the MMAT report formats and supporting documentation.

Because many of these topics are interrelated, a particular topic may appear in more than one part of the manual.

Description of the Test Battery

The *Missouri Mastery and Achievement Tests* (MMAT) is a battery of criterion-referenced achievement tests that evaluate educational objectives identified by the Missouri Department of Elementary and Secondary Education in *Core Competencies and Key Skills for Missouri Schools* (1985). The entire test battery, developed by the Center for Educational Assessment at the University of Missouri-Columbia (CEA) in cooperation with the Department, consists of 34 distinct subject tests assessing student performance in grades 2 through 10. The test for grade 2 covers two subjects:

- Reading/Language Arts
- Mathematics

Tests for grades 3 through 10 cover four subjects:

- Reading/Language Arts (grades 3-6) or English/Language Arts (grades 7-10)
- Mathematics
- Science
- Social Studies/Civics

The test is divided into three increasingly specific levels: subjects, clusters, and skills. The broadest level, subjects, is composed of the four academic areas listed above. Subjects are divided into clusters of related content. Each cluster is then further divided into the individual Key Skills necessary for mastering that cluster. (Appendix A presents the Key Skills which make up each cluster.) Grade 2 mathematics is not divided into clusters. Table 1 below presents the clusters for each subject.

Table 1

Subjects and Clusters

Language Arts/Reading/English	Mathematics
Language Arts Reading Writing	Understanding Numbers Computation Measurement and Geometry Interpretation and Application
Science	Social Studies/Civics
Life Science Earth Science Physical Science Cross Disciplinary	Geography History Government Economics Civics

The battery consists of multiple-choice items with four options each. The items are distributed across three equivalent test forms (A, B, and C) for grades 3, 6, 8, and 10 and two equivalent test forms (D and E) for grades 2, 4, 5, 7, and 9. Each of the subject tests is administered in two parts, thereby allowing students a rest break during each subject test.

The subject tests for each grade are bound together in a booklet. Grade 2 tests are presented in machine-scored booklets that have been specially designed to accommodate the requirements of students at that level and that eliminate the need for separate answer sheets. The color-coded test booklets for grades 3 through 10 have matching machine-scorable answer sheets.

Excellence in Education Act of 1985

The MMAT is the product of an extensive effort to fulfill the provisions of the state's Excellence in Education Act of 1985 (160.257 RSMo). This broadly based legislative mandate was initiated to enhance and improve instruction and curriculum in Missouri schools. Of the Act's 37 provisions, Section 4, the testing provision, probably has the broadest impact in that it requires annual, statewide, criterion-referenced testing on educational objectives identified by the Department.

Specifically, Section 4 of the Act requires that the Department identify Key Skills in the areas of English/reading/language arts, science, mathematics, and social studies/civics. All local school districts are required to implement an assessment program which includes the annual administration of approved criterion-referenced tests to all pupils in at least two nonconsecutive

cutive grade levels within the grade span two through six, and to all pupils in at least two nonconsecutive grade levels within the grade span seven through ten. Districts are expected to use the resultant data in monitoring student progress on the Key Skills and in identifying areas for instructional improvement. In addition, the Act requires that the Department administer annually a criterion-referenced achievement test to a randomly selected sample of public school pupils to assess their performance on the Key Skills. The Department is to report these results to the General Assembly each year.

Even before the act became law, the Department set out to identify Core Competencies and Key Skills for the major subjects in grades 2 through 10. To assure that the Key Skills were commonly recognized among Missouri educators as essential to the academic progress of the student and central to the respective subjects, the Department invited teachers and administrators throughout the state to participate in their development. The four Core Competencies Committees, formed by the Department on February 27, 1985, completed their major tasks on October 17, 1985. The results of their work have since been brought together in *Core Competencies and Key Skills for Missouri Schools*, a publication of the Department which is available upon request from CEA. The *Core Competencies and Key Skills for Missouri Schools* was published in October 1986, and it will be reviewed periodically.

The Core Competencies and Key Skills are arranged into three levels. The first level reflects those Key Skills thought to represent core elements in most districts' curriculum. These are not minimums; rather, they identify learning outcomes deemed essential to successful progress at a particular grade level and in succeeding grade levels. The MMAT focuses exclusively upon the first level of Key Skills. The second level contains learning outcomes of major importance for local assessment. The third level contains related, but non-essential, competencies that students might know. Many outcomes on the Key Skills in levels two and three are not amenable to assessment with multiple-choice tests, so the MMAT does not address them. Districts are instructed to monitor levels two and three locally.

The Core Competencies and Key Skills Oversight Committee convened for the first time on February 28, 1985. Composed of 47 members, all of whom are educators, leaders in business and industry, or state officials, the committee was set up to oversee efforts for implementing the testing provision of the Excellence in Education Act.

In November 1985, the Department contracted with CEA to develop a criterion-referenced test based on the Key Skills. CEA had been involved previously in the Department's test development activities with the *Basic Essential Skills Test* and the *Missouri Kindergarten Inventory of Developmental Skills*. In addition, CEA has printed, distributed, and scored tests as well as reported results for the Department.

Formerly known as the Missouri Testing and Evaluation Service, a division of the University of Missouri-Columbia College of Education, CEA has engaged in test development, administration, and scoring for more than fifty years. Its various other activities range from optical scanning of examination and evaluation forms to psychom-

etric research. Presently, it supplies testing services and material to school districts, colleges, professional groups, and businesses across the United States and in several foreign countries. A catalog describing the array of CEA testing and educational material is available upon request.

Test Development

This section describes how the MMAT was developed. It presents in detail the purpose of the test and links the test and its development to the Excellence in Education Act of 1985. It describes the groups involved in the development of the MMAT and the process of test development. Last, it describes the completed version of the test administered to students in Missouri schools.

Purpose of the Tests

The MMAT battery, like the other components of the Excellence in Education Act of 1985, is devised to encourage and promote quality in Missouri schools. The test battery contributes to this goal by furnishing data for an annual report made to the Missouri General Assembly by the Department of Elementary and Secondary Education. This report presents summary data concerning the academic achievement of students in the state and identifies overall trends in test scores. A representative sample of students, approximating ten percent of those taking the tests during the spring administration, serves as the basis for this monitoring of scholastic progress in Missouri. The "state sample" is selected from school districts throughout the state by means of a statistical procedure developed by CEA. The sample is representative of particular characteristics of the state's entire student population. The sampling plan accounts for location of districts by geographic region, size of districts, ethnic composition of districts, and urban versus rural location of districts. To preclude bias in the sample identification, codes rather than names are used to designate individual buildings and districts. No school district, building, or pupil is identified by name. The result is a statistically valid representation of all students in the state.

The scores derived from the tests are also used at the local level to accomplish the aims of the Excellence in Education Act of 1985. The score information can assist teachers in tailoring personalized instruction that reflects each student's actual achievement in particular areas. MMAT scores can be used to assist counselors as they guide students into the most constructive class schedules. In addition, the test scores provide school principals and district officials with a gauge calibrated to both state and national standards to assist them in making decisions about curriculum and instruction. In serving these functions and many others ordinarily associated with standardized achievement tests, the MMAT is improving the quality of Missouri's schools.

Chronology

The MMAT was developed in two phases. Phase I, grades 3, 6, 8, and 10, was developed during 1985 through 1987. Phase II, grades 2, 4, 5, 7, and 9, was developed during 1986 through 1988. Virtually identical development processes were used in the two phases. The tests for grades 3, 6, 8, and 10 were first administered in 1987, while the tests for grades 2, 4, 5, 7, and 9 were first administered in 1988.

Groups Involved

A test development effort such as this one requires large numbers of individuals to complete specified tasks at each stage of test development. While many individuals were involved in developing the MMAT, the groups described below were formally identified to perform the functions indicated.

The extent of participation by teachers in the development of the MMAT is quite remarkable. In addition to enhancing the relationship between curriculum and assessment, widespread involvement of Missouri educators ensured that the tests were crafted with a firsthand knowledge of the state's school children and a special sensitivity to the educational growth of its young people. After the test objectives were determined, creation of the test battery itself required contributions by hundreds of Missouri educators at every level and scrutiny and revision of numerous drafts by CEA and the Missouri Department of Elementary and Secondary Education.

More than 200 teachers and administrators from schools throughout Missouri were involved in item writing. Another group of more than 200 teachers reviewed the items written by their colleagues for agreement with the Key Skills and for quality of the item stem and its choices. Teachers and administrators were involved in every phase of test development. In addition, a sample of students who had participated in the field trials was interviewed with structured-interview protocols to assess which features of the items had prompted their responses.

Consultants from the Department of Elementary and Secondary Education were involved with test development at all stages. Department consultants worked with teachers, administrators, and CEA staff in defining test content and in writing items, as well as producing the field trial and final versions.

The CEA staff supplemented its expertise with numerous consultants who worked on specific aspects of test development. A group of 12 nationally noted college professors provided special expertise in curriculum, content, and learning theory. This group worked initially to define the test content specifications and then reviewed the item development as the process progressed. A panel of five testing experts from across the country served as the Technical Advisory Committee and provided input on issues concerning test construction, scaling, equating, norming, item analysis, reliability, and validity. Members of that panel also individually completed specialized tasks and studies during test development. A nationally recognized specialist in sampling design devised the sampling plans employed in this effort. A number of individuals examined items for bias. In short, CEA utilized the skills of no fewer than 18 nationally recognized individuals to amplify its professional efforts.

Finally, state officials and leaders in business, industry, and education were involved as an Oversight Committee during the implementation of Section 4 of the Excellence in Education Act of 1985. Thus, implementation of the testing provisions of the Act was completed by utilizing the services of the broadest sampling of the community involved in education.

Development Process

Test Content Specifications

A committee composed of 12 professors, three each from reading, social studies, mathematics, and science, wrote content specifications for each of the 688 Key Skills in both phases of development (Phase I, 353; Phase II, 335). Particular Key Skills were assigned according to an individual's content or curricular expertise. The professors were instructed to elaborate, interpret, and clarify each Key Skill so that its particular characteristics might be clearly and uniformly understood by everyone involved. These elaborations formed the initial draft of the test content specifications.

Next, approximately 120 teachers from all grades throughout Missouri reviewed and further refined the specifications drafted by the college content area specialists. Test content specifications writing workshops were convened for both phases: December 4-5, 1985, for Phase I; and September 7-9, 1986, for Phase II. Workshop participants reviewed specifications in terms of their applicability to classroom instruction. For each phase, teachers met in 16 groups; a group of five to seven teachers met for each grade level and for each content area. Phase I groups reviewed reading and English/language arts, mathematics, science, and social studies specifications for grades 3, 6, 8, and 10. Phase II groups reviewed specifications for grades 2, 4, 5, 7, and 9. Curriculum consultants from the Department of Elementary and Secondary Education assisted the teachers in reviewing specifications. The reviewers' comments and the exemplary items they produced were compiled to describe each skill according to content domain and grade level.

This revised version of the test content specifications was then reviewed by the college content area specialists, CEA editorial staff, and Department consultants. Finally, each test content specification was reviewed by Department consultants and two to five teachers for specific content at a particular grade level. In total, more than 1,000 reviews of test content specifications occurred before they reached their final form. The final test content specifications, examples of which are included in Appendix B, "Samples of Test Content Specifications," were prepared; the complete document, *Test Content Specifications, Grades 2-10*, is available from CEA.

Item Writing

A test item writing workshop was conducted from January 6 through 10, 1986, for Phase I. For Phase II separate grade level workshops were held from September through November 1986. As with the Test Content Specifications Conference, teachers were grouped according to their subject areas and grade levels. Test specifications for the grade level and the content area were available to each group.

Hundreds of test items, obtained from public domain item banks, were available to item writers during the workshop. The teacher/authors decided that these items were not

appropriate for assessing the Key Skills, and that the content of the items themselves was not suitable. The items were abandoned in favor of developing original items referenced directly to the test content specifications and Key Skills.

Two test item construction experts taught participants techniques of item writing during the workshops. In addition, consultants for the Department and from CEA worked closely with the teachers in writing items.

By the close of the workshops, 12 to 16 items for each of the 688 test content specifications had been developed. This totaled approximately 4,800 items for each phase, or just under 10,000 items during the two phases.

Item Review and Revision

In each of the two phases of test development, another group of approximately 120 teachers reviewed the items. Content specialists, including college professors, teachers, private consultants, curriculum consultants from the Department, and the CEA staff also reviewed the items. This process insured at least five reviews per item. Each systematic review was conducted using a scorecard which contained specific criteria relating to congruence between the item and the key skill, appropriateness of the reading level, item wording, item format, and appropriateness of distractors. CEA editorial staff and consultants from the Department of Elementary and Secondary Education analyzed the reviews and synthesized the results. Modifications of particular items were made when appropriate. In some instances, these exhaustive reviews resulted in the removal of test items from further consideration.

A reading specialist completed studies of the readability of items by using standard formulas such as Flesch, Dale-Chall, SMOG, and FOG. Additionally, for grades 2 and 3, a panel of 15 classroom teachers reviewed the test items and the directions to ensure that the syntax and vocabulary would be understood by children of this age. Of course, some items, such as those that involved straight mathematical computation, could not be reviewed for readability.

For both phases, item bias reviews, both judgmental and empirical, were undertaken. A panel with nationally accepted expertise in test item bias judged items for bias. Approximately 150 classroom teachers were also asked to rate items systematically with specially constructed rating sheets to point out stereotyping or other objectionable wording which might bias items. Each item was rated approximately five times by classroom teachers who examined how each item might affect the way a child with a culturally or linguistically different background might respond. For each bias review, clearly delineated criteria were used. Items identified as biased were either discarded or flagged as needing editorial repair. Additionally, CEA editorial staff adjusted items to ensure a mix of genders in the name references in items, a mix of traditional and non-traditional gender roles, and a variety of ethnic backgrounds. Similar, though not identical, processes were used in both phases of test construction.

Item Pool

The initial test item pool contained approximately 9,600 items. The text alone for each phase required some 70,000 lines. Just under one-third of the items required graphics, and a large number required text to accompany items. Five graphic artists completed approximately 1,500 graphics for the items for each phase. More than 800 galleys were set.

Field Testing

During Phase I, four field test forms were created: W, X, Y, and Z. Items were assigned to forms according to the following rules:

1. Form W would always have four items per key skill tested. The remaining items for Key Skills would be distributed across Forms X, Y, and Z so that each had equivalent numbers of items, content, and complexity of items.
2. Items which might give clues to correct answers to other items in the pool would not be placed in the same form.
3. The number of items linked to a passage or a graphic would increase as the grade level increased.

A total of 64 field test booklets was constructed: four forms for four content areas for four grade levels. Each booklet went through 26 steps of editing, revision, and production. Each question went through two rounds of editing with at least four persons reading each item. A data base was established for every item, referencing response options, graphics, and text to the item stem. The same Key Skills were covered in each of the four forms. Four or more items were assigned to each Key Skill in every form. Graphics and art work were further edited to satisfy the demands of page design.

Field tests were conducted between April 28 and May 2, 1986, with a representative sample of school districts throughout the state. Geographic location, school size, and socio-economic status comprised the criteria for selection within the sampling design. The field trial included approximately 3,500 students from about 350 schools in 250 school districts. Each form was administered at its designated grade level and at the grade level below and above to assure measurement of a broad range of abilities. In addition, the *Iowa Tests of Basic Skills* (ITBS) for grades 3, 6, and 8 and the *Tests of Achievement and Proficiency* (TAP) for grade 10 were administered to provide criterion-related validity information and pilot-study data for obtaining Chapter 1 usable norms. Each student in the field trial sample was tested in only one subject area of the MMAT and the corresponding area of the ITBS or TAP.

For Phase II, field trial Forms M, N, and O were created using the rules and procedures previously described. These forms were administered from April 20 to 24, 1987. A sampling

design similar to that used in Phase I was employed, and numbers of students taking the test were virtually identical.

Further Review and Revision

Preliminary Statistical Analysis

Initial item response data were utilized to evaluate items in the field testing forms. Items with distractors drawing no responses, those which had more than 80 percent of pupils responding correctly, or those which had less than 40 percent of pupils responding correctly were identified as needing revision or as inappropriate for use in the first round of testing.

Additional Reviews

During Phase I, about 400 items were chosen for review with students who had taken the test during May, 1986. These items were neither excellent nor poor; rather, they were in the middle group, an unknown quantity. For three weeks, members of the CEA staff and Department consultants visited three to four classrooms a day throughout the state to discuss with students their opinions and reactions to particular test items. These in-class discussions were conducted at grades 3, 6, 8, and 10, and from them reviewers tried to determine the quality and validity of particular test items. Using a structured interview form, interviewers discussed responses to items with the students. Each interviewer completed a protocol for every item discussed. Information gathered in this way was used to conduct a qualitative analysis of some of the items.

Teams of two teachers and one Department consultant reviewed items during June, 1986 for Phase I and during June, 1987 for Phase II, using the Key Skills, the test content specifications, teacher review comments, item statistics from the field testing, and classroom review data. From 12 to 16 items per Key Skill, the teams narrowed the number of items to seven for Phase I and six for Phase II. By the end of this process, nearly half of the original 9,600 items had been discarded. This exhaustive process of item development, review, and revision resulted in a pool of items of exceptional quality. The language of each item was clarified; its congruence to a specific skill was verified; and its ability to be understood by students was examined. Rarely is such a thorough process of item development used, and rarer still is the result a pool of items with such exemplary characteristics.

Final Version of MMAT

The final version of the MMAT consists of thirty-four criterion-referenced tests which assess student performance on Key Skills in grades 2 through 10. The grade 2 tests cover Language Arts/Reading and Mathematics. Subject tests for grades 3 through 10 include Reading/English/Language Arts, Mathematics, Science, and Social Studies/Civics. The battery consists of multiple-choice items with four options each. There are three equivalent test

forms (A, B, and C) for grades 3, 6, 8, and 10 and two equivalent test forms (D and E) for grades 2, 4, 5, 7, and 9. Each subject test is administered in two parts to allow a rest break. In each part items are ordered from easy to difficult.

CEA used a common-items equating design to ensure that the final forms for each grade level were indeed parallel. Forms A, B, and C for grades 3, 6, 8, and 10 underwent a field trial during the fall of 1986 in order to obtain information about testing time, administration procedures, and equivalency of forms within a grade level.

Administration of Forms A, B, and D

In the Spring of 1987, approximately 240,000 students in grades 3, 6, 8, and 10 were administered Form A. In the Spring of 1988, approximately 240,000 students in grades 3, 6, 8, and 10 were administered Form B and approximately 200,000 students in grades 2, 4, 5, 7, and 9 were administered Form D.

Score Reports

MMAT results are reported on a variety of complementary report forms to help make the MMAT useful for educators at every level. The Individual Student Report presents one student's results for each subject and its associated clusters, as well as Key Skill mastery data. The Individual Student Score Label, a self-adhesive label for the permanent record, lists a student's subject scaled score and estimated comparable national percentile rank. The Pupil List Report, a roster of all students in a grade, presents key skill mastery information for every student. The Grade Level Key Skill Report, the Grade Level Cluster Report, and the Summary Report present aggregate data for Key Skills, clusters, and subjects, respectively, at the building and district level. The optional Chapter 1 Eligibility List identifies students eligible for Chapter 1 services in reading, language arts, and/or mathematics. In addition, a school district may elect to receive other reports, such as summary data organized by particular variables or pupil list reports broken down by classrooms. Reports are also available on computer diskette or tape.

Norms, Scores, and Comparability

This section describes the procedures used to develop norms, derive scores, and ensure comparability among forms and grade levels. It is divided into three subsections: the norm group (its selection and specific MMAT performance characteristics); the scores themselves and how they were derived; and the methods of establishing comparability (specifically, the methods used to link the MMAT to other tests and to examine for concordance among forms).

Norms

The *Standards for Educational and Psychological Testing* (AERA/APA/NCME, 1985) state that “interpreting scores can often be aided by the availability of scoring scales and norms that relate raw scores to defined theoretical or empirical distributions” (p. 31). Norms for the MMAT are based on the performance of the state sample. This subsection will describe the process used in selecting this group. Next, it will report the summary statistics for the performance of the group. Finally, it will discuss distributions of p -values, biserial correlations between item responses and total score, and item thresholds.

State Sample

At each of the grades 2 through 10, about 6,000 students representing about 10 percent of the students at a specific grade level are designated each year to be part of the “state sample.” The Department of Elementary and Secondary Education reports their performance on the Key Skills to the Missouri General Assembly, as well as utilizing the data for various technical analyses.

A stratified random cluster technique in which schools are stratified by geographic location, school size, and socio-economic status is used to select students for participation in the sample. Within strata, the equivalent of two-stage cluster sampling is used with schools as clusters. Schools are selected within each stratum by simple random sampling, and proportional allocation of school sample size to strata is used.

Summary Statistics

Appendix C, “Raw Score and Scaled Score Statistics for Subject Areas,” presents summary statistics resulting from the 1987 administration of Form A and the 1988 administrations of Forms B and D to the state sample. The number of items, raw score mean, raw score standard deviation, KR_{20} reliability, scaled score mean computed by item response theory (IRT), scaled score standard deviation, IRT reliability estimate, and the range of IRT standard errors of measurement are listed for each subject.

The number of items per subject ranges from 32 for the Mathematics Test (Grade 2) to 112 for the Reading Test (Grades 9 and 10). The mean raw score for Mathematics, Grade 2 is

25.38; for Reading, Grade 9 it is 78.45. These values indicate that the tests are neither exceptionally difficult nor exceptionally easy for students at the targeted grade levels. The number of items per cluster ranges from 16 to 68. However, raw score means and standard deviations are not reported for clusters.

Appendix D, "Distribution of Scaled Score Statistics," reports data for both subjects and clusters. At grades 3, 6, 8, and 10 one may assess for growth in mean performance by comparing mean scaled score performance on Form A to that on Form B (the comparable forms administered to students in grades 3, 6, 8, and 10 in 1987 and 1988). Growth is shown in all clusters but Cross Disciplinary in the Science test at grade 10. Here, the difference is -3 scaled score points. In no other cluster are the differences smaller than +6 (Language Arts, Grade 6), or larger than +68 (Computation, Grade 6). For subjects, growth ranges from a minimum of +8 (Reading, Grade 6) to a maximum of +36 (Mathematics, Grade 8).

The difference in relative magnitude of raw score and scaled score standard deviations can be accounted for by the number of students giving correct responses to items of higher difficulty and greater discrimination. A student answering correctly the same number of items as another may have a higher scaled score because he or she responded correctly to more difficult items with greater discrimination.

Distribution of *p*-Values, Biserials, and Thresholds

Appendix E, "Distribution of *p*-Values, Biserials, and Thresholds," reports item statistics for subject and cluster scores for Forms A, B, and D. *P*-values and biserials are traditional measures of item difficulty and item discrimination, whereas thresholds are an IRT measure of item difficulty.

***p*-Values and Thresholds**

P-values simply represent the proportion of examinees responding correctly to an item. Because *p*-values are computed at the item level without regard to other scores, the means for subjects will approximate the means for clusters. The minimum mean *p*-value for subjects is 0.50 for Science, Grade 9, Form D, while the maximum mean *p*-value is 0.83 for Mathematics, Grade 3, Form B. This indicates that on the average somewhere between 50 and 83 percent of the examinees select the correct responses to items. Midpoints among mean difficulty levels for Reading Tests are approximately 0.71; for Mathematics Tests, approximately 0.65; for Science Tests, approximately 0.58; and for Social Studies Tests, approximately 0.68.

At the item level without regard to averages, the minima ranged from 0.09 for Mathematics, Grade 6, Form A, to 0.58 for Mathematics, Grade 2, Form D. The maxima ranged from 0.85 for Mathematics, Grade 9, Form D, to 0.99 for Reading, Grade 8, Form A. The smallest range of difficulties for Reading is 0.32 to 0.93 (mean = 0.68) for Grade 4, Form D. The largest range for Reading is 0.16 to 0.95 (mean = 0.68) for Grade 10, Form A. For Mathematics, the smallest range of difficulties is 0.58 to 0.97 (mean = 0.80) for Grade 2, Form D; the largest

is 0.09 to 0.96 (mean = 0.65) for Grade 6, Form A. For Science, the smallest range is 0.34 to 0.96 (mean = 0.63) for Grade 4, Form D; the largest is 0.17 to 0.97 (mean = 0.74) for Grade 3, Form B. Finally, for Social Studies, the smallest range is 0.37 to 0.92 (mean = 0.65) Grade 10, Form A; the largest is 0.19 to 0.98 (mean = 0.70) for Grade 6, Form A.

To examine item difficulty further, one must also examine threshold values. Thresholds are the points on the IRT item characteristic curve where the probability of obtaining a correct response is 0.50. For subjects and clusters in which item difficulties are approximately normally distributed and are distributed across the range of difficulties, the threshold values typically range from -3 to +3, with -3 indicating easy items and +3 indicating difficult items. MMAT item difficulties are not normally distributed and are not distributed across the range of difficulties. Further, the maximum likelihood method is used to compute scores, amplifying the effect of extremes on the values of thresholds, so that values greatly exceeding -3 or +3 are to be expected. Because IRTs are computed separately for subjects and clusters, the maximum and minimum reported for thresholds in Appendix D for subjects will be different for subjects and clusters.

Mean subject threshold values vary from -1.80 for Mathematics, Grade 3, Form B to 0.57 for Science, Grade 9, Form D. For each of the subjects, the midpoint of the mean thresholds is as follows: -0.97 for Reading, -0.69 for Mathematics, -0.13 for Science, and -0.87 for Social Studies.

The minima range from -6.15 for Science, Grade 3, Form B to -1.92 for Social Studies, Grade 3, Form A. The maxima range from -2.05 for Mathematics, Grade 6, Form B to 16.08 for Reading, Grade 10, Form A. The smallest range of thresholds for Reading is -3.32 to -0.28 (mean = -1.48) for Grade 3, Form B; the largest is -3.68 to 16.08 (mean = -0.74) for Grade 10, Form A. For Mathematics, the smallest range of thresholds is -3.93 to -0.23 (mean = -1.80) for Grade 3, Form B; the largest is -2.00 to 5.16 (mean = -0.24) for Grade 8, Form A. For Science, the smallest range of thresholds is -4.42 to 1.26 (mean = -.75) for Grade 4, Form D; the largest is -3.39 to 10.96 (mean = 0.45) for Grade 10, Form A. Finally, for Social Studies, the smallest range is from -1.92 to 2.16 (mean = -.64) for Grade 3, Form A; the largest is -2.92 to 8.59 (mean = 0.45) for Grade 4, Form D.

Mean thresholds ranges are as follows: Reading/English, -1.48 for Grade 3, Form B to -0.74 for Grade 10, Form A; Mathematics, -1.80 for Grade 3, Form B to -0.24 for Grade 8, Form A; Science, -1.38 for Grade 3, Form B, to 0.57 for Grade 9, Form D; and for Social Studies, -1.21 for Grade 6, Form B, to 0.45 for Grade 10, Form A.

As stated above, individual item thresholds for the MMAT range beyond the usual ± 3.00 . In fact, the minima range from -6.15 for Science, Grade 3, Form B to -1.92 for Social Studies, Grade 3, Form A, while the maxima range from -2.05 for Mathematics, Grade 6, Form B to 16.08 for Reading, Grade 10, Form A. While this range is large, it is not unusual. Values for clusters follow a similar pattern.

One may conclude from examining the distribution of p -values that on the average 50 to 85 percent of the examinees who took the MMA T responded correctly to items. This is verified by the threshold values obtained for the same data. There is less variability at the subject level than at the cluster level. The items appear to be of sufficient difficulty to estimate an examinee's level of achievement and of sufficient ease to keep an individual motivated to continue responding. A sufficient number of items with higher difficulty test the limits of achievement, yet enough items of moderate difficulty provide reference to the Key Skills.

Biserial Correlations

A biserial correlation is the correlation between students' correct responses to a particular item and students' total scores. For clusters, the biserial represents the correlation between passing the item within the cluster and the total scores on the cluster; for subjects, the biserial represents the correlation between passing the item and the subject total scores. Because clusters are composed of 16 to 52 items, the biserials are quite varied. The difference in numbers of items, as well as the brevity of some of the clusters compared with the number for subjects, tend to weaken the magnitude of the cluster biserials.

Subject mean biserials (the arithmetic mean computed for each of the biserials within the subject) range from 0.28 for Science, Grade 9, Form D to 0.61 for Reading, Grade 3, Form A. For Reading, the mean biserials range from 0.45 for Grade 10, Form B to 0.60 for Grade 3, Form B, with a median of approximately 0.51. For Mathematics, biserials range from 0.46 for Grade 8, Form A to 0.56 for Grade 3, Form B, with a median of approximately 0.50. For Science, biserials range from 0.28 for Grade 9, Form D to 0.49 for Grade 3, Form B, with a median of approximately 0.36. Finally, in Social Studies, biserials range from 0.46 for Grade 5, Form D to 0.55 for Grade 8, Form B, with a median of approximately 0.53. This indicates that, at the subject level, items are discriminating well among the varying degrees of performance reflected by the total scores. Typically, mean biserials of 0.35 are considered to reflect acceptable evidence of adequate discrimination.

Minima and maxima for biserials by subject, however, show greater variation. The minima range from -0.30 for Reading, Grade 10, Form A to 0.35 for Mathematics, Grade 3, Form B. The maxima range from 0.52 for Science, Grade 5, Form D to 0.98 for Reading, Grade 3, Form A. For Reading, the smallest range of biserials is 0.15 to 0.73 for Grade 8, Form A (mean = 0.48); the largest range is -0.30 to 0.80 for Grade 10, Form A (mean = 0.47). For Mathematics, the smallest range is 0.35 to 0.68 for Grade 2, Form D (mean = 0.55); the largest range is -0.10 to 0.73 for Grade 10, Form A (mean = 0.46). For Science, the smallest range is 0.18 to 0.67 for Grade 4, Form D (mean = 0.41); the largest range is 0.06 to 0.85 for Grade 3, Form A (mean = 0.49). Finally, for Social Studies, the smallest range is 0.25 to 0.78 for Grade 8, Form B (mean = 0.55); the largest range is -0.04 to 0.79, for Grade 4, Form D (mean = 0.49). This represents a reasonable spread of biserials. Of course, some of the low biserial values are the product of the level of difficulty of items. Extremely high or low difficulty levels attenuate the magnitude of biserials. Nevertheless, the magnitudes, on the whole, as

reflected by the minimum and maximum values along with the means, are quite acceptable and support the contention that the meticulous effort at item specification and construction was reflected in examinee performance. The care and energy spent on specifying content and writing items seems to have yielded a pool of items that function well in assessing the achievement of students taking the MMAT.

Scores

Three types of scores are reported for individual students taking the MMAT: subject and cluster scaled scores, Key Skill mastery information, and estimated comparable national percentile ranks. Each is based on the number of correct responses. (Beginning in 1989 with Forms C and E, cluster scores for individual students will be reported as High, Medium, or Low.)

Subject and Cluster Scaled scores

Subject and cluster scaled scores are derived using IRT because IRT scaling yields more accurate results than does the commonly used number correct scaling, which is based on classical test theory. A two-parameter logistic model, as implemented by the computer program BILOG (Mislevy and Bock, 1984), is used to compute subject and cluster scaled scores. This particular model makes possible the best estimate of the student's true ability as it relates to the difficulty and the discrimination of items. In fact, the score is a maximized value based on the mathematical relationship of item difficulty distribution and item discrimination distribution.

The model purports to avoid dependence on the particular examinee sample when estimating item difficulty and item discrimination, not to be limited to using the same or parallel test items for estimates of ability, and to yield better estimates of reliability and true ability at all levels of examinee performance. As a result, there is a direct relationship, but not a one-to-one correspondence, between a student's raw score and the scaled score. A student's scaled score represents the relationship of the student's pattern of responses on the entire set of items to the specific characteristics of each item. Cluster scores are scaled independently from subject scores.

For Forms A and D, the means and standard deviations of subject and cluster scaled score distributions for individual students were 300 and 65, respectively. (Beginning in 1989 with Forms C and E, cluster scores for individual students will be reported as High, Medium, or Low.) Subject and cluster means and standard deviations for Form B are given in Appendix C. Means and standard deviations for Forms B, C, and E will differ from 300 and 65 because they reflect growth in student performance from the base years of administration. The base year for Form A, which was administered to grades 3, 6, 8, and 10, was Spring 1987; for Form D, which was administered to grades 2, 4, 5, 7, and 9, the base year was Spring 1988.

Key Skill Mastery Information

Each Key Skill is measured by four items. A student must answer at least three items correctly in order to demonstrate mastery of the Key Skill. The Key Skills mastered along with the number of items passed is reported. The Key Skills not mastered are also reported. The decision to require that three of four items per Key Skill be answered correctly is judgmental rather than empirical. However, because this is a mastery designation, setting the mastery level any lower than 70 percent correct would be problematic. After considering standards based on several different methods, the CEA staff, in consultation with the Commissioner's Advisory Committee on Testing, decided that three out of four correct responses per skill would yield a defensible mastery score and would be compatible with the item difficulties.

Estimated Comparable National Percentile Ranks

Estimated comparable national percentile ranks are derived by comparing the performance of students in the state sample on the MMAT to their performance on the *Iowa Tests of Basic Skills* (ITBS), grades 2 through 8, or the *Tests of Achievement and Proficiency* (TAP), grades 9 and 10. The exact method for establishing comparability is described later in the section on Concordance. Raw scores are used for this procedure because the authors of the ITBS and TAP derive national percentile ranks from raw scores on these two tests. The process used to obtain comparable scores involves relating the sample's distribution of MMAT raw scores to the sample's distribution of ITBS or TAP raw scores. Using the equipercentile equating method, the comparable national percentile ranks derived from state sample data are then used to estimate comparable national percentile ranks for all students taking the MMAT. For Reading and Language Arts, estimated comparable national percentile ranks are derived using cluster raw scores, rather than subject raw scores, in order to serve Chapter 1 needs.

Each student in grade 2 receives an estimated comparable national percentile rank in Reading, Language Arts, and Mathematics. Each student in grades 3 through 9 receives an estimated comparable national percentile rank in Reading, Language Arts, Mathematics, Science, and Social Studies. Each student in grade 10 receives an estimated comparable national percentile rank in all of these areas except Language Arts.

The estimated comparable national percentile rank is not equivalent to a norm-referenced test score, but it can be used appropriately for various purposes associated with Chapter 1 programs, such as identifying students eligible for Chapter 1 services.

Comparability

This section examines the comparability of scores on alternate forms of the MMAT. Forms are considered parallel, and are linked to each other by the process described below.

Linking

In developing multiple forms of the MMAT, it is necessary to link forms to establish their equivalence. In order to accomplish this linkage, any two forms administered contiguously at a grade level share 50 percent of their items in common. In this way, year to year administrations are linked directly to each other and through each other to previous years.

The IRT procedure used to derive subject and cluster scaled scores is also used to link forms. If the data fit the model, one can estimate each examinee's achievement independently of the form of the test administered and regardless of variance in difficulty of the forms. The IRT procedure estimates an examinee's true score on the domain tested, and that true score estimator is independent of the sample of items from which the score was derived. All that is necessary is that the data fit the IRT model and that there be a sufficiently large sample from which to derive true score estimates. With the MMAT, both requirements are met. Samples are in the thousands of examinees, and the two-parameter IRT model fits the data well.

In the linking procedure, two different groups of students take two different forms of the MMAT. The forms share 50 percent of their items, thus creating an anchor test between administrations. For anchored items, ability estimates are computed on each item for each group that took the test. The item characteristic curves for the anchored subset are summed to obtain the ability estimate on the items. Only a transformation is necessary to calibrate score estimates of the anchored items on the same scale. However, because 50 percent of each form's items are mutually independent, ability estimates for these remaining items must be computed. The item characteristic curves for these items are summed and added to those of the anchored portion. A transformation to a common scale enables direct comparison of scores on the two forms.

Concordance

The model employed for generating national norm-referenced information from local objective-referenced information requires a concurrent administration of a norm-referenced test (NRT) and an objective-referenced test (ORT) to a representative sample of the student population. The scores from the two tests are equated using any one of a number of equating methods. The norm-referenced scores are then estimated for the group student population using their objective-referenced scores. Once the equating has been done, only the ORT needs to be administered during subsequent assessments.

Form G of both the ITBS (grades 2 through 8) and the TAP (grades 9 and 10) were chosen for equating with the MMAT. These two vertically-linked NRTs were chosen primarily because they measure content similar to the Key Skills, a critical factor in the success of ORT-NRT equating.

Equipercntile equating was selected for the Missouri application of the model. Preliminary data collected in pilot studies during field tests in which the ITBS/TAP and field test versions of the MMAT were administered indicated that the psychometric properties of the two

batteries were different. Furthermore, ITBS/TAP national percentile ranks are derived from raw score tables, so the equating method selected had to utilize raw scores. Thus, it was not possible to equate using the item response theory two-parameter logistic model that was used to derive MMAT subject-scaled scores.

In order to improve the accuracy of individual student NRT scores, the NRT and the ORT are equated each year rather than only once. A single-group design is used for equating, in an attempt to minimize error. The students composing the state sample are randomly assigned to a group. One subject test of the ITBS/TAP and the entire MMAT are administered to each group. Counterbalanced administrations are not systematically incorporated into the procedure because of logistical constraints. It is possible, however, that some degree of counterbalancing results even without such a provision.

Validity

Validity, according to the *Standards for Educational and Psychological Testing* (1985), “refers to the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores” (p. 9). That is, generalizations, inferences, hypotheses, predictions, and evaluations based upon data gathered with tests must have evidence of appropriateness and efficacy. Evidence for effective use of information generated by scores on the MMAT is classified into three areas:

- Content-related
- Criterion-related
- Construct-related

Content-Related Evidence

Content-related evidence of validity “demonstrates the degree to which the sample of items, tasks or questions on a test are representative of some defined universe or domain of content” (AERA/APA/NCME, 1985, p. 10). This type of validity was built into the MMAT during its development. The test is designed to assess mastery of a well-defined domain of skills outlined in the *Core Competencies and Key Skills for Missouri Schools*. The section of this manual entitled “Test Development” documents the meticulous care taken to reference content to the Key Skills. First, teachers, assisted by content specialists, identified the core competencies and Key Skills. Next, content experts wrote test content specifications for the Key Skills. Teachers from across the state refined and revised the specifications. Their revisions were reviewed by content and measurement specialists. Finally, the test content specifications were published and distributed to school districts throughout the state.

Test items were developed directly from the test content specifications by teachers who were trained to write test items. In order to ensure content validity, each item was reviewed by three teachers and/or content experts using a structured rating schedule to assess congruence with the appropriate Key Skill. Item clarity, readability, and potential for bias were reviewed at the same time. In addition, a panel of linguistics and bias experts reviewed items for systematic miscue and for offensive content. Readability of items was empirically assessed by an independent expert. By the end of this review process, over half of the 9,600 multiple-choice test items originally developed were discarded.

This exhaustive process of item development, review, and revision resulted in a pool of items of exceptional quality. The language of each item was clarified, its congruence to a specific skill was verified, and its ability to be understood by students was examined. Rarely is such a thorough process of item development used, and rarer still is the result a pool of items with such exemplary characteristics. Finally, field tests generated empirical data to assist test developers in selecting items for final forms. All items selected for the final forms were judged to be valid Key Skill measures by at least three content experts.

Criterion-Related Evidence

Criterion-related evidence of validity is demonstrated by data relating the scores of interest to some outside outcome criterion. In this instance, criterion-related evidence can be defined as the degree to which scores on the MMAT agree with two independent measures of the same attribute in contiguous administrations. Form G of the ITBS (grades 2 through 8) and the TAP (grades 9 and 10) were chosen because they were judged to be norm-referenced tests that measured content similar to the Key Skills, a necessary condition in establishing the validity of a new measure.

Appendix F, "Pearson Product-Moment Coefficients for Corresponding ITBS/TAP and MMAT Examinees," reports agreement for concurrent administrations of the MMAT and the ITBS/TAP. Pearson Product Moment Correlation Coefficients range from 0.713 for third-grade Science to 0.860 for sixth-grade Mathematics on Form A; from 0.703 for third-grade Science to 0.871 for sixth-grade Mathematics on form B, and from 0.701 for ninth-grade Science to 0.856 for seventh-grade Mathematics on Form D. These correlations are of acceptable magnitude and indicate that the MMAT measures significantly similar constructs as do the ITBS and the TAP. In fact, the percent of shared variance ranges from a low of 49.1 to a high of 75.8. The magnitudes of the correlations reported provide strong evidence that performance on the MMAT is related to performance on an outside outcome criterion (the ITBS/TAP).

Construct-Related Evidence

Construct-related evidence of validity focuses on the extent to which test scores reflect the psychological characteristics being measured. For purposes of the MMAT, it is important that the items assessing each content area actually measure that content domain and that there is as little cross-over among domains as possible. Therefore, items in a specific test must measure that subject content alone: reading items must measure reading, mathematics must measure mathematics, and so on. In addition, reading ability must only minimally influence performance in the other subject areas. This latter requirement is very important for accurate application of the scaling technique employed; IRT requires as an assumption that each subject test measures only one trait.

To provide construct-related evidence of validity, several factor analytic studies were undertaken. Responses of examinees included in the state sample were submitted to the principal factors technique with oblique rotation. Responses of approximately 6,000 examinees were included in each analysis, exceeding the general rule that the number of respondents should equal 10 times the number of items.

As a first step, an analysis of responses for Grade 10, Form A was completed using data from the 1987 administration. Item intercorrelations were computed for each grade level. These intercorrelation matrices were submitted to principal factors analysis. Squared multiple correlations were used as prior communality estimates. Results of these analyses are reported in Appendix G, "Common Factor Analyses, Grade 10, Form A."

The findings may be summarized in this way. A five-factor solution was generated, accounting for 59.158 percent of the variance. Factors intercorrelated, as would be expected with this solution. Factor 1 correlated with the other four factors with magnitudes of 0.711 to 0.745, indicating that elements measured by items associated with Factor 1 were associated substantially with the other factors as well. Factor 2 correlated very highly with Factor 5, substantially with Factor 1, and moderately with Factors 3 and 4. Factor 3 correlated substantially with Factor 1 and moderately with Factors 2, 4, and 5. Factor 4 correlated moderately with Factors 2 and 3 and substantially with Factors 1 and 5. Finally, Factor 5 had a high correlation with Factor 2, substantial correlations with Factors 1 and 4, and a moderate correlation with Factor 3. From the intercorrelations and the pattern of factor loadings, five factors were identified as reflecting these areas:

- Factor 1— English/Language Arts
- Factor 2— Mathematics
- Factor 3— Inductive Reasoning Involved in Social Studies and Mathematics
- Factor 4— Social Studies/Civics
- Factor 5— Deductive Reasoning Involved in Science, Mathematics, and Social Studies/Civics

Both principal components analysis and principal factors solutions were employed to describe the structure of the battery at each grade level and the subjects within a grade level.

Reliability and Errors of Measurement

According to *Standards* (1985) “Reliability refers to the degree to which test scores are free from errors of measurement” (p. 19). To determine the reliability of the MMAT, several different scores must be examined. To address the reliability of subject and cluster scaled scores, the reliability of subject raw scores will be examined. Then, scaled score reliabilities for subjects will be examined, followed by cluster scaled score reliabilities. Finally, Key Skills mastery classifications will be addressed.

Raw and Scaled Score Subject Reliabilities

Internal consistency estimates (KR_{20}) of subject raw score reliabilities range from 0.846 to 0.950 for Form A, from 0.870 to 0.949 for Form B, and from 0.810 to 0.947 for Form D. These are uniformly high reliability coefficients, and indicate excellent internal consistency reliability. IRT reliability estimates range from 0.891 to 0.957 for Form A, from 0.892 to 0.956 for Form B, and from 0.851 to 0.956 for Form C. Again, these uniformly high coefficients give good evidence of reliability based upon internal consistency estimates. Complete listings of subject reliabilities can be found in Appendix C.

Scaled Score Cluster Reliabilities

IRT reliability estimates for cluster scaled scores are computed by the BILOG program. For Form A they range from 0.668 to 0.944. For grade 3 they range from 0.770 to 0.930; for grade 6 they range from 0.726 to 0.903; for grade 8 they range from 0.752 to 0.920; and for grade 10 they range from 0.668 to 0.944. For Form B, IRT estimates for cluster scaled score reliabilities range from 0.689 to 0.940. For grade 3 they range from 0.724 to 0.916; for grade 6 they range from 0.728 to 0.902; for grade 8 they range from 0.769 to 0.916; and for grade 10 they range from 0.689 to 0.940. For form D, IRT estimates range from 0.556 to 0.920. For grade 2 they range from 0.768 to 0.869; for grade 4 they range from 0.688 to 0.932; for grade 5 they range from 0.689 to 0.896; for grade 7 they range from 0.766 to 0.920; and for grade 9 they range from 0.556 to 0.917. These coefficients are acceptable for all of the clusters, even for those having a smaller number of items. This indicates that pupils performed fairly consistently within each cluster.

Key Skills Information

Reliability estimates for Key Skills are based on the beta binomial model as developed by Huynh (1980). They are reported for MMAT, Form A only.

Raw agreement indices estimate the proportion of examinees who would be classified in the same category in two contiguous testings. Kappa adjusts estimates for the dependency between two equivalent sets of data. Values for raw agreement indices (p) range from 0 to 1, as do Kappa indices. For English/Reading/Language Arts Key Skills across grade levels,

p ranges from 0.565 to 0.870, while Kappa ranges from 0.032 to 0.762. For Mathematics Key Skills, p ranges from 0.550 to 0.909, while Kappa ranges from 0.033 to 0.686. For Science Key Skills, p ranges from 0.004 to 0.875, while Kappa ranges from 0.012 to 0.467. Finally, for Social Studies Key Skills, p ranges from 0.550 to 0.868, while Kappa ranges from 0.078 to 0.495.

The complete set of data is reported in Appendix H, "Mastery Classification Reliability Estimates." These data indicate that the mastery classifications have sufficient indications of reliability to warrant confidence.

Manuals and Reports

The *Standards* (1985) clearly require that the publishers of a test provide sufficient information so that those who use and review the test can do so responsibly. Such material must describe how the test is administered and scored, how scores can be interpreted, and how information is reported to the user. These materials are to be complete, accurate, and clear.

Manuals

The thirteen manuals accompanying the battery describe content, administration, interpretation, and use of the tests and reports comprising the MMAT. Most were developed using teacher input, and all were reviewed by teachers and administrators prior to publication. The manuals, available for distribution to test users and reviewers, are listed in Appendix I, "MMAT Manuals and Reports."

Missouri Mastery and Achievement Tests Score Reports

The MMAT results are reported through a variety of complementary forms, each appropriate for specific educational applications. The following score reports are part of the basic services provided to MMAT users:

- Individual Student Report
- Individual Student Score Label
- Pupil List Report
- Grade Level Key Skill Report
- Grade Level Cluster Report
- Summary Report

Classroom teachers will find the Individual Student Report and the Pupil List Report useful in improving instruction for individual students. The Individual Student Score Label provides a practical mechanism for entering MMAT scores in students' permanent records. The Summary Report and the Grade Level Cluster Report for a district are probably most valuable for district-level administrators who wish to evaluate the district's curriculum and effectiveness of instruction. Building-level administrators can use the Summary, Grade Level Cluster, and Grade Level Key Skill Reports for a building in order to evaluate instructional effects. The district Grade Level Key Skill Report shows student performance on each Key Skill, and curriculum specialists can use this report to review the scope and sequence of curriculum content and to plan changes in instructional emphasis. The Chapter 1 Eligibility

List, an optional report, provides building and district Chapter 1 personnel with information useful in determining eligibility and program effects.

Descriptions of the MMAT basic score reports and the optional Chapter 1 Eligibility List follow.

Individual Student Report

The Individual Student Report presents a student's MMAT results. The Individual Student Report for grade 2 has two sections, one for each subject tested: Reading/Language Arts and Mathematics. The Individual Student Report for grades 3 through 10 is divided into four sections, one for each subject tested: Reading/English/Language Arts; Mathematics; Science; and Social Studies/Civics. Two copies of the Individual Student Report are sent to the district test coordinator so that a copy can be given to the student's parent or guardian.

Most useful for determining the instructional needs of individual students, the Individual Student Report provides information about specific Key Skills mastered and not mastered by the student for each subject. It reports student information, special notes, subject-scaled scores, cluster scores (High, Medium, or Low beginning in 1989), and Key Skills information.

The student's name (and identification number if one was provided on the answer sheet), the building and district names with corresponding code number, and the testing date are listed in the top section of this report. The test form administered and the student's grade are also identified.

Special notes pertaining to a student's performance follow the identifying information. For example, a student who was identified by district personnel as receiving special education services will be designated as an IEP (Individualized Education Program) student. An explanatory note will also be given for a student who attempted to answer less than 70 percent of the test items in a subject.

The subject and cluster scores achieved by a student are presented in the middle sections of this report. Clusters are groupings of Key Skills within a subject. A list of the Key Skills making up the clusters for each grade is presented in Appendix A of this manual.

The bottom sections of the Individual Student Report presents the Key Skills in each subject in two categories—those mastered and those not mastered by the student. Key Skills are listed by a code number and a brief descriptor. The code number is the same as that assigned to the Key Skill in *Core Competencies and Key Skills for Missouri Schools* and the *MMAT Test Content Specifications*. Each Key Skill is tested by four questions. Mastery of a Key Skill is demonstrated when a student answers three or four of these questions correctly. Following each Key Skill descriptor, in parentheses, is the number of questions a student answered correctly.

The information presented in the Key Skills sections of the Individual Student Report is extremely useful for improving instruction. Teachers should develop strategies for teaching the Key Skills which were not mastered. The *Core Competencies and Key Skills for Missouri Schools* and the *MMAT Test Content Specifications* are valuable resources for developing Key Skill instructional strategies.

Individual Student Score Label

Two copies of the self-adhesive Individual Student Score Label are provided for districts that wish to affix such a label to students' permanent records. Identifying information printed on the label includes the student's name (and identification number if one was provided on the answer sheet), the building code number, the student's grade level, and the testing date. The student's scaled score and an estimated comparable national percentile rank in each subject are also listed on the label.

The estimated comparable national percentile rank ranges from 1 to 99. This norm-referenced information is obtained by statistically comparing scores of a sample of students in the state sample on the *Iowa Tests of Basic Skills* (ITBS) (grades 2 through 8) and the *Tests of Achievement and Proficiency* (TAP) (grades 9 and 10) to their scores on the MMAT. The results of these comparisons are then used to estimate a national norm for every student taking the MMAT. Estimated comparable national percentile ranks are reported on the label for students in grade 2 in Reading, Language Arts, and Mathematics. Estimated comparable national percentile ranks are reported for students in grades 3 through 8 in Reading, Language Arts, Mathematics, Science, and Social Studies/Civics, and for students in grades 9 and 10 in Reading, Mathematics, Science, and Social Studies/Civics.

The estimated comparable national percentile rank is not equivalent to a norm-referenced test score, but it can be used for various purposes associated with Chapter 1 programs, such as identifying students eligible for Chapter 1 services.

Pupil List Report

The Pupil List Report shows in roster format each student's performance on each Key Skill in a subject for the grade tested. This report also presents the total number of Key Skills mastered by a student and the total number of students mastering each Key Skill.

Schools receive a separate Pupil List Report for each grade level subject in a building. A single page of this report can present data for 35 students. One copy of each report is sent to the district test coordinator. Because it is organized as a roster, the Pupil List Report provides quick access to an individual student's status on each Key Skill. The top section of the Pupil List Report identifies the school building, district, and corresponding code number. The testing date, the grade, the subject, and the test form administered are also listed.

Each subject's Key Skills are listed by code number across the top of the student roster section of the Pupil List Report. (A brief descriptor for each Key Skill code number is given in

Appendix A.) The total number of Key Skills tested in the subject is noted to the left of the code numbers. Student names are listed alphabetically down the left side of the report. Key Skill mastery is indicated by a plus sign (+); nonmastery is indicated by a minus sign (-).

The total number of Key Skills in the subject mastered by each student is noted at the right of the page. At the end of the Pupil List Report, a table presents the total number of students in the grade demonstrating mastery of each Key Skill. The number of students for whom data are given is also listed at the end of the report.

Before the name of each student identified by the district as receiving special education services is the notation "IEP." The notation "LEP" is used for those students demonstrating limited English proficiency. An asterisk (*) designates a student who attempted to answer less than 70 percent of the test items for the subject. No mastery/nonmastery data are presented for a student who did not attempt at least five items on the subject test.

For the convenience of those needing information to screen students for placement in programs for the academically talented, a double asterisk (**) designates a student whose subject-scaled score equalled or exceeded that corresponding to the 90th percentile for the state sample.

Grade Level Key Skill Report

The Grade Level Key Skill Report presents the percentage of students mastering each Key Skill, scaled score averages for the subject and its clusters, and a code number and brief descriptor for each Key Skill. A separate Grade Level Key Skill Report is provided for each subject for a grade in a building and in a district. Two copies of each report are sent to the district test coordinator.

The percentage and the number of students in the building or district by grade demonstrating mastery of a corresponding cumulative total of Key Skills is also reported. It is likely that all students in a building or district will master several Key Skills, and only a few students will master all the Key Skills. Thus, the number and percentage of students demonstrating mastery decreases as the cumulative total of Key Skills increases.

The Grade Level Key Skill Report provides information useful for determining the typical strengths and weaknesses of students in a grade within a building or a district. It also presents summary data about students' mastery of the Key Skills. This report can be a useful tool in evaluating a building's or district's curriculum and instruction, because it focuses on the achievement of all the students in a grade.

In the upper left section of this report, the building and/or the district are identified, and the corresponding code number is given. The grade, the test form administered, and the testing date are also listed. The number of students whose scores were used to compute the data shown on the report is indicated. The scores of IEP students and students who attempted less than 70 percent of the items on the subject test are not represented in the data. The number

and types of students whose scores are excluded are listed in the section designated "Note." Beginning in 1989, scores for Limited English Proficiency (LEP) students will not be included in the Grade Level Key Skill Report.

The upper right section of this report shows the subject-scaled score averages and cluster designations for the building and/or district and the state. The Key Skills making up each cluster in a subject for a grade level are listed in Appendix A. Subject scaled scores are computed independently of cluster scaled scores.

Using these data, the building's or district's cluster and subject-scaled score averages can be compared in order to determine relative strengths and weaknesses in the students' performance. This is done by comparing cluster averages to the overall subject average. In general, a scaled score difference of 25 points or more constitutes a meaningful difference between building or district subject and cluster averages.

Similarly, a building's strengths and weaknesses relative to the district and the state can be determined by examining the difference between respective grade-level averages. This is done by comparing the respective subject-scaled score averages and the respective cluster scaled score averages. In general, a scaled score difference of 25 points or more suggests a meaningful difference between respective averages.

Also, a district's cluster and subject-scaled score averages can be compared to determine relative strengths and weaknesses in students' performance, just as with the building's averages. As with the building, a scaled score difference of 25 points or more constitutes a meaningful difference between district subject and cluster averages.

A district's strengths and weaknesses relative to the state can be determined by examining the difference between respective grade level subject and cluster scaled score averages. In general, a difference of 25 scaled score points or more suggests a meaningful difference between corresponding district and state averages.

Grade Level Cluster Report

The Grade Level Cluster Report presents information about students' performance on each MMAT subject and its clusters. A separate Grade Level Cluster Report is provided for each grade in each building and in each district. Two copies of each report are sent to the district test coordinator. The Grade Level Cluster Report can be a useful tool in evaluating students' performance on the subject and clusters because it graphically depicts average scores as well as the degree to which students' scores vary from this average. Thus, the data presented can be used to compare visually the typical performance as well as the degree of variation in scores in all subjects and clusters at a grade within a building or a district.

The upper section of the Grade Level Cluster Report identifies the building and/or the district and gives the corresponding code number. The testing date, the grade, the test form administered, and any special notes pertaining to the report are also listed here.

The subject and cluster scores of individual students at a grade can be averaged for a building or a district. The Key Skills making up each cluster in each subject at each grade are listed in Appendix A. The subject and cluster scaled score averages in a grade for a building and/or a district, and the state, are reported in the right hand section of this report.

To the left of the averages, the performance of students in a specific grade within a building or a district is graphically depicted. The number of students used to compute the data is listed for each subject and its clusters. The scores of IEP students and students who attempted less than 70 percent of the items on a subject test are not represented in the data. Beginning in 1989, scores for LEP students will not be included in the Grade Level Cluster Report.

Summary Report

The Summary Report provides information about an entire building's or district's performance at grades 2 through 10 on each of the subjects tested by the MMAT: Reading/English/Language Arts; Mathematics; Science; and Social Studies/Civics. Three copies of the Summary Report for a district and two copies of the Summary Report for a building are sent to the district test coordinator.

This report offers information useful for evaluating the general effectiveness of the building's or district's curriculum and instruction in teaching the students to master the Key Skills. It also can be a valuable tool for monitoring changes over time in MMAT performance as well as for providing achievement data to improve instruction. The number and type of students whose scores are excluded are listed in the section designated "Notes."

Chapter 1 Eligibility List

The Chapter 1 Eligibility List identifies students eligible to receive Chapter 1 services, based on their estimated comparable national percentile rank(s). Students in grades 2 through 8 can be eligible in one, two, or three subject(s): Reading, Language Arts, and Mathematics. Students in grades 9 and 10 can be eligible in one or two subject(s): Reading and Mathematics. The Chapter 1 Eligibility List is an optional report and can be ordered for each grade in a building. One copy of each report is sent to the district test coordinator. Each page of the Chapter 1 Eligibility List presents data for a maximum of 35 students, so the report may be several pages long.

All students in a grade are ranked according to their performance on the Reading portion of the test. For each subject, the student's estimated comparable national percentile rank and the corresponding normal curve equivalent (NCE) are provided. "IEP" appears before the name of a student identified by the district as receiving special education services. "LEP" appears before the name of a student who has demonstrated limited English proficiency. An asterisk (*) follows the score of a student who attempted to answer less than 70 percent of the test items for a subject.

Conclusions

This technical manual has described the *Missouri Mastery and Achievement Tests* and its origin in the Excellence in Education Act of 1985. In addition to providing the user with information concerning the test, its appropriateness, and its technical adequacy, the manual demonstrates that every effort was made during the development of the MMAT to ensure that the battery would meet the *Standards for Educational and Psychological Testing* (AERA/ APA/NCME, 1985). At each step, the guiding principle was that the MMAT must yield reliable and valid measures of student achievement on the Key Skills for Missouri Schools.

Three types of evidence of validity were presented. A great deal of care was taken in chronicling development efforts that demonstrate content-related evidence of validity. Empirical studies were conducted to provide criterion-related evidence and construct-related evidence of validity. And finally, care was taken to avoid techniques of item construction and selection that would disadvantage students from differing socio-cultural backgrounds. Evidence points to the fact that scores can be used to generalize achievement of the Key Skills.

Reliability was addressed in several ways, employing traditional techniques and IRT methods. Standard errors were computed and reported with both methodologies. Evidence indicates that scores are relatively free of error. When they are interpreted within the standard error estimates, scores can yield reliable information upon which to base decisions.

Detailed descriptions of norming procedures and scaling techniques were presented. Finally, methods of establishing comparability were described. In these descriptions, evidence for the interpretability of test scores was presented. The several scores all refer to levels of achievement of academic content described in the Key Skills. The scores assist in interpreting level of performance.

Pertinent information can be generated from results of the MMAT for students, parents, teachers, school administrators, and curriculum specialists. The Missouri Department of Elementary and Secondary Education can use the data generated from test performance to report composite test information to the Missouri General Assembly on a consistent yearly basis.

References

- AERA/APA/NCME (1985). Standards for educational and psychological testing. Washington, D.C.: American Psychological Association.
- Mislevy, R. J., & Bock, R. D. (1984). BILOG: Item analysis and test scoring with binary logistic models. Mooresville, IN: Scientific Software, Inc.
- Huynh, H. (1980). Computation and inference for two reliability indices in mastery testing based on the beta-binomial model. In H. Huynh and J. C. Saunders (eds.) Solutions for some technical problems in domain-referenced mastery testing. Columbia, SC: Department of Educational Research and Psychology, College of Education, University of South Carolina.

Appendix A
Key Skill\Abbreviations and Clusters

Grade 2

READING/LANGUAGE ARTS

Cluster: *Language Arts*

- B-2 Synonyms/Antonyms
- B-3 Prefixes/Suffixes
- B-4 New Word Meaning
- D-2 Book Parts

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Elements
- C-3 Fact/Fantasy
- C-4 Cause-effect
- C-6 Main Idea
- C-7 Outcome Prediction
- D-3 Maps/Charts
- D-4 Directions

Cluster: *Writing*

- B-5 Antecedents/Pronouns
- G-4 Effective Writing
- G-5 Synonyms/Overused Words
- G-6 Sentence Combining
- G-7 Capitalization
- G-8 Punctuation
- G-9 Usage

Grade 3

READING/LANGUAGE ARTS

Cluster: *Language Arts*

- B-2 Prefixes/Suffixes
- B-3 New Word Meaning
- B-4 Synonyms/Antonyms
- D-2 Book Parts

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Detail
- C-2 Story Title
- C-3 Fact/Fantasy
- C-4 Cause-effect
- C-5 Character Comparison
- C-6 Main Idea
- C-7 Summarization
- C-8 Outcome Prediction
- C-9 Conclusions/Generalizations
- C-10 Problem/Solution
- D-3 Maps/Charts
- D-4 Directions

Cluster: *Writing*

- G-4 Effective Writing
- G-6 Word Choice
- G-7 Sentence Combining
- G-8 Capitalization
- G-9 Punctuation
- G-10 Grammatical Usage

Grade 4

READING/LANGUAGE ARTS

Cluster: *Language Arts*

- B-2 Prefixes/Suffixes
- B-3 Synonyms/Antonyms
- D-1 Learning Resources
- D-2 Book Parts

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Elements
- C-2 Author's Purpose
- C-3 Cause-effect
- C-4 Character Comparison
- C-5 Main Idea
- C-6 Summarization
- C-7 Outcome Prediction
- C-8 Conclusions/Generalizations
- C-9 Problem/Solution
- C-10 Figurative Language
- D-3 Maps/Charts/Tables
- D-6 Directions

Cluster: *Writing*

- G-2 Effective Writing
- G-3 Draft Revision
- G-4 Spelling
- G-5 Capitalization
- G-6 Punctuation
- G-7 Usage

Grade 5

READING/LANGUAGE ARTS

Cluster: *Language Arts*

- B-2 New Word Meaning
- B-3 Synonyms/Antonyms
- B-4 Prefixes/Suffixes
- C-11 Figurative Language
- D-1 Learning Resources

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Elements
- C-2 Author's Purpose
- C-3 Fact/Opinion
- C-4 Cause-effect
- C-5 Character Comparison
- C-6 Main Idea
- C-7 Summarization
- C-8 Outcome Prediction
- C-9 Conclusions/Generalizations
- D-2 Maps/Graphs/Time Lines
- D-6 Directions

Cluster: *Writing*

- G-3 Effective Writing
- G-5 Spelling
- G-6 Capitalization
- G-7 Punctuation
- G-8 Grammatical Usage

Grade 6

READING/LANGUAGE ARTS

Cluster: *Language Arts*

- B-2 New Word Meaning
- B-4 Synonyms/Antonyms
- C-11 Point of View
- C-12 Figurative Language
- D-1 Learning Resources
- D-5 Appropriate Sources

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Sequence
- C-2 Author's Purpose
- C-3 Fact/Opinion
- C-4 Cause-effect
- C-5 Character Comparison
- C-6 Main Idea
- C-7 Summarization
- C-8 Outcome Prediction
- C-9 Conclusions/Generalizations
- C-10 Story Elements
- D-2 Maps/Charts/Tables
- D-6 Directions

Cluster: *Writing*

- G-3 Effective Writing
- G-4 Draft Revision
- G-5 Spelling
- G-6 Capitalization
- G-7 Punctuation
- G-8 Grammatical Usage

Grade 7

ENGLISH/LANGUAGE ARTS

Cluster: *Language Arts*

- C-11 Euphemisms
- C-12 Figurative Language
- D-1 Learning Resources
- D-2 Pictures/Maps/Charts
- D-6 Comparison/Point of View

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Line
- C-2 Elements/Plot Explanation
- C-3 Author's Purpose
- C-4 Fact/Opinion
- C-5 Cause-effect
- C-7 Main Idea
- C-8 Summarization
- C-9 Inferences/Outcomes
- C-10 Conclusions/Generalizations
- D-2 Pictures/Maps/Charts
- D-5 Directions
- D-6 Comparison/Point of View
- E-1 Routes/Schedules/Timetables
- E-2 Instructions

Cluster: *Writing*

- G-2 Main Idea/Details
- G-3 Effective Writing
- G-4 Main Idea/Supporting Details
- G-6 Sentence Combining
- G-8 Spelling
- G-9 Capitalization
- G-10 Punctuation
- G-11 Grammatical Usage

Grade 8

ENGLISH/LANGUAGE ARTS

Cluster: *Language Arts*

- B-2 New Word Meaning
- C-11 Figurative Language
- D-1 Learning Resources
- D-3 Study Aids
- D-7 Appropriate Sources

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Elements
- C-2 Author's Purpose
- C-3 Fact/Opinion
- C-4 Cause-effect
- C-5 Character Change
- C-6 Main Idea
- C-7 Summarization
- C-8 Time/Place Relationships
- C-9 Inferences/Conclusions
- C-10 Propaganda Techniques
- D-2 Graphic Sources
- D-6 Directions
- E-1 Schedules/Timetables
- E-2 Instructions/Labels

Cluster: *Writing*

- G-2 Detail Organization
- G-3 Story Detail
- G-4 Organizational Elements
- G-5 Persuasive Rhetoric
- G-6 Sentence Combining
- G-7 Spelling
- G-8 Capitalization
- G-9 Punctuation
- G-10 Grammatical Usage

Grade 9

ENGLISH/LANGUAGE ARTS

Cluster: *Language Arts*

- B-2 New Word Meaning
- B-3 Prefixes/Suffixes
- C-4 Metaphorical Language
- D-1 Learning Resources
- D-3 Print/Nonprint Sources
- D-4 Information Organization

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Theme
- C-2 Characterization Methods
- C-3 Plot Structures
- C-6 Author's Purpose
- C-7 Character Inclusion
- C-8 Cultural/Social Context
- C-9 Cause-effect
- C-10 Main Idea
- C-11 Inferences/Conclusions
- D-2 Graphic Sources
- D-8 Comparison/Point of View
- D-9 Directions
- E-1 Business Correspondence
- E-2 Occupational Information

Cluster: *Writing*

- G-1 Prewriting Strategies
- G-2 Discourse Modes
- G-4 Sentence Revision
- G-6 Spelling/Usage
- G-7 Capitalization
- G-8 Punctuation
- G-9 Vocabulary Expansion
- G-10 Subject-Verb Agreement

Grade 10

ENGLISH/LANGUAGE ARTS

Cluster: *Language Arts*

- B-2 New Word Meaning
- B-3 Prefixes/Suffixes
- C-10 Literary Forms
- C-11 Figurative Language
- D-1 Learning Resources
- D-3 Specialized Sources
- D-4 Study Aids
- D-8 Appropriate Sources

Cluster: *Reading*

- B-1 Contextual Word Meaning
- C-1 Story Theme
- C-2 Author's Purpose
- C-4 Cause-effect
- C-5 Character Attributes
- C-6 Main Idea
- C-7 Plot Elements
- C-8 Outcome Prediction
- C-9 Implicit Assumptions
- C-13 Examples
- C-14 Inferences/Conclusions
- D-2 Graphic Sources
- D-9 Directions
- E-1 Business Information
- E-2 Occupational Information

Cluster: *Writing*

- G-2 Paragraph Construction
- G-3 Sentence Revision
- G-5 Spelling/Grammatical Usage
- G-6 Capitalization
- G-7 Punctuation
- G-9 Sentence Structure

Grade 2

MATHEMATICS

- A-2 Number Comparison
- A-3 Place Value
- B-2 Addition
- B-3 Subtraction
- B-5 Number Sentences
- D-1 Time
- F-2 Graphs
- G-4 Word Problems

Grade 3

MATHEMATICS

Cluster: *Understanding Numbers*

- A-2 Number Comparison
- A-3 Place Value
- A-5 Odd/Even Numbers
- A-9 Fractional Parts

Cluster: *Computation*

- B-2 Addition/Subtraction
- B-4 Number Sentences
- B-5 Money Calculations
- B-7 Multiplication
- B-8 Division

Cluster: *Measurement and Geometry*

- D-1 Time
- D-2 Temperature
- D-5 Linear Measurement
- E-1 Geometric Shapes

Cluster: *Interpretation and Application*

- F-1 Graphs
- F-2 Coordinates
- G-3 Word Problems
- H-2 Money Problems

Grade 4

MATHEMATICS

Cluster: *Understanding Numbers*

- A-2 Number Comparison
- A-3 Place Value
- A-5 Fractional Comparison
- A-6 Number Lines/Graphs

Cluster: *Computation*

- B-1 Addition/Subtraction
- B-3 Multiplication
- B-4 Fraction Computation
- B-5 Number Sentences

Cluster: *Measurement and Geometry*

- A-6 Number Lines/Graphs
- D-2 Linear Measurement
- E-1 Lines/Rays/Segments/Angles
- H-4 Time Problems

Cluster: *Interpretation and Application*

- B-5 Number Sentences
- F-2 Graphs
- H-3 Money Problems
- H-4 Time Problems

Grade 5**MATHEMATICS****Cluster: *Understanding Numbers***

- A-2 Place Value
- A-5 Fractional Comparison
- A-10 Number Lines/Graphs
- A-12 Fractional Parts

Cluster: *Computation*

- B-2 Multiplication
- B-3 Division
- B-6 Fraction Computation
- B-7 Number Sentences

Cluster: *Measurement and Geometry*

- D-4 Measurement Conversion
- D-6 Rectangle Area/Perimeter
- E-1 Angle Classification
- E-2 Line Classification

Cluster: *Interpretation and Application*

- F-1 Graphs
- F-3 Ratios
- D-4 Measurement Conversion
- D-6 Rectangle Area/Perimeter
- G-2 Multistep Word Problems

Grade 6**MATHEMATICS****Cluster: *Understanding Numbers***

- A-2 Place Value
- A-4 Fractional Comparison
- A-6 Decimal Comparison
- A-8 Number Lines/Graphs
- A-11 Mixed Numbers
- A-12 Arithmetic Sequences

Cluster: *Computation*

- B-1 Whole Number Computation
- B-2 Decimal Computation
- B-3 Fraction Computation
- B-6 Order of Operations
- B-7 Number Sentences
- B-8 Proportions
- C-2 Fraction Magnitude

Cluster: *Measurement and Geometry*

- D-1 Measurement Units
- D-2 Linear Measurement
- D-4 Measurement Conversion
- D-5 Rectangle Area/Perimeter
- D-6 Area/Perimeter Problems
- E-2 Line Classification

Cluster: *Interpretation and Application*

- F-1 Graphs
- F-3 Ratios
- F-4 Averages
- G-2 Multistep Word Problems
- G-3 Essential Data
- H-1 Money Problems
- H-2 Time Problems

Grade 7**MATHEMATICS****Cluster: *Understanding Numbers***

- A-1 Number Comparison
- A-2 Equivalent Numbers
- A-4 Prime/Composite
- A-6 Place Value
- A-7 Number Lines/Graphs

Cluster: *Computation*

- B-2 Fraction Computation
- B-4 Order of Operations
- B-6 Scientific Notation
- B-9 Percent Problems
- B-10 Number Sentences
- C-3 Estimation

Cluster: *Measurement and Geometry*

- D-1 Linear Measurement
- D-3 Triangle Perimeter/Area
- D-4 Circle Circumference/Area
- D-6 Appropriate Unit Selection
- E-2 Right Triangle Identification

Cluster: *Interpretation and Application*

- F-1 Probability Calculation
- F-2 Probability Interpretation
- G-2 Number Sentences
- G-3 Multistep Word Problems
- H-1 Perimeter/Area/Volume
- H-2 Percent Problems
- H-3 Ratio Problems

Grade 8

MATHEMATICS

Cluster: *Understanding Numbers*

- A-1 Number Comparison
- A-2 Equivalent Numbers
- A-3 Geometric Sequences
- A-4 Prime Factors

Cluster: *Computation*

- B-1 Integer Computation
- B-2 Fraction Computation
- B-3 Decimal Computation
- B-5 Order of Operations
- B-8 Multiples/Factors
- C-2 Estimation

Cluster: *Measurement and Geometry*

- D-1 Linear Measurement
- D-3 Triangle Perimeter/Area
- D-4 Circle Circumference/Area
- E-1 Angle Classification
- E-2 Triangle Classification
- E-3 Quadrilateral Classification
- E-4 Polygon Classification

Cluster: *Interpretation and Application*

- F-1 Circle Graphs
- F-2 Coordinates
- G-2 Flow Charts
- H-1 Pay Computation
- H-2 Interest Computation
- H-3 Cost Computation
- H-4 Area Problems
- H-5 Ratio Problems

Grade 9

MATHEMATICS

Cluster: *Computation*

- A-2 Inverse Properties
- B-1 Scientific Notation
- B-3 Rational Computation
- B-4 Expression Evaluation
- B-5 Linear Equations
- C-3 Estimation

Cluster: *Measurement and Geometry*

- D-1 Cylinder/Prism Volumes
- D-2 Irregularly Shaped Figures
- E-2 Angle Complement/Supplement
- E-4 Point/Graph/Equation
- E-6 Slope

Cluster: *Interpretation and Application*

- F-1 Mean/Median/Mode/Range
- F-2 Probability
- G-1 Verbal/Algebraic Translation
- G-3 Word Problems
- H-1 Area Problems
- H-2 Ratio Problems

Grade 10

MATHEMATICS

Cluster: *Computation*

- A-1 Number Properties
- B-1 Scientific Notation
- B-2 Algebraic Expressions
- B-3 Linear Equations
- B-4 Formula Manipulation
- C-5 Estimation

Cluster: *Measurement and Geometry*

- D-1 Cylinder/Prism Volumes
- D-2 Irregularly Shaped Figures
- D-3 Pythagorean Theorem
- E-1 Angle Properties
- E-2 Third Angle Calculation
- E-3 Linear Equation Graphs
- E-4 Slope
- E-6 Similar Figures

Cluster: *Interpretation and Application*

- F-1 Mean/Median/Mode/Range
- F-2 Probability
- F-3 Inference Limitations
- G-1 Logical Conclusions
- G-2 Conclusion Validity
- G-3 Verbal/Algebraic Translation
- H-2 Banking Problems
- H-3 Gross Pay Computation
- H-4 Net Pay Computation

Grade 3

SCIENCE

Cluster: Life Science

- A-2 Hazardous Situations
- A-3 Health/Nutrition
- B-3 Plant/Animal Development
- B-4 Seasonal Activities
- B-10 Visual Sequences
- D-1 Seedling Height
- E-3 Survival Characteristics
- F-1 Seed Growth Conditions
- F-2 Organism Habitats

Cluster: Earth Science

- A-2 Hazardous Situations
- A-6 Natural Sky Objects
- B-4 Seasonal Activities
- B-5 Illumination/Reflection
- B-10 Visual Sequences
- D-3 Temperature/Precipitation
- F-4 Landform Changes

Cluster: Physical Science

- A-2 Hazardous Situations
- B-10 Visual Sequences
- F-5 Vibration & Pitch
- F-6 Heat Effects
- F-7 Motion Prediction

Grade 4

SCIENCE

Cluster: Life Science

- A-4 Pollution & Plants/Animals
- A-12 Major Body Organs
- A-15 Common Missouri Animals
- B-1 Object Classification
- C-3 Time Determination
- D-2 Living/Nonliving Things

Cluster: Earth Science

- A-4 Pollution & Plants/Animals
- A-11 Weather Instruments
- B-1 Object Classification
- C-3 Time Determination
- C-4 Mass Measurement
- C-5 Temperature Measurement
- D-1 Properties of Soil

Cluster: Physical Science

- B-1 Object Classification
- C-1 Change of State Temperature
- C-3 Time Determination
- C-4 Mass Measurement
- C-5 Temperature Measurement
- F-1 Heat Conductors/Insulators
- F-2 Predict Magnetic Strength

Grade 5

SCIENCE

Cluster: *Life Science*

- A-7 Daily Allowances in Diet
- A-8 Common Missouri Plants
- B-4 Plant/Animal Cells
- C-3 Measuring Devices
- D-3 Substance Use/Abuse

Cluster: *Earth Science*

- A-1 Common Constellations
- A-3 Conditions for Lunar Eclipse
- A-4 Barometer Readings/Weather
- A-5 Major Cloud Types
- C-3 Measuring Devices
- D-8 Evidence from Fossils
- F-2 Erosion/Weathering

Cluster: *Physical Science*

- B-5 Changes in Matter
- B-7 Simple/Compound Machines
- C-3 Measuring Devices
- E-1 Particulate Nature of Matter
- G-1 Effects of Heat on Matter
- G-2 Surface Texture & Friction
- G-3 Changes in Pitch

Grade 6

SCIENCE

Cluster: *Life Science*

- A-2 Life Requirements
- A-7 Major Body Organs
- B-2 Plant/Animal Classification
- B-3 Nutrient Classification
- B-4 Food Chain Roles
- D-1 Photosynthesis
- D-2 Plant Parts
- D-13 Plant/Animal Cells
- D-14 Endangered Species
- E-1 Animal Types
- F-1 Growth Rate Prediction
- F-2 Food Chain Alteration

Cluster: *Earth Science*

- A-2 Life Requirements
- E-4 Fossils
- F-6 Precipitation
- G-2 Erosion/Weathering
- G-7 Evaporation Rate

Cluster: *Physical Science*

- B-6 Acids/Bases
- B-8 Conductors/Nonconductors
- B-9 Simple/Compound Machines
- B-10 Kinetic/Potential Energy
- B-11 Physical/Chemical Changes
- F-3 Series/Parallel Circuits
- G-4 Electromagnets
- G-7 Evaporation Rate

Grade 7

SCIENCE

Cluster: *Life Science*

- A-1 Vertebrate Characteristics
- B-1 Food Web Members
- B-2 Plants/Animals in Biomes
- C-3 Metric Measurement
- D-15 Asexual/Sexual Reproduction
- F-7 Substance Use/Abuse
- G-1 Healthy Body Variables
- I-4 Experimental Methods

Cluster: *Earth Science*

- A-3 Planet Identification
- B-6 Basic Cloud Types
- C-3 Metric Measurement
- C-4 Energy Production/Use
- D-5 Mineral Identification Tests
- D-14 Kinetic Energy in Water Cycle
- D-16 Causes of Seasons
- E-1 Crustal Plate Boundaries
- F-3 Weather Changes & Fronts
- I-4 Experimental Methods

Cluster: *Physical Science*

- B-5 States of Common Elements
- C-3 Metric Measurement
- D-4 Atomic Parts/Structure
- D-14 Kinetic Energy in Water Cycle
- G-4 Pitch Alteration
- I-4 Experimental Methods

Grade 8**SCIENCE****Cluster: Life Science**

- A-1 Animal Cells
- A-3 Plant Structure
- B-1 Plant/Animal Classification
- D-1 Cell Parts
- D-2 Human Body Systems
- E-2 Pollution Effects
- J-3 Overpopulation Effects

Cluster: Earth Science

- B-15 Renewable/Nonrenewable Energy
- C-1 Weather Data
- D-7 Heat Transfer
- E-2 Pollution Effects
- F-2 Weather Prediction
- J-3 Overpopulation Effects

Cluster: Physical Science

- A-6 Electrolytic Cells
- B-10 Wave Characteristics
- B-12 Kinetic/Potential Energy
- B-13 Atoms/Elements/Compounds
- B-15 Renewable/Nonrenewable Energy
- D-6 Static Charges
- D-7 Heat Transfer
- D-11 Periodic Table
- G-2 Pendulums

Grade 9**SCIENCE****Cluster: Life Science**

- B-2 Chemical Reactions
- B-5 Covalent/Ionic Bonds
- D-1 Technology & Environment
- D-6 Mass & Weight
- D-9 Factors of Fertile Topsoil
- G-1 Variable in an Experiment
- G-2 Symbols of Elements
- J-1 Erosion Prevention

Cluster: Earth Science

- B-2 Chemical Reactions
- D-1 Technology & Environment
- D-6 Mass & Weight
- D-9 Factors of Fertile Topsoil
- G-1 Variable in an Experiment
- G-2 Symbols of Elements
- J-1 Erosion Prevention

Cluster: Physical Science

- B-1 Kilowatt/Kilowatt-hour
- B-2 Chemical Reactions
- B-5 Covalent/Ionic Bonds
- B-6 Lenses & Images
- C-2 Calculate Density
- C-3 Velocity/Wavelength/Frequency
- D-1 Technology & Environment
- D-3 Volume & Pressure
- D-4 Temperature & Volume
- D-6 Mass & Weight
- E-2 Conservation of Energy
- E-4 Archimedes' Principle
- E-5 Direction of Resultant Force
- G-1 Variable in an Experiment
- G-2 Symbols of Elements

Cluster: Cross-Disciplinary

- B-1 Kilowatt/Kilowatt-hour
- D-3 Volume & Pressure
- G-1 Variable in an Experiment
- J-1 Erosion Prevention

Grade 10**SCIENCE****Cluster: Life Science**

- A-1 Mitosis/Meiosis
- B-1 Earth Region Classification
- B-2 Symbiosis
- C-7 Organism Growth Charts
- D-3 Element Cycles/Food Webs
- H-2 Vegetation Growth Factors

Cluster: Earth Science

- B-1 Earth Region Classification
- B-4 Precipitation
- C-5 Scientific Notation
- D-6 Wind/Atmospheric Pressure
- D-8 Weather Fronts
- F-5 Energy Needs
- J-1 Toxic Waste

Cluster: Physical Science

- B-10 Nuclear Fission/Fusion
- C-4 Kilowatt-hour Calculation
- D-14 Matter State Changes
- D-15 Boyle's Law
- E-4 Energy Transformations

Cluster: Cross-Disciplinary

- C-5 Scientific Notation
- F-2 Semipermeable Membranes
- F-5 Energy Needs
- G-1 Experimental Variables
- J-1 Toxic Waste
- J-4 Human/Marine Interaction

Grade 3

SOCIAL STUDIES/CIVICS

Cluster: *Geography*

- A-3 Maps: Features
- B-1 Results of Regional Changes
- C-2 Spatial Interaction
- D-1 Maps: Place Location
- D-2 Maps: Interpretation
- V-2 Inferences

Cluster: *Government*

- G-1 Democratic Decision Making
- H-2 Governmental Function
- J-1 Responsible Behavior
- K-1 Respect for Others
- V-2 Inferences

Cluster: *Economics*

- E-4 Past/Present Comparison
- M-2 Economic Choices
- N-1 Types of Economic Resources
- O-2 Sources of Household Income
- T-1 Meeting Human Needs
- V-2 Inferences

Cluster: *Civics*

- G-1 Democratic Decision Making
- H-2 Governmental Function
- J-1 Responsible Behavior
- K-1 Respect for Others
- M-2 Economic Choices
- T-1 Meeting Human Needs

Grade 4

SOCIAL STUDIES/CIVICS

Cluster: *Geography*

- A-2 Maps: Features
- A-3 Maps: Place Location
- A-4 Characteristics of Regions
- B-1 Results of Regional Changes
- D-2 Maps: Interpretation
- O-2 Results of Economic Change
- P-2 Variations Among Needs
- R-1 Regional Specialization
- V-1 Appropriate Sources
- V-2 Inferences

Cluster: *Government*

- E-1 Time Line Interpretations
- G-1 Democratic Decision Making
- H-1 Make/Enforce/Interpret Rules
- L-1 Fact/Opinion
- V-1 Appropriate Sources
- V-2 Inferences

Cluster: *Economics*

- M-2 Trade-Offs in Decisions
- N-1 Types of Economic Resources
- O-2 Results of Economic Changes
- P-2 Variations Among Needs
- R-1 Regional Specialization
- V-2 Inferences

Cluster: *Civics*

- G-1 Democratic Decision Making
- H-1 Make/Enforce/Interpret Rules
- L-1 Fact /Opinion
- M-2 Trade-Offs in Decisions
- V-1 Appropriate Sources
- V-2 Inferences

Grade 5

SOCIAL STUDIES/CIVICS

Cluster: *Geography*

- A-2 Maps: Features
- A-3 Maps: United States
- A-4 Maps: World
- B-1 Results of Regional Changes
- D-2 Maps: Interpretation
- R-1 Regional Specialization
- V-1 Appropriate Sources

Cluster: *Government*

- G-3 Democratic Practices/Values
- H-1 Governmental Functions
- I-2 Rights/Responsibilities
- K-1 Political Issue Analysis
- L-1 Fact/Opinion

Cluster: *Economics*

- M-2 Trade-offs in Decisions
- N-1 Production Terms
- P-1 Price & Supply/Demand
- R-1 Regional Specialization
- V-3 Inferences

Cluster: *History*

- E-2 Historical Cause-effect
- E-3 Lifestyles in History
- E-4 Impact of Innovation
- F-1 Comparison of Viewpoints
- V-1 Appropriate Sources
- V-3 Inferences

Cluster: *Civics*

- G-3 Democratic Practices/Values
- H-1 Governmental Functions
- I-2 Rights/Responsibilities
- K-1 Political Issue Analysis
- L-1 Fact/Opinion
- M-2 Trade-offs in Decisions
- V-1 Appropriate Sources
- V-3 Inferences

Grade 6

SOCIAL STUDIES/CIVICS

Cluster: *Geography*

- A-2 Maps: Features
- A-3 Maps: United States
- A-4 Maps: World
- B-1 Results of Regional Changes
- D-3 Maps: Interpretation
- R-1 Regional Specialization
- V-1 Appropriate Sources

Cluster: *Government*

- G-1 Democratic Practices/Values
- H-1 Governmental Functions
- I-1 Rights/Responsibilities
- L-1 Fact/Opinion

Cluster: *Economics*

- M-1 Trade-offs in Decisions
- N-1 Production Terms
- O-2 Government & Economy
- P-1 Price & Supply/Demand
- R-1 Regional Specialization

Cluster: *History*

- E-2 Historical Cause-effect
- E-3 Past/Present Comparison
- E-4 Impact of Innovation
- F-1 Comparison of Viewpoints
- V-1 Appropriate Sources

Cluster: *Civics*

- G-1 Democratic Practices/Values
- H-1 Governmental Functions
- I-1 Rights/Responsibilities
- L-1 Fact/Opinion
- M-1 Trade-offs in Decisions
- O-2 Government & Economy
- S-3 Social Roles
- T-1 Meeting Human Needs

Grade 7

SOCIAL STUDIES/CIVICS

Cluster: *Geography*

- A-1 Maps: United States
- A-2 Maps: World
- B-1 Results of Regional Changes
- D-3 Maps: Interpretation
- R-1 Regional Specialization
- V-1 Methods to Research Topics
- V-2 Appropriate Sources
- V-3 Inferences

Cluster: *Government*

- G-1 Democratic Practices/Values
- H-1 Governmental Functions
- I-1 Rights/Responsibilities
- J-2 Citizen Influence
- L-1 Fact/Opinion

Cluster: *Economics*

- M-1 Economic Decision Making
- N-1 Production Terms
- P-1 Changes in Supply/Demand
- R-1 Regional Specialization
- V-3 Inferences

Cluster: *History*

- E-1 Time Line Interpretations
- E-2 Cause-effect Relationships
- V-1 Methods to Research Topics
- V-2 Appropriate Sources
- V-3 Inferences

Cluster: *Civics*

- G-1 Democratic Practices/Values
- H-1 Governmental Functions
- I-1 Rights/Responsibilities
- J-2 Citizen Influence
- L-1 Fact/Opinion
- M-1 Economic Decision Making
- V-1 Methods to Research Topics
- V-2 Appropriate Sources
- V-3 Inferences

Grade 8

SOCIAL STUDIES/CIVICS

Cluster: *Geography*

- A-1 Maps: United States
- A-2 Maps: World
- B-1 Results of Regional Changes
- D-3 Maps: Interpretation
- N-1 Results of Economic Changes
- Q-1 Economic Growth/Decline

Cluster: *Government*

- G-1 Democratic Practices
- G-3 Democratic Values
- H-1 Legislative Procedures
- H-3 Governmental Offices
- I-1 Constitutional Rights
- J-2 Citizen Influence
- K-1 Political Issue Analysis

Cluster: *Economics*

- M-2 Economic Decision Making
- N-1 Results of Economic Changes
- P-1 Changes in Supply/Demand
- Q-1 Economic Growth/Decline

Cluster: *History*

- B-1 Results of Regional Changes
- E-1 Early U.S. History
- E-2 19th Century U.S. History
- V-1 Research Methods/Sources

Cluster: *Civics*

- G-1 Democratic Practices
- G-3 Democratic Values
- H-1 Legislative Procedures
- H-3 Governmental Offices
- I-1 Constitutional Rights
- J-2 Citizen Influence
- K-1 Political Issue Analysis
- M-2 Economic Decision Making
- N-1 Results of Economic Changes
- Q-1 Economic Growth/Decline
- V-1 Research Methods/Sources

Grade 9

SOCIAL STUDIES/CIVICS

Cluster: *Geography*

- A-1 Maps: United States
- A-2 Maps: World
- B-1 Results of Regional Changes
- D-1 Maps: Place Location
- D-2 Map Selection for Purposes
- V-1 Appropriate Sources
- V-4 Inferences

Cluster: *Government*

- G-1 Federal/State Governments
- H-2 Legislative Process
- H-3 Separation of Powers
- I-1 Constitutional Rights
- I-3 Rights vs. Needs of Others
- J-3 Citizen Influence
- K-1 Political Issue Analysis

Cluster: *Economics*

- M-2 Economic Decision Making
- O-1 Government & Economy
- O-2 Market & Command Economies
- Q-2 GNP & Living Conditions
- V-4 Inferences

Cluster: *History*

- B-1 Results of Regional Changes
- E-3 Cause-effect Relationships
- V-1 Appropriate Sources
- V-2 Analysis of History Sources

Cluster: *Civics*

- G-1 Federal/State Governments
- H-2 Legislative Process
- H-3 Separation of Powers
- I-1 Constitutional Rights
- I-3 Rights vs. Needs of Others
- J-3 Citizen Influence
- K-1 Political Issue Analysis
- M-2 Economic Decision Making
- O-1 Government & Economy
- O-2 Market & Command Economies
- Q-2 GNP & Living Conditions
- V-1 Appropriate Sources
- V-4 Inferences

Grade 10

SOCIAL STUDIES/CIVICS

Cluster: *Geography*

- A-1 Maps: United States
- A-2 Maps: World
- B-1 Results of Regional Changes
- C-1 Spatial Interactions
- D-3 Maps: Conclusions
- Q-1 GNP & Living Conditions
- Q-2 Influences on GNP

Cluster: *Government*

- G-1 Government Comparison
- G-2 Democratic Principles
- H-2 Legislative Procedures
- H-3 Government Officials
- I-1 Constitutional Rights
- J-3 Citizen Influence
- K-1 Political Issue Analysis
- K-2 Constitutional Principles
- L-3 Propaganda Techniques
- V-1 Research Methods/Sources

Cluster: *Economics*

- M-3 Economic Decision Making
- O-1 Market & Command Economies
- O-5 Government & Economy
- P-4 Price & Supply/Demand
- Q-1 GNP & Living Conditions
- Q-2 Influences on GNP
- V-1 Research Methods/Sources

Cluster: *History*

- E-2 Historical Causes
- E-3 Historical Consequences
- E-4 Historical Sequences
- K-2 Constitutional Principles
- V-1 Research Methods/Sources

Cluster: *Civics*

- G-1 Government Comparison
- G-2 Democratic Principles
- H-2 Legislative Procedures
- H-3 Government Officials
- I-1 Constitutional Rights
- J-3 Citizen Influence
- K-1 Political Issue Analysis
- K-2 Constitutional Principles
- L-3 Propaganda Techniques
- M-3 Economic Decision Making
- O-1 Market & Command Economies
- O-5 Government & Economy
- P-4 Price & Supply/Demand
- Q-1 GNP & Living Conditions
- Q-2 Influences on GNP
- S-5 Prejudice/Discrimination
- V-1 Research Methods/Sources

Appendix B
Samples of Test Content Specifications

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Reading

Level II

Code C-3

Distinguish fact from fantasy and justify response.

Given a brief passage, the student will determine whether it is fact or fantasy and justify that response. Excerpts may come from poems, stories, or factual texts.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Writing

Level II

Code G-6

Combine sentences with simple conjunctions.

Given two sentences that can be combined with a coordinating conjunction, the student will identify the combination which most appropriately retains the meaning of the original sentences.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Math

Level III

Code B-8

Divide using one-digit divisors (through the tables of 5) that result in one-digit quotients with no remainders.

Given a division problem with a two-digit dividend and a one-digit divisor greater than or equal to two and less than six, the student will solve the problem. The solution will consist of a one-digit quotient with no remainder.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Science

Level III

Code F-6

Predict the changes in matter as an object is heated or cooled.

The student will identify what happens to an object when it is heated or cooled. For example, the student will recognize that ice will melt on a sidewalk during a warm day, that molten candy will solidify as its temperature drops, and that air will decrease in volume as it cools.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Social Studies

Level III

Code H-2

Identify major offices, activities, and locations of three levels of government: local, state, and national.

The student will make correct associations regarding the locations and duties of major government offices on the national, state, and local levels. For example, the student will know that Washington, D.C. is where the President and Congress work, and that the local government is responsible for providing schools and libraries.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Reading

Level V

Code C-9

Draw conclusions and make generalizations from material read.

Given a passage, the student will arrive at valid conclusions and logical generalizations based on the information presented. An understanding of elements such as character, setting, purpose, and theme may be necessary.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Writing

Level V

Code G-3

Produce various forms of writing: paragraphs with topic sentences and supporting details, journals, stories, poems, autobiographical and biographical pieces, how-to activities, social notes, informal and business letters, and simple reports.

Given one or more excerpts from various forms of writing, the student will identify the particular forms. The student will also identify various parts of a written passage, such as the topic sentence, supporting details, and conclusion of a paragraph. Material may include journals, stories, poems, autobiographical and biographical pieces, how-to activities, social notes, personal and business letters, and simple reports.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Math

Level IV

Code A-2

Compare numbers through five digits using $<$, $>$, and $=$.

Given any two positive whole numbers of five digits or less, the student will express their relationship by using $<$, $>$, or $=$.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Science

Level IV

Code A-15

Identify examples of common Missouri animals.

Given an illustration of a common Missouri animal, wild or domesticated, the student will identify its common name. Examples of common Missouri animals include deer, raccoon, opossum, squirrel, rabbit, mouse, rat, lizard, eagle, hawk, owl, frog, turkey, cardinal, duck, goose, pig, horse, cow, goat, chicken, sheep, butterfly, mosquito, wood tick, beetle, earthworm, spider, crayfish, catfish, trout, snake, and moth.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Social Studies

Level V

Code V-3

Draw inferences from various sources: simple tables and graphs, observations, photographs, works of art, stories, interviews, and primary sources (artifacts, records, and documents from the past).

Given a graph, table, illustration, or an excerpt from a story or primary source, the student will draw inferences from the information presented.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Reading

Level VII

Code C-4

Distinguish between fact and opinion and justify response.

Given statements of fact and opinion, the student will distinguish between fact and opinion and justify the response.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Writing

Level VII

Code G-6

Use sentence combining to show relationship between two kernel sentences (compounding, subordination, apposition, etc.).

Given two or more simple sentences, the student will identify the best combination into a single sentence. Methods of combination will include subordination, coordination, and apposition.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Math

Level VIII

Code H-3

Compute the actual cost of an item purchased on an installment plan.

Given a word problem that provides information about an item purchased on the installment plan, including such details as the cost of the item, tax, carrying charge, down payment, monthly payment, and time span, the student will determine the actual cost of the item.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Science

Level IX

Code D-9

Communicate ways in which plant decay and animal activity make fertile topsoil (holding water, aerating soil, enriching soil, rock decomposition, soil nutrients, etc.).

Given illustrations or descriptions of plants or animals living in or on the soil, the student will identify how these organisms affect the composition of the soil. Examples of such interaction include the processes by which plants help prevent erosion, decaying animals enhance the fertility of the soil, and micro-organisms break down organic components.

MISSOURI MASTERY AND ACHIEVEMENT TESTS

TEST CONTENT SPECIFICATIONS

Subject Social Studies

Level X

Code V-1

Identify and evaluate methods and resources for investigating given topics related to other cultures, history, current politics, or economics.

The student will demonstrate a knowledge of various sources and methods for investigating topics that may be studied in secondary social studies programs. The student will identify appropriate primary and secondary sources for a given topic, evaluating them according to timeliness, authority, and relevance. The student will also demonstrate a knowledge of research tools like the Reader's Guide to Periodical Literature, almanacs, and atlases. For example, the student will indicate that interviewing a Japanese exchange student would be an appropriate method for investigating differences between American and Japanese high schools and that articles listed in the latest Reader's Guide on that topic would be more useful than an entry on Japan in an encyclopedia published in 1953.

Appendix C
Raw Score and Scaled Score Statistics for Subject Areas

Grade 2, Form D Statistics

Spring 1988 Administration

Subject	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	76	55.82	14.36	.946	300	65	.942
Mathematics	32	25.38	5.48	.867	300	65	.851

Grade 3, Form A Statistics

Spring 1987 Administration

Subject	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	84	64.46	14.58	.946	300	65	.948
Mathematics	68	53.69	10.79	.922	300	65	.929
Science	64	45.05	9.35	.889	300	65	.909
Social Studies	56	37.96	9.87	.909	300	65	.924

Grade 4, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	88	59.77	15.64	.940	300	65	.950
Mathematics	52	39.20	8.28	.887	300	65	.901
Science	52	32.90	8.16	.857	300	65	.877
Social Studies	64	40.32	11.42	.914	300	65	.927

Grade 5, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	84	61.98	12.70	.921	300	65	.929
Mathematics	60	38.83	10.83	.912	300	65	.921
Science	68	39.38	9.75	.858	300	65	.873
Social Studies	80	50.50	13.72	.920	300	65	.932

Grade 6, Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	92	62.38	14.50	.928	300	65	.937
Mathematics	104	67.82	17.61	.946	300	65	.955
Science	92	51.32	12.92	.891	300	65	.914
Social Studies	84	59.34	14.21	.935	300	65	.943

Grade 7, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	100	70.57	16.64	.941	300	65	.946
Mathematics	92	53.47	17.95	.947	300	65	.956
Science	76	43.92	10.97	.877	300	65	.895
Social Studies	72	44.82	12.91	.921	300	65	.931

Grade 8, Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	108	73.02	17.87	.942	300	65	.951
Mathematics	100	59.34	17.33	.939	300	65	.949
Science	72	40.09	10.22	.862	300	65	.901
Social Studies	72	50.96	12.84	.932	300	65	.941

Grade 9, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	112	78.45	18.09	.945	300	65	.951
Mathematics	68	39.10	12.82	.922	300	65	.937
Science	68	33.88	8.46	.810	300	65	.861
Social Studies	80	52.91	13.45	.923	300	65	.935

Grade 10, Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	112	75.72	17.71	.940	300	65	.948
Mathematics	92	53.85	15.98	.933	300	65	.951
Science	80	41.69	10.27	.846	300	65	.891
Social Studies	100	64.89	18.70	.950	300	65	.957

Grade 3, Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	84	66.25	13.25	.937	315	68	.941
Mathematics	68	56.27	9.51	.913	318	77	.916
Science	64	47.21	8.80	.883	311	72	.895
Social Studies	56	40.07	8.99	.896	325	73	.905

Grade 6, Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	92	64.63	13.87	0.925	308	72	.937
Mathematics	104	71.86	17.89	0.949	332	77	.956
Science	92	53.79	13.91	0.908	332	84	.926
Social Studies	84	61.49	13.32	0.931	321	77	.942

Grade 8, Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	108	76.33	17.42	.941	322	70	.950
Mathematics	100	64.87	17.60	.945	336	84	.953
Science	72	44.53	10.43	.874	337	88	.906
Social Studies	72	52.54	12.47	.929	325	77	.937

Grade 10, Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Raw Score			Scaled Score		
		Mean	Standard Deviation	Reliability	Mean	Standard Deviation	Reliability
Reading	112	77.35	16.63	.933	315	72	.940
Mathematics	92	55.19	16.51	.938	313	72	.952
Science	80	42.15	11.09	.870	326	82	.892
Social Studies	100	67.65	17.85	.947	314	70	.952

Appendix D
Distribution of Scaled Score Statistics

Grade 2, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	76	300	65	.942	13 - 33
Language Arts	16	300	65	.768	
Reading	32	300	65	.869	
Writing	28	300	65	.851	
Mathematics	32	300	65	.851	19 - 39

Grade 3, Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	84	300	65	.948	10 - 30
Language Arts	16	300	65	.806	
Reading	52	300	65	.930	
Writing	24	300	65	.821	
Mathematics	68	300	65	.929	12 - 31
Numbers	16	300	65	.811	
Computation	20	300	65	.822	
Measurement/Geometry	16	300	65	.783	
Interpretation	16	300	65	.811	
Science	64	300	65	.909	13 - 35
Life Science	36	300	65	.846	
Earth Science	28	300	65	.785	
Physical Science	20	300	65	.770	
Social Studies	56	300	65	.924	13 - 50
Geography	24	300	65	.853	
Government	20	300	65	.831	
Economics	24	300	65	.833	
Civics	24	300	65	.878	

Grade 3, Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	84	315	68	.941	9 - 23
Language Arts	16	336	109	.762	
Reading	52	323	76	.916	
Writing	24	315	91	.823	
Mathematics	68	318	77	.916	14 - 22
Numbers	16	333	114	.791	
Computation	20	350	130	.769	
Measurement/Geometry	16	391	180	.785	
Interpretation	16	324	103	.776	
Science	64	311	72	.895	16 - 26
Life Science	36	319	79	.808	
Earth Science	28	313	88	.782	
Physical Science	20	313	87	.724	
Social Studies	56	325	73	.905	15 - 32
Geography	24	329	88	.813	
Government	20	341	107	.807	
Economics	24	321	75	.770	
Civics	24	334	86	.820	

Grade 4, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	88	300	65	.950	12 - 37
Language Arts	16	300	65	.694	
Reading	52	300	65	.932	
Writing	24	300	65	.838	
Mathematics	52	300	65	.901	17 - 34
Numbers	16	300	65	.744	
Computation	16	300	65	.807	
Measurement/Geometry	16	300	65	.688	
Interpretation	16	300	65	.770	
Science	52	300	65	.877	21 - 42
Life Science	24	300	65	.729	
Earth Science	28	300	65	.818	
Physical Science	28	300	65	.841	
Social Studies	64	300	65	.927	15 - 46
Geography	40	300	65	.875	
Government	24	300	65	.831	
Economics	24	300	65	.825	
Civics	24	300	65	.839	

Grade 5, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	84	300	65	.929	14 - 28
Language Arts	20	300	65	.774	
Reading	48	300	65	.896	
Writing	20	300	65	.727	
Mathematics	60	300	65	.921	16 - 40
Numbers	16	300	65	.855	
Computation	16	300	65	.774	
Measurement/Geometry	16	300	65	.803	
Interpretation	20	300	65	.798	
Science	68	300	65	.873	23 - 39
Life Science	20	300	65	.689	
Earth Science	28	300	65	.697	
Physical Science	28	300	65	.777	
Social Studies	80	300	65	.932	15 - 37
Geography	28	300	65	.830	
History	24	300	65	.805	
Government	20	300	65	.786	
Economics	20	300	65	.776	
Civics	32	300	65	.842	

Grade 6, Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	92	300	65	.937	5 13 - 31
Language Arts	24	300	65	.815	5
Reading	52	300	65	.903	5
Writing	24	300	65	.806	5
Mathematics	104	300	65	.955	5 11 - 31
Numbers	24	300	65	.834	5
Computation	28	300	65	.857	5
Measurement/Geometry	24	300	65	.814	5
Interpretation	28	300	65	.878	5
Science	92	300	65	.914	5 17 - 35
Life Science	48	300	65	.860	
Earth Science	20	300	65	.726	
Physical Science	32	300	65	.740	5
Social Studies	84	300	65	.943	12 - 30
Geography	28	300	65	.831	
History	20	300	65	.824	
Government	16	300	65	.759	
Economics	20	300	65	.814	
Civics	32	300	65	.861	

Grade 6, Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	92	308	72	.937	13 - 32
Language Arts	24	306	86	.821	
Reading	52	315	76	.902	
Writing	24	307	80	.809	
Mathematics	104	332	77	.956	13 - 30
Numbers	24	318	107	.839	
Computation	28	368	96	.873	
Measurement/Geometry	24	335	75	.814	
Interpretation	28	315	88	.881	
Science	92	332	84	.926	19 - 40
Life Science	48	328	82	.874	
Earth Science	20	344	96	.728	
Physical Science	32	342	95	.809	
Social Studies	84	321	77	.942	13 - 26
Geography	28	333	91	.828	
History	20	315	90	.822	
Government	16	333	120	.793	
Economics	20	331	93	.794	
Civics	32	329	90	.855	

Grade 7, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	100	300	65	.946	12 - 29
Language Arts	20	300	65	.793	
Reading	60	300	65	.920	
Writing	32	300	65	.837	
Mathematics	92	300	65	.956	12 - 42
Numbers	20	300	65	.854	
Computation	24	300	65	.842	
Measurement/Geometry	20	300	65	.839	
Interpretation	28	300	65	.864	
Science	76	300	65	.895	33 - 74
Life Science	32	300	65	.804	
Earth Science	40	300	65	.805	
Physical Science	24	300	65	.779	
Social Studies	72	300	65	.931	15 - 42
Geography	32	300	65	.859	
History	20	300	65	.814	
Government	20	300	65	.766	
Economics	20	300	65	.804	
Civics	36	300	65	.861	

Grade 8, Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	108	300	65	.951	12 - 28
Language Arts	20	300	65	.782	
Reading	60	300	65	.920	
Writing	36	300	65	.867	
Mathematics	100	300	65	.949	12 - 38
Numbers	16	300	65	.797	
Computation	24	300	65	.840	
Measurement/Geometry	28	300	65	.830	
Interpretation	32	300	65	.859	
Science	72	300	65	.901	18 - 34
Life Science	28	300	65	.814	
Earth Science	24	300	65	.752	
Physical Science	36	300	65	.795	
Social Studies	72	300	65	.941	13 - 37
Geography	24	300	65	.853	
History	16	300	65	.789	
Government	28	300	65	.854	
Economics	16	300	65	.814	
Civics	44	300	65	.899	

Grade 8, Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	108	322	70	.950	13 - 29
Language Arts	20	324	81	.780	
Reading	60	325	80	.916	
Writing	36	331	79	.869	
Mathematics	100	336	84	.953	14 - 39
Numbers	16	350	114	.786	
Computation	24	326	84	.823	
Measurement/Geometry	28	340	88	.847	
Interpretation	32	350	104	.886	
Science	72	337	88	.906	23 - 45
Life Science	28	340	100	.810	
Earth Science	24	339	100	.769	
Physical Science	36	336	89	.822	
Social Studies	72	325	77	.937	14 - 37
Geography	24	325	81	.816	
History	16	330	86	.771	
Government	28	328	96	.842	
Economics	16	331	104	.806	
Civics	44	327	90	.894	

Grade 9, Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	112	300	65	.951	12 - 24
Language Arts	24	300	65	.828	
Reading	60	300	65	.917	
Writing	32	300	65	.846	
Mathematics	68	300	65	.937	14 - 50
Computation	24	300	65	.856	
Measurement/Geometry	20	300	65	.829	
Interpretation	24	300	65	.834	
Science	68	300	65	.861	24 - 49
Life Science	32	300	65	.798	
Earth Science	28	300	65	.788	
Physical Science	60	300	65	.833	
Cross-Disciplinary	16	300	65	.556	
Social Studies	80	300	65	.935	14 - 34
Geography	28	300	65	.843	
History	16	300	65	.747	
Government	28	300	65	.846	
Economics	20	300	65	.744	
Civics	52	300	65	.901	

Grade 10, Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	112	300	65	.948	13 - 28
Language Arts	32	300	65	.853	
Reading	60	300	65	.910	
Writing	24	300	65	.800	
Mathematics	92	300	65	.951	11 - 39
Computation	24	300	65	.859	
Measurement/Geometry	32	300	65	.886	
Interpretation	36	300	65	.866	
Science	80	300	65	.891	20 - 42
Life Science	24	300	65	.703	
Earth Science	28	300	65	.703	
Physical Science	20	300	65	.668	
Cross-Disciplinary	24	300	65	.769	
Social Studies	100	300	65	.957	11 - 37
Geography	28	300	65	.846	
History	20	300	65	.813	
Government	40	300	65	.903	
Economics	28	300	65	.867	
Civics	68	300	65	.944	

Grade 10, Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	Scaled Score			
		Mean	Standard Deviation	Reliability	Range of Standard Errors of Measurement
Reading	112	315	72	.940	13 - 27
Language Arts	32	319	84	.839	
Reading	60	324	77	.882	
Writing	24	316	83	.804	
Mathematics	92	313	72	.952	13 - 39
Computation	24	308	91	.862	
Measurement/Geometry	32	314	84	.875	
Interpretation	36	310	81	.865	
Science	80	326	82	.892	24 - 48
Life Science	24	358	103	.716	
Earth Science	28	310	91	.711	
Physical Science	20	343	102	.689	
Cross-Disciplinary	24	297	83	.744	
Social Studies	100	314	70	.952	12 - 31
Geography	28	325	83	.838	
History	20	321	88	.792	
Government	40	326	81	.899	
Economics	28	314	85	.864	
Civics	68	313	72	.940	

Appendix E
Distribution of p-Values, Biserials, and Thresholds

Grade 2 Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	76	.39	.96	.73	.15	.89	.57	-3.12	.65	-1.18
Language Arts	16	.39	.91	.66	.15	.69	.49	-2.58	.63	-1.14
Reading	32	.53	.96	.78	.37	.90	.59	-3.23	-.19	-1.38
Writing	28	.54	.92	.72	.41	.72	.54	-2.01	-.20	-1.11
Mathematics	32	.58	.97	.80	.35	.68	.55	-3.58	-.29	-1.64

Grade 3 Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	p values			Biserials			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	84	.42	.96	.77	.23	.98	.61	- 3.39	- .12	- 1.26
Language Arts	16	.53	.94	.77	.43	.74	.60	- 2.36	- .13	- 1.21
Reading	52	.42	.96	.78	.24	.99	.65	- 3.33	- .52	- 1.30
Writing	24	.47	.95	.71	.18	.74	.50	- 2.43	- .19	- 1.10
Mathematics	68	.51	.97	.78	.27	.76	.55	- 4.22	- .04	- 1.51
Numbers	16	.63	.89	.77	.38	.75	.57	- 2.27	- .48	- 1.27
Computation	20	.60	.97	.81	.29	.83	.56	- 4.53	- .36	- 1.70
Measurement/Geometry	16	.61	.94	.80	.28	.71	.50	- 2.87	- .40	- 1.82
Interpretation	16	.51	.92	.75	.35	.78	.60	- 1.99	- .60	- 1.16
Science	64	.17	.94	.70	.06	.85	.49	- 2.24	4.87	- .88
Life Science	36	.22	.94	.72	.02	.84	.50	- 2.29	4.20	- 1.03
Earth Science	28	.17	.94	.65	.03	.68	.40	- 2.27	5.24	- .62
Physical Science	20	.44	.94	.75	.09	.80	.50	- 2.28	.39	- 1.27
Social Studies	56	.24	.91	.69	.14	.87	.54	- 1.92	2.16	- .64
Geography	24	.43	.85	.67	.28	.75	.53	- 1.45	.50	- .81
Government	20	.40	.86	.71	.23	.81	.56	- 1.80	.85	- .92
Economics	24	.24	.91	.68	.13	.87	.52	- 2.02	2.53	- .61
Civics	24	.40	.91	.74	.21	.89	.64	- 1.83	.94	- .88

Grade 3 Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	84	.54	.98	.79	.18	.97	.60	- 3.32	- .28	- 1.48
Language Arts	16	.58	.94	.81	.41	.73	.56	- 2.77	- .33	- 1.60
Reading	52	.54	.98	.81	.19	.97	.62	- 3.36	- .26	- 1.54
Writing	24	.57	.90	.74	.32	.88	.53	- 2.16	- .42	- 1.21
Mathematics	68	.54	.97	.83	.35	.81	.56	- 3.93	- .23	- 1.80
Numbers	16	.70	.91	.80	.37	.72	.54	- 2.52	- .76	- 1.51
Computation	20	.63	.97	.86	.37	.76	.56	- 3.79	- .66	- 2.00
Measurement/Geometry	16	.74	.96	.86	.37	.78	.55	- 3.13	- 1.03	- 1.98
Interpretation	16	.55	.96	.79	.31	.74	.56	- 2.59	- .23	- 1.42
Science	64	.20	.97	.74	.17	.79	.49	- 6.15	2.23	- 1.38
Life Science	36	.20	.97	.77	.15	.79	.48	- 5.80	2.47	- 1.60
Earth Science	28	.20	.97	.69	.16	.58	.41	- 6.40	2.21	- 1.20
Physical Science	20	.36	.97	.78	.16	.69	.47	- 5.63	1.50	- 1.72
Social Studies	56	.25	.98	.72	.09	.90	.51	- 4.82	1.91	- 1.06
Geography	24	.39	.91	.73	.09	.74	.50	- 3.11	1.71	- 1.13
Government	20	.45	.95	.77	.27	.81	.55	- 2.68	.39	- 1.28
Economics	24	.25	.98	.72	.09	.81	.43	- 4.99	2.70	- 1.16
Civics	24	.45	.98	.78	.26	.91	.58	- 4.49	.40	- 1.45

Grade 4 Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	88	.32	.93	.68	.15	.79	.53	- 2.28	2.17	- .82
Language Arts	16	.40	.90	.68	.23	.50	.40	- 2.18	.56	- 1.02
Reading	52	.38	.93	.71	.28	.79	.58	- 2.31	.53	- .95
Writing	24	.32	.87	.59	.11	.66	.46	- 1.78	2.69	- .34
Mathematics	52	.41	.96	.75	.12	.72	.51	- 3.14	- .16	- 1.40
Numbers	16	.57	.96	.78	.25	.68	.51	- 2.43	- .40	- 1.41
Computation	16	.59	.91	.79	.34	.71	.58	- 2.75	- .28	- 1.52
Measurement/Geometry	16	.41	.94	.70	.12	.53	.37	- 2.98	.50	- 1.21
Interpretation	16	.54	.94	.75	.39	.62	.50	- 3.04	- .15	- 1.32
Science	52	.34	.96	.63	.18	.67	.41	- 4.42	1.26	- .75
Life Science	24	.38	.96	.64	.17	.55	.34	- 4.34	1.20	- .84
Earth Science	28	.34	.90	.65	.18	.69	.43	- 2.29	1.41	- .66
Physical Science	28	.44	.90	.63	.24	.67	.46	- 2.15	.47	- .64
Social Studies	64	.14	.91	.63	-.04	.79	.49	- 2.92	8.59	- .45
Geography	40	.39	.91	.66	.30	.65	.47	- 3.16	.73	- .82
Government	24	.19	.86	.61	.19	.76	.46	- 1.77	2.78	- .40
Economics	24	.14	.91	.64	-.05	.72	.46	- 3.06	8.80	- .40
Civics	24	.19	.86	.62	.18	.75	.48	- 1.77	2.88	- .40

Grade 5 Form D Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	84	.35	.96	.74	-.06	.75	.49	-3.25	2.13	-1.25
Language Arts	20	.51	.94	.78	.29	.65	.49	-2.61	-.04	-1.55
Reading	48	.35	.96	.73	.12	.76	.51	-3.12	.32	-1.19
Writing	20	.35	.95	.69	-.10	.61	.40	-3.23	1.55	-.99
Mathematics	60	.28	.96	.58	.21	.74	.47	-4.27	1.93	-.81
Numbers	16	.40	.95	.68	.29	.71	.54	-2.99	.86	-.97
Computation	16	.38	.88	.71	.37	.64	.50	-2.17	.60	-1.03
Measurement/Geometry	16	.28	.85	.54	.19	.65	.45	-4.07	.77	-.35
Interpretation	20	.28	.96	.58	.22	.59	.43	-3.86	1.85	-.62
Science	68	.23	.92	.58	.02	.52	.35	-3.66	2.57	-.39
Life Science	20	.38	.92	.65	.21	.51	.34	-4.14	1.23	-1.02
Earth Science	28	.23	.92	.50	-.02	.42	.27	-4.33	1.70	-.02
Physical Science	28	.26	.92	.61	.17	.54	.37	-3.48	2.80	-.54
Social Studies	80	.24	.90	.62	.02	.74	.46	-2.19	2.48	-.60
Geography	28	.34	.84	.65	.23	.60	.44	-1.79	1.27	-.80
History	24	.24	.83	.60	.22	.69	.42	-1.86	2.54	-.46
Government	20	.40	.82	.60	.23	.60	.43	-1.68	.66	-.48
Economics	20	.24	.90	.63	.07	.60	.41	-2.11	2.76	-.59
Civics	32	.24	.90	.62	.21	.66	.44	-2.04	2.59	-.55

Grade 6 Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	92	.24	.93	.69	-.09	.80	.49	-2.75	7.08	-.79
Language Arts	24	.24	.92	.66	.10	.67	.46	-2.24	2.87	-.67
Reading	52	.36	.93	.77	.28	.83	.55	-2.77	.94	-1.35
Writing	24	.26	.79	.56	-.10	.65	.40	-1.52	7.20	.13
Mathematics	104	.09	.96	.65	.06	.76	.51	-3.47	6.94	-.63
Numbers	24	.42	.95	.71	.17	.64	.50	-2.76	.33	-1.08
Computation	28	.09	.96	.64	.15	.71	.47	-3.75	6.00	-.53
Measurement/Geometry	24	.17	.87	.58	.21	.58	.43	-2.31	2.97	-.27
Interpretation	28	.49	.92	.70	.31	.75	.58	-2.14	.48	-.90
Science	92	.21	.94	.56	-.07	.64	.36	-2.96	7.42	-.13
Life Science	48	.21	.87	.57	.06	.58	.39	-2.40	3.63	-.41
Earth Science	20	.29	.83	.58	-.05	.49	.33	-2.92	2.41	-.40
Physical Science	32	.21	.94	.53	-.03	.54	.28	-2.92	6.49	.29
Social Studies	84	.19	.98	.70	.12	.78	.53	-3.77	3.77	-1.02
Geography	28	.19	.98	.70	.14	.69	.48	-3.94	3.57	-.96
History	20	.42	.93	.70	.32	.71	.53	-2.15	.49	-.89
Government	16	.47	.92	.71	.21	.70	.46	-2.51	.27	-1.04
Economics	20	.40	.93	.70	.21	.77	.52	-2.16	.87	-.95
Civics	32	.47	.95	.77	.22	.81	.55	-2.90	.31	-1.33

Grade 6 Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	92	.23	.97	.70	-.12	.81	.49	-3.04	8.95	-.94
Language Arts	24	.23	.93	.71	-.11	.69	.46	-2.92	7.91	-.83
Reading	52	.43	.97	.77	.25	.89	.54	-2.92	.45	-1.39
Writing	24	.23	.84	.58	.07	.64	.40	-1.83	3.38	-.19
Mathematics	104	.26	.94	.69	.14	.75	.53	-4.22	-2.05	-.96
Numbers	24	.44	.93	.74	.34	.65	.54	-2.02	.27	-1.15
Computation	28	.37	.94	.68	.13	.79	.51	-4.24	.64	-1.10
Measurement/Geometry	24	.26	.87	.62	.17	.59	.45	-2.45	1.89	-.59
Interpretation	28	.46	.91	.74	.30	.76	.59	-2.38	.26	-1.09
Science	92	.24	.90	.59	-.01	.61	.40	-3.97	5.78	-.07
Life Science	48	.29	.89	.60	.16	.66	.41	-3.29	3.87	.20
Earth Science	20	.25	.89	.57	.12	.54	.33	-2.93	3.37	.13
Physical Science	32	.24	.90	.58	.08	.55	.35	-2.62	5.55	.17
Social Studies	84	.26	.98	.73	.13	.87	.53	-4.15	1.43	-1.21
Geography	28	.26	.98	.73	.12	.77	.49	-4.14	1.28	-1.38
History	20	.45	.97	.70	.30	.77	.52	-2.92	.39	-.93
Government	16	.44	.91	.73	.20	.74	.53	-2.49	.56	-1.00
Economics	20	.46	.97	.73	.24	.75	.51	-2.75	.26	-1.11
Civics	32	.44	.97	.78	.18	.90	.57	-3.43	.59	-1.47

Table

Grade 7 Form D StatisticsSpring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	100	.40	.95	.71	.19	.81	.50	-3.71	.55	-1.15
Language Arts	20	.41	.94	.74	.29	.75	.49	-2.63	.55	-1.22
Reading	60	.40	.95	.68	.18	.74	.49	-3.77	.55	-1.07
Writing	32	.42	.94	.71	.17	.76	.46	-2.58	.38	-1.13
Mathematics	92	.25	.90	.58	.20	.70	.51	-2.56	2.10	-.38
Numbers	20	.28	.78	.57	.21	.70	.63	-1.13	2.12	-.26
Computation	24	.25	.90	.53	.23	.59	.48	-2.35	1.82	-.16
Measurement/Geometry	20	.28	.87	.63	.27	.67	.49	-2.74	1.09	-.65
Interpretation	28	.39	.85	.60	.31	.66	.50	-1.77	.77	-.51
Science	76	.22	.95	.58	-.10	.61	.36	-3.42	9.33	-.30
Life Science	32	.35	.95	.67	.18	.59	.38	-3.15	1.09	-1.04
Earth Science	40	.22	.86	.53	-.10	.53	.31	-2.15	9.41	.19
Physical Science	24	.30	.83	.54	.15	.53	.37	-1.65	2.65	-.10
Social Studies	72	.20	.90	.62	-.01	.69	.47	-3.44	7.38	-.50
Geography	32	.33	.90	.64	.24	.65	.46	-2.19	1.19	-.69
History	20	.33	.88	.61	.16	.65	.47	-2.02	1.96	-.47
Government	20	.20	.88	.64	.02	.63	.41	-3.26	6.14	-.46
Economics	20	.33	.79	.59	.27	.65	.46	-1.35	1.30	-.37
Civics	36	.20	.88	.65	.00	.68	.45	-3.58	6.96	-.58

Grade 8 Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	108	.30	.99	.68	.15	.73	.48	- 3.05	2.53	- .77
Language Arts	20	.47	.89	.68	.29	.64	.45	- 2.39	.20	- .92
Reading	60	.40	.99	.71	.16	.71	.50	- 3.27	1.36	- 1.03
Writing	36	.30	.87	.59	.13	.65	.49	- 1.88	2.85	- .32
Mathematics	100	.14	.88	.59	.08	.70	.46	- 2.00	5.16	- .24
Numbers	16	.32	.76	.61	.31	.63	.49	- 1.27	1.29	- .48
Computation	24	.14	.82	.57	.12	.65	.47	- 1.71	5.09	- .14
Measurement/Geometry	28	.17	.84	.56	.05	.58	.42	- 1.63	3.72	- .03
Interpretation	32	.28	.88	.60	.10	.66	.44	- 2.02	1.75	- .42
Science	72	.16	.93	.56	-.12	.63	.36	- 2.81	11.89	.03
Life Science	28	.24	.93	.62	.11	.62	.39	- 2.42	2.76	- .49
Earth Science	24	.30	.83	.57	.06	.57	.33	- 2.63	1.23	- .43
Physical Science	36	.16	.79	.52	-.11	.56	.31	- 3.01	11.14	.32
Social Studies	72	.27	.96	.71	.11	.79	.55	- 2.84	2.46	- .87
Geography	24	.32	.96	.72	.10	.74	.55	- 3.17	2.49	- .95
History	16	.27	.84	.61	.29	.61	.48	- 1.55	1.81	- .37
Government	28	.38	.94	.72	.31	.69	.53	- 2.18	.80	- 1.02
Economics	16	.54	.84	.72	.40	.74	.56	- 1.43	-.17	- .95
Civics	44	.38	.94	.70	.32	.74	.55	- 2.36	.88	- 1.03

Grade 8 Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	108	.35	.98	.71	.08	.81	.50	-2.85	2.72	-1.03
Language Arts	20	.38	.86	.71	.26	.64	.45	-2.11	.85	-1.12
Reading	60	.36	.98	.73	.17	.84	.51	-2.93	.65	-1.22
Writing	36	.35	.90	.66	.07	.68	.48	-1.86	2.66	-.75
Mathematics	100	.23	.96	.65	.14	.72	.49	-3.15	2.18	-.69
Numbers	16	.55	.85	.71	.37	.60	.50	-2.05	-.22	-.98
Computation	24	.23	.93	.62	.31	.58	.46	-3.39	1.78	-.66
Measurement/Geometry	28	.29	.88	.62	.26	.58	.48	-2.12	1.13	-.60
Interpretation	32	.26	.96	.66	.38	.75	.52	-3.06	2.20	-.65
Science	72	.23	.92	.62	.00	.63	.38	-2.69	6.48	-.54
Life Science	28	.40	.92	.69	.20	.63	.42	-2.67	.96	-1.04
Earth Science	24	.35	.88	.64	.62	.11	.36	-2.76	2.29	-.84
Physical Science	36	.23	.83	.57	.01	.52	.34	-2.72	6.59	-.30
Social Studies	72	.29	.98	.73	.25	.78	.55	-4.07	1.30	-1.10
Geography	24	.51	.98	.74	.22	.74	.49	-4.15	-.12	-1.28
History	16	.29	.77	.60	.23	.61	.45	-1.31	1.18	-.43
Government	28	.58	.92	.75	.33	.76	.55	-2.58	-.34	-1.30
Economics	16	.59	.89	.77	.31	.83	.59	-1.52	-.39	-1.21
Civics	44	.58	.92	.76	.31	.79	.58	-2.74	-.34	-1.31

Table

Grade 9 Form D StatisticsSpring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	112	.25	.96	.70	.01	.78	.51	-3.52	4.43	-.97
Language Arts	24	.43	.91	.68	.19	.68	.49	-1.91	.39	-.91
Reading	60	.33	.96	.73	.08	.78	.52	-3.61	2.75	-1.16
Writing	32	.25	.92	.67	.02	.67	.46	-2.13	4.06	-.69
Mathematics	68	.20	.85	.58	-.01	.75	.48	-2.96	5.27	-.27
Computation	24	.33	.83	.62	.15	.69	.50	-2.88	1.12	-.64
Measurement/Geometry	20	.20	.85	.50	.32	.69	.47	-1.30	2.28	.10
Interpretation	24	.27	.84	.59	.01	.67	.44	-2.04	4.68	-.24
Science	68	.18	.87	.50	-.17	.55	.28	-3.93	7.02	.57
Life Science	32	.22	.85	.55	-.08	.53	.32	-2.63	7.11	.27
Earth Science	28	.22	.85	.56	-.07	.53	.32	-2.79	7.07	.25
Physical Science	60	.18	.87	.48	-.14	.48	.27	-3.71	6.93	.69
Cross-Disciplinary	16	.21	.77	.44	.02	.34	.19	-2.09	4.59	1.18
Social Studies	80	.15	.94	.66	.14	.73	.47	-2.41	3.08	-.71
Geography	28	.45	.94	.74	.30	.70	.50	-2.61	.29	-1.24
History	16	.30	.89	.67	.02	.61	.42	-2.51	2.62	-.67
Government	28	.33	.90	.66	.25	.69	.48	-2.17	1.15	-.72
Economics	20	.15	.94	.61	.22	.53	.38	-2.76	2.66	-.57
Civics	52	.15	.94	.65	.21	.70	.47	-2.41	2.86	-.70

Grade 10 Form A Statistics

Spring 1987 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	112	.16	.95	.68	-.30	.80	.47	-3.68	16.08	-.74
Language Arts	32	.43	.92	.70	.17	.75	.48	-2.38	.84	-1.03
Reading	60	.28	.95	.73	.23	.74	.49	-3.61	2.01	-1.27
Writing	24	.16	.75	.50	-.24	.70	.36	-1.11	14.02	.77
Mathematics	92	.22	.92	.59	-.10	.73	.46	-2.60	6.46	-.25
Computation	24	.26	.81	.56	.21	.73	.47	-2.66	1.80	-.27
Measurement/Geometry	32	.24	.88	.55	.13	.73	.48	-1.72	3.56	-.07
Interpretation	36	.22	.92	.63	-.11	.69	.44	-2.54	6.44	-.45
Science	80	.18	.95	.52	-.07	.60	.31	-3.39	10.46	.45
Life Science	24	.31	.89	.56	.02	.50	.28	-2.04	2.04	-.21
Earth Science	28	.19	.95	.53	-.03	.49	.24	-3.61	8.60	.43
Physical Science	20	.18	.65	.42	-.04	.45	.24	-.89	9.72	1.44
Cross-Disciplinary	24	.21	.95	.58	.02	.57	.36	-2.24	3.68	-.24
Social Studies	100	.37	.92	.65	.12	.77	.53	-3.76	.95	-.70
Geography	28	.42	.92	.66	.30	.68	.49	-2.04	.35	-.77
History	20	.37	.81	.61	.10	.63	.47	-1.29	.90	-.48
Government	40	.37	.85	.65	.12	.75	.53	-3.37	1.00	-.67
Economics	28	.37	.84	.62	.11	.67	.51	-1.44	.67	-.54
Civics	68	.37	.88	.64	.13	.78	.54	-3.63	.95	-.65

Grade 10 Form B Statistics

Spring 1988 Administration

Subject Cluster	Number of Items	p values			Biserial			Thresholds		
		Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean
Reading	112	.27	.95	.69	.00	.75	.45	-4.36	5.06	-.89
Language Arts	32	.39	.93	.73	.21	.65	.47	-4.01	.00	-1.09
Reading	60	.27	.95	.74	-.02	.73	.44	-3.99	5.26	-1.13
Writing	24	.33	.80	.53	.19	.74	.40	-1.74	1.82	-.04
Mathematics	92	.21	.93	.60	-.08	.73	.47	-3.83	4.41	-.39
Computation	24	.31	.84	.58	.18	.74	.52	-2.57	1.50	-.39
Measurement/Geometry	32	.28	.90	.59	.14	.69	.47	-1.90	1.65	-.34
Interpretation	36	.21	.93	.65	-.10	.70	.44	-3.36	6.18	-.42
Science	80	.16	.95	.63	.01	.64	.34	-3.11	5.85	.19
Life Science	24	.27	.84	.57	.14	.47	.32	-2.17	2.93	-.23
Earth Science	28	.19	.86	.50	-.01	.40	.26	-3.38	4.97	.49
Physical Science	20	.16	.68	.48	.09	.48	.30	-1.18	5.70	.48
Cross Disciplinary	24	.28	.95	.58	-.01	.59	.34	-2.27	2.74	-.20
Social Studies	100	.35	.95	.68	.09	.79	.53	-2.91	2.31	-.89
Geography	28	.35	.95	.67	.08	.63	.47	-2.73	2.44	-.83
History	20	.44	.83	.67	.22	.64	.45	-1.70	.37	-.84
Government	40	.45	.91	.71	.33	.77	.55	-2.42	.38	-.94
Economics	28	.35	.87	.68	.25	.72	.53	-1.82	1.21	-.83
Civics	68	.35	.92	.70	.27	.81	.56	-2.63	1.11	-.92

Appendix F
Pearson Product Moment Coefficients for
Corresponding ITBS/TAP and MMAT Subject Tests

Pearson Product Moment Coefficients for Corresponding ITBS/TAP and

MMAT, Form A, Subjects

Spring 1987 Administration

Subject	Grade			
	3	6	8	10
Reading	.799	.786	.808	.756
Language Arts	.799	.781	.792	—
Mathematics	.809	.860	.870	.856
Science	.713	.809	.743	.786
Social Studies	.759	.800	.789	.843

Pearson Product Moment Coefficients for Corresponding ITBS/TAP and

MMAT, Form D, Subjects

Spring 1988 Administration

	Grade				
Subject	2	4	5	7	9
Reading	.838	.836	.826	.839	.813
Language Arts	.795	.804	.816	.796	—
Mathematics	.751	.820	.836	.856	.833
Science	—	.747	.731	.803	.701
Social Studies	—	.772	.806	.813	.846

Pearson Product Moment Coefficients for Corresponding ITBS/TAP and

MMAT, Form B, Subjects

Spring 1988 Administration

	Grade			
Subject	3	6	8	10
Reading	.782	.776	.791	.747
Language Arts	.796	.807	.801	—
Mathematics	.829	.871	.867	.853
Science	.703	.799	.727	.777
Social Studies	.689	.742	.740	.849

Appendix G
Common Factor Analysis
Grade 10, Form A

A principal factors analysis was undertaken using the results of the state sample, Form A, Grade 10. Communalities were set to the squared multiple correlation with all other variables. Inspection of the 25 factors, to which the intercorrelation matrix was reduced, revealed that a five factor solution would give maximum useable information. The proportion of variance asymptoted at eigenvalues of 1.5. The variance explained by each of the three factors is listed below:

Variance Explained by Each Factor

Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
45.998	5.691	3.219	2.328	1.922

Final Communality Estimates: Total = 59.158

Oblique rotation using the Kaiser procedure was employed to reduce the factor pattern. Interfactor correlations are reported below:

Interfactor Correlations

	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	0.711	0.727	0.745	0.715
Factor 2		0.557	0.695	0.828
Factor 3			0.682	0.601
Factor 4				0.792

As expected the factors are correlated. Factor 1 relates more strongly to each of the other factors than do any of the others. Factor 2 relates more to Factors 1 and 5. Factor 3 relates more to Factor 1. Factor 4 relates more to Factors 1 and 5. And Factor 5 relates more to Factors 1, 2, and 4.

The standard regression coefficients indicate the magnitude to which each item loads on a factor. By identifying which items have loadings greater than .15, from which subtests these items come, and the tasks which these items comprise, a conclusion about what the factor measures can be drawn. Listed below are the numbers of items for each subtest with major loadings on each factor as well as the magnitude of the loading.

Number of Items Loading on Each Factor by Magnitude of Standard Regression Coefficient

Magnitude	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
SRC>.24					
English	53	1	5	1	11
Mathematics	0	53	10	0	12
Science	0	2	6	0	31
Social Studies	0	0	20	53	29
.14<SRC<.25					
English	48	9	8	15	20
Mathematics	2	21	12	0	20
Science	1	2	6	3	21
Social Studies	0	8	5	22	10
.11<SRC<.15					
English	6	6	9	6	6
Mathematics	1	7	6	1	7
Science	2	8	4	5	5
Social Studies	1	9	9	1	8
-.11>SRC>-.15					
English	1	6	6	2	3
Mathematics	0	1	6	6	1
Science	1	5	5	1	0
Social Studies	9	2	4	0	0
-.14>SRC>-.25					
English	0	12	4	2	7
Mathematics	0	0	10	20	1
Science	0	1	7	4	1
Social Studies	13	0	7	0	4
-.24>SRC					
English	0	0	0	0	0
Mathematics	0	0	2	4	2
Science	1	1	0	0	0
Social Studies	4	0	1	0	1

Factor 1 seems to load with items from the English/Language Arts area with few loadings of items from other areas. The five highest loading items may be described in this way:

Subject	Stimulus	Task
English	Text	Establishing Word meaning in context
English	Text	Identifying causality
English	Stem	Identifying correct punctuation
English	Sample Index	Identifying where something is found
English	Text	Identifying point of view

Three items from other areas which load on this factor may be described in this way:

Subject	Stimulus	Task
Mathematics	Text	Drawing a conclusion
Science	Text	Selecting order of experiment
Social Studies	Stem	Selecting inference

Based on the content and the tasks required of examinees, this factor has been labeled English.

Factor 2 seems to load items from the Mathematics area with items from the Social Studies and Science areas contributing some, followed by a few English Subject items. Items with the five highest loadings on this factor may be described in this way:

Subject	Stimulus	Task
Mathematics	Expression	Selecting algebraic equivalent
Mathematics	Expression	Selecting algebraic equivalent
Mathematics	Illustration	Finding a perimeter
Mathematics	Expression	Selecting algebraic equivalent
Mathematics	Expression	Solving for a value

Items from other subjects which load on this factor can be described in this way:

Subject	Stimulus	Task
English	Text	Sequencing sentences
Science	Stem	Expressing in scientific notation
Social Studies	Figure	Drawing conclusion from data in figure

On the basis of these loadings and the tasks required of the examinee this factor has been labeled Mathematics.

Factor 3 seems to be composed of items from the Social Studies and Mathematics areas with some English and Science items. The five items contributing significantly to this factor may be described as follows:

Subject	Stimulus	Task
Mathematics	Word Problem	Solving for a value
Mathematics	Text	Selecting data which supports a conclusion
Social Studies	Stem	Giving directions
English	Illustration	Interpreting what information is requested on forms
Science	Map	Concluding effects of an action taken

Three items contributing in a lesser way may be described in this way:

Subject	Stimulus	Task
Social Studies	Stem	Identifying an effect of an action
Math	Word Problem	Solving for a value
English	Text	Identifying causality

Based on the loadings and on the tasks required of an examinee, this factor is labeled Inductive Reasoning in Social Studies and Mathematics.

Factor 4 seems to be composed primarily of items from the Social Studies area with a few from English. The five items contributing most to this factor may be described in this way:

Subject	Stimulus	Task
Social Studies	Text	Identifying differences
Social Studies	Text	Identifying purpose of a persuasive presentation
Social Studies	Text	Drawing a conclusion from information
Social Studies	Stem	Determining which is not an example
Social Studies	Text	Selecting an interpretation

Three items from subjects other than Social Studies which load on this factor may be described in this way:

Subject	Stimulus	Task
Science	Text	Evaluating a hypothesis
English	Text	Selecting a complete sentence
English	Stem	Selecting the best revision

Because the loadings seem to come primarily from the Social Studies/Civics area and because the tasks from other subjects require responses similar to those for Social Studies, this factor has been labeled Social Studies/Civics.

Factor 5 is composed primarily of items from the Social Studies/Civics and the Science areas, although both Mathematics and English contribute a large number of items with minor loadings. What is required of an examinee on the five highest loading items may be described in this way:

Subject	Stimulus	Task
Social Studies	Map	Locating a characteristic
Social Studies	Map	Locating a characteristic
Social Studies	Map	Locating a characteristic
Science	Diagram	Drawing a conclusion
Social Studies	Time Line	Identifying the location of an event along the line

Three other contributing items may be described in this way:

Subject	Stimulus	Task
Math	Word Problem	Solving for a value
English	Text	Drawing a conclusion
Science	Stem	Drawing a conclusion

Based on these descriptions and the pattern of loadings, this factor has been labeled Deductive Reasoning in Science, Social Studies/Civics, Mathematics, and English/Language Arts.

Appendix H
Mastery Classification Reliability Estimates

Mastery Classification Reliability Estimates

MMAT, Form A

Spring 1987 Administration

	P		Kappa	
	Low	High	Low	High
Reading				
Grade 3	0.651	0.853	0.108	0.762
Grade 6	0.565	0.870	0.032	0.439
Grade 8	0.568	0.863	0.074	0.440
Grade 10	0.592	0.813	0.128	0.435
Mathematics				
Grade 3	0.734	0.909	0.195	0.665
Grade 6	0.657	0.856	0.144	0.686
Grade 8	0.614	0.811	0.072	0.500
Grade 10	0.550	0.768	0.033	0.443
Science				
Grade 3	0.004	0.791	0.012	0.467
Grade 6	0.571	0.875	0.028	0.360
Grade 8	0.567	0.733	0.031	0.376
Grade 10	0.544	0.768	0.025	0.321
Social Studies				
Grade 3	0.660	0.868	0.227	0.495
Grade 6	0.550	0.864	0.078	0.481
Grade 8	0.658	0.829	0.176	0.442
Grade 10	0.599	0.740	0.165	0.442

Appendix I
MMAT Manuals and Reports

Manuals

Thirteen manuals have been developed for distribution to the schools:

- *Core Competencies and Key Skills for Missouri Schools*
- *Test Content Specifications: Missouri Mastery and Achievement Tests, Grade 2, Language Arts/Reading and Mathematics*
- *Test Content Specifications: Missouri Mastery and Achievement Test, Grade 3, Language Arts/Reading, Mathematics, Science, and Social Studies/Civics*
- *Test Content Specifications: Missouri Mastery and Achievement Test, Grade 4, Language Arts/Reading, Mathematics, Science, and Social Studies/Civics*
- *Test Content Specifications: Missouri Mastery and Achievement Test, Grade 5, Language Arts/Reading, Mathematics, Science, and Social Studies/Civics*
- *Test Content Specifications: Missouri Mastery and Achievement Test, Grade 6 Language Arts/Reading, Mathematics, Science, and Social Studies/Civics*
- *Test Content Specifications: Missouri Mastery and Achievement Test, Grade 7 Language Arts/Reading, Mathematics, Science, and Social Studies/Civics*
- *Test Content Specifications: Missouri Mastery and Achievement Test, Grade 8 Language Arts/Reading, Mathematics, Science, and Social Studies/Civics*
- *Test Content Specifications: Missouri Mastery and Achievement Test, Grade 9 Language Arts/Reading, Mathematics, Science, and Social Studies/Civics*
- *Test Content Specifications: Missouri Mastery and Achievement Test, Grade 10 Language Arts/Reading, Mathematics, Science, and Social Studies/Civics*
- *Missouri Mastery and Achievement Tests: Examiner's Manual and Directions*
- *Missouri Mastery and Achievement Tests: Guide to Score Interpretation and Use*
- *Missouri Mastery and Achievement Tests: District Test Coordinator's Directions*

Reports

Reports

Eleven reports have been developed for the MMAT:

- Academic Achievement Report
- Chapter 1 Pre/Post Evaluation Report
- Chapter 1 Eligibility List
- Grade Level Cluster Report
- Grade Level Key Skill Report
- Individual Student Label
- Individual Student Report
- Longitudinal Trend Report
- Pupil List Report
- Score Distribution Report
- Summary Report