The Illinois State Assessment
2000 Technical Manual

Illinois State Board of Education
Assessment Division

CONTENTS

1. PURPOSE AND DESIGN OF THE ISAT TESTING PROGRAM

In February 2000, students in grades 3, 5, 8, and 10 took Illinois Standards Achievement Tests (ISAT) in reading, mathematics, and writing. Students in grades 4 and 7 took ISAT tests in science and social science. Approximately 800,000 students enrolled in public elementary and secondary schools across the state participated in the testing program. ISAT measures the extent to which students are meeting the Illinois Learning Standards. Illinois teachers and curriculum experts developed the ISAT tests in cooperation with the Illinois State Board of Education (ISBE).

This manual provides technical information about the 2000 test administration. It describes the tests and assessment approaches and addresses technical concerns. Other reports, documents, or publications issued by the Illinois State Board of Education (ISBE) provide additional information about interpreting test results (*Guide to the 2000 Illinois State Assessment*, *Understanding Your Child's ISAT Scores*) that is not included here.

General Procedures

Each ISAT test is designed to ensure that its results validly and fairly assess the Illinois Learning Standards. The selection of items and assembly of each test is guided by a set of specifications. These specifications were developed by Illinois educators to help ensure that test content corresponds to the purposes, objectives, and skills framed by the learning standards.

Illinois teachers and administrators participate in all phases of the test development process: item writing, item selection, bias review, and test assembly. The State Board of Education convenes a series of advisory committees to ensure that test development is continually informed and guided by the recommendations of content authorities, measurement specialists, and practitioners. The following evaluation criteria are applied to all assessment material used in the Illinois program:

Content. Every item is screened for alignment with the Illinois Learning Standards, grade-level appropriateness, importance, and clarity. Incorrect choices (for multiple-choice items) are reviewed for plausibility. In tests other than reading, the complexity of the text of the questions is kept to the minimum necessary to state the problem.

Difficulty. Items are pilot tested on large samples of students prior to their inclusion in tests to develop a statistical profile for each item. Items that are too easy or too difficult and, therefore, provide little or no information are omitted.

Precision. Point-biserial (i.e., item-test) correlations evaluate the extent to which an item distinguishes between less proficient and more proficient students. Reviewers usually omit items with a point-biserial of less than .30 and select items with the highest point-biserial.

Fairness. Test items and forms undergo regular sensitivity reviews and statistical analyses to ensure that all materials meet fairness criteria with respect to the cultural and ethnic diversity of Illinois public schools.

ISBE takes several precautions to help ensure test security. Test materials shipped to schools are packaged and sealed. Each test booklet is barcoded so that it can be accounted for. The administration of tests is standardized. A series of manuals provides guidance on security and other issues to the district testing coordinator, school testing coordinator, and classroom test administrator. After administration, all materials are removed from schools and returned to a central facility for processing and secure destruction of unneeded materials.

Reading

The ISAT reading test assesses material defined by standards associated with three state learning goals. The standards were developed using the 1985 State Goals for Language Arts, various state and national standards drafts, and local education standards contributed by team members. These learning standards are designed to guide language arts instruction in Illinois schools. This alignment of assessment to curriculum insures consistency and strengthens the influence of standards and assessment on improved teaching and learning. These standards are:

- Goal 1: Read with understanding and fluency.
  - 1A. Apply word analysis and vocabulary skills to comprehend selections.
  - 1B. Apply reading strategies to improve understanding and fluency.
  - 1C. Comprehend a broad range of reading materials.

- Goal 2: Read and understand literature representative of various societies, eras and ideas.
  - 2A. Understand how literary elements and techniques are used to convey meaning.
  - 2B. Read and interpret a variety of literary works.

- Goal 5: Write to communicate for a variety of purposes.
  - 5A. Locate, organize, and use information from various sources to answer questions, solve problems and communicate ideas.
  - 5B. Analyze and evaluate information acquired from various sources.
  - 5C. Apply acquired information, concepts and ideas to communicate in a variety of formats.

The reading test has two formats. The grade 3 reading assessment is given in three 35-minute sessions. One of these sessions consists of 12-15 word analysis questions and one passage followed by 15-17 multiple-choice questions. The two remaining sessions include one passage followed by 15-20 multiple-choice questions, and one short answer question.

The reading tests for grades 5, 8, and 10 are also given in three 35-minute sessions. One of these sessions consists of a longer passage with 15-20 multiple-choice questions. The other

two sessions each include one passage with 15-20 multiple choice questions and one short answer question.

The reading passages and accompanying questions reflect two of the most frequent purposes for reading—reading to gain information and reading for literary experience. The sources for these passages range from high interest, grade-appropriate periodicals to newspapers, short stories, and novels. Illinois teachers reviewed and selected the material for these tests.

The multiple-choice questions require students to select one correct response from four possibilities presented to them. Again, teachers in Illinois played an active part in writing, reading, and editing these test questions. Questions must meet both content and statistical criteria for inclusion in the test.

The short answer questions on the reading test require students not only to read and understand a text, but also to analyze, evaluate, and interpret the text as a means of making connections and conclusions related to the text. The rubric used to score the short answer responses is a holistic scoring rubric. It describes characteristics of different levels of achievement in reading. The levels of achievement on the reading rubric range from 0 to 4 (4 being the highest score). Responses with scores of 0 indicate that the student response is insufficient to effectively determine evidence of achievement in reading. Responses with scores of 1 and 2 indicate developing levels of achievement in reading. Responses with scores of 3 indicate a developed level of achievement in reading. Finally, responses with scores of 4 represent a well-developed level of achievement in reading. The rubric was developed with Illinois educators. The reporting of the short answer item scores is different than that of the other questions. For the first two years of the assessment, schools and districts will receive the short answer item scores for informational purposes rather than accountability purposes.

In addition to an overall reading score, results are reported in terms of the percent of items correctly answered within five "standard sets" (six at grade 3). These scores are as follows:

- *Comprehension: Literary Works:* Understanding of passages taken from sources such as novels, short stories, and periodicals. (Standards 1B, 1C, 2A, 2B, 5A, 5B, 5C)

- *Comprehension: Informational Sources:* Understanding of non-fiction texts such as student periodicals, newspapers, and trade journals. (Standards 1B, 1C, 2A, 2B, 5A, 5B, 5C)

- *Application of Strategies: Explicit Ideas:* Identifying important information directly stated in the text. (Standards 1B, 5A)

- *Application of Strategies: Inferences from Text:* Analyzing important information in the text to draw logical conclusions about the text. (Standards 1C, 2A, 2B, 5B, 5C)

- *Vocabulary:* Using contextual clues and other skills to understand key words, phrases, and concepts in literary and informational texts. (Standard 1A)

- *Word Analysis (3rd grade only):* Using phonics, word pattern, and other word analysis skills to recognize new words. (Standard 1A)

Mathematics

People use mathematics to identify, describe and investigate the patterns and challenges of everyday living. Mathematics helps us to understand events that have occurred and to predict and prepare for events to come so that we can more fully understand our world and more successfully live in it. Mathematics encompasses arithmetic, measurement, algebra, geometry, trigonometry, statistics, probability and other fields. It deals with numbers, quantities, shapes and data, as well as numerical relationships and operations. Confronting, understanding and solving problems is at the heart of mathematics. Mathematics is much more than a collection of concepts and skills; it is a way of approaching new challenges through investigating, reasoning, visualizing and problem-solving with the goal of communicating the relationships observed and problems solved to others.

The ISAT mathematics tests are designed to measure the following learning standards.

- Goal 6: Demonstrate and apply a knowledge and sense of numbers, including numeration and operations (addition, subtraction, multiplication, division), patterns, ratios and proportions.
  - 6A. Demonstrate knowledge and use of numbers and their representations in a broad range of theoretical and practical settings.
  - 6B. Investigate, represent and solve problems using number facts, operations (addition, subtraction, multiplication, division) and their properties, algorithms and relationships.
  - 6C. Compute and estimate using mental mathematics, paper-and-pencil methods, calculators and computers.
  - 6D. Solve problems using comparison of quantities, ratios, proportions and percents.

- Goal 7: Estimate, make and use measurements of objects, quantities and relationships and determine acceptable levels of accuracy.
  - 7A. Measure and compare quantities using appropriate units, instruments and methods.
  - 7B. Estimate measurements and determine acceptable levels of accuracy.
  - 7C. Select and use appropriate technology, instruments and formulas to solve problems, interpret results and communicate findings.

- Goal 8: Use algebraic and analytical methods to identify and describe patterns and relationships in data, solve problems and predict results.
  - 8A. Describe numerical relationships using variables and patterns.
  - 8B. Interpret and describe numerical relationships using tables, graphs and symbols.
  - 8C. Solve problems using systems of numbers and their properties.
  - 8D. Use algebraic concepts and procedures to represent and solve problems.

- Goal 9: Use geometric methods to analyze, categorize and draw conclusions about points, lines, planes and space.

9A. Demonstrate and apply geometric concepts involving points, lines, planes and space.
9B. Identify, describe, classify and compare relationships using points, lines, planes and solids.
9C. Construct convincing arguments and proofs to solve problems.
9D. Use trigonometric ratios and circular functions to solve problems.

- Goal 10: Collect, organize and analyze data using statistical methods; predict results; and interpret uncertainty using concepts of probability.
  10A. Organize, describe and make predictions from existing data.
  10B. Formulate questions, design data collection methods, gather and analyze data and communicate findings.
  10C. Determine, describe and apply the probabilities of events.

Illinois teachers developed the Illinois Learning Standards for Mathematics. These goals, standards and benchmarks are an outgrowth of the 1985 Illinois State Goals for Learning influenced by the latest thinking in school mathematics. This includes the National Council of Teachers of Mathematics; Curriculum and Evaluation Standards for School Mathematics; ideas underlying recent local and national curriculum projects; results of state, national, and international assessment findings; and the work and experiences of Illinois school districts and teachers.

The mathematics assessment includes 80 multiple-choice items administered in three test sessions. In addition, the tests contain two short-answer/problem-solving tasks. As is true in reading, for the first two years of the assessment, schools and districts will receive the short answer item scores for informational purposes rather than accountability purposes.

In addition to an overall mathematics score, results are reported in terms of the percent of items correctly answered within eight standard sets. These scores are as follows:

- *Estimation/Number Sense/Computation:* Demonstrating an understanding of numbers, their representations, and number operations of addition, subtraction, multiplication, division, percentages, and fractions as appropriate to grade level. (Standards 6A, 6B, 6C, 6D, 8C)

- *Algebraic Patterns/Variables:* Identifying, describing, and extending algebraic, geometric, and numeric patterns and constructing and solving problems using variables. (Standards 8A, 8C, 8D)

- *Algebraic Relationships/Representations:* Representing and interpreting algebraic concepts with words, diagrams, tables, coordinate graphs, equations, and inequalities. (Standards 8B, 8C)

- *Geometric Concepts:* Identifying and describing points, lines, two- and three-dimensional shapes and their properties, such as parallel; symmetry; perpendicular; and number of sides, faces, and vertices. (Standards 8C, 9A)

- *Geometric Relationships:* Sorting, classifying, comparing, and contrasting geometric figures. This category includes such properties as similarity and congruency. (Standards 8C, 9B, 9D)

- *Measurement:* Estimating, measuring, and comparing quantities using appropriate units and acceptable levels of accuracy. At higher grades, this category encompasses conversions within measurement systems. (Standards 7A, 7B, 7C, 8C)

- *Data Organization/Analysis:* Creating, analyzing, displaying, and interpreting data using a variety of graphs (pictures, tallies, tables, charts, bar graphs, Venn diagrams), and computing the mean, median, mode, and range of given data. (Standards 8C, 10A, 10B)

- *Probability:* Determining, describing, and applying elementary probability theory and fundamental counting principles. At higher grades, this category encompasses combinations and permutations of simple and complex events. (Standard 10C)

Writing

The state goal for writing states that the student will be able to write standard English in a grammatical, well-organized, and coherent manner for a variety of purposes. The learning standards associated with the goal are as follows:

3A. Use correct grammar, spelling, and punctuation
3B. Compose well-organized and coherent writing
3C. Communicate ideas in writing to accomplish a variety of purposes

The writing assessment uses three types of prompts, which represent persuasive, expository, and narrative discourse modes. Persuasive topics require students to take a position on an issue or to state a problem and solution. Expository topics require students to explain, interpret, or describe something objectively and clearly. Narrative topics require students to reflect upon and describe an experience or event from personal knowledge. Readers evaluate each paper with respect to its focus, support/elaboration, organization, and conventions. They also evaluate how effectively the paper integrates these features.

Students in grades 5, 8, and 10 wrote one assigned essay. All students within a grade received the same assignment. They then selected a second topic (or prompt) from a list of two and wrote a second essay. Third-grade students received one of three topics and wrote an essay on the assigned topic.

Readers score all papers with respect to four specific features (focus, support/elaboration, organization, and conventions) and a holistic feature (integration). Descriptions of these features follow:

- *Focus:* the degree to which the subject, issue, theme, or unifying event of the composition is clear and maintained.

- *Support/Elaboration:* the quality of the detail or support through reasons and explanations.

- *Organization:* the extent to which a clear structure or plan of development is maintained and the points logically related to each other and the text structure.

- *Conventions:* the extent to which the writer demonstrates adequate knowledge of standard English.

- *Integration:* the extent to which the paper as a whole uses the four features (focus, support, organization, and conventions) to address the assignment.

Readers rate a paper's first three features and its overall integration on a scale from 1 (absent) to 6 (well developed). The conventions feature is evaluated as either 1 (not developed) or 2 (developed). A composite writing score is derived from the raw feature scores according to the following formula:

Focus + Support + Organization + Conventions + (2 x Integration)

The overall writing score ranges from 6 to 32. For students who wrote more than one essay (grades 6, 8, 10), writing scores for each essay were averaged and then rounded up. Thus, individual student scores at all grades are reported as whole numbers. Scores for schools, districts, and the state are reported to one decimal place.

Science

Science is a creative endeavor of the human mind. It offers a special perspective on the natural world in terms of understanding and interaction. The Illinois Learning Standards for Science are organized by goals that inform one another and depend upon one another for meaning. Expectations for learners related to the inquiry process are presented in standards addressing the doing of science and elements of technological design.

The ISAT science tests are designed to measure the following three learning standards.

- Goal 11: Understand the process of scientific inquiry and technological design to investigate questions, conduct experiments and solve problems.
  - 11A. Know and apply the concepts, principles and processes of scientific inquiry.
  - 11B. Know and apply the concepts, principles and processes of technological design.

- Goal 12: Understand the fundamental concepts, principles and interconnections of the life, physical and earth/space sciences.
  - 12A. Know and apply concepts that explain how living things function, adapt and change.
  - 12B. Know and apply concepts that describe how living things interact with each other and with their environment.
  - 12C. Know and apply concepts that describe properties of matter and energy and the interactions between them.
  - 12D. Know and apply concepts that describe force and motion and the principles that explain them.

12E. Know and apply concepts that describe the features and processes of the earth and its resources.
12F. Know and apply concepts that explain the composition and structure of the universe and earth's place in it.

- Goal 13: Understand the relationships among science, technology and society in historical and contemporary contexts.
  13A. Know and apply the accepted practices of science.
  13B. Know and apply concepts that describe the interaction between science, technology and society.

The science assessment consists of single-right-answer, multiple-choice items. In addition to an overall score, results are reported in terms of the percent of items correctly answered within five standard sets. These scores are as follows:

- *Scientific Inquiry:* Understanding and applying knowledge of experimental and technological design, including data analysis, use of scientific instruments, and the metric system. (Standards 11A and 11B)

- *Life Sciences:* Understanding and applying knowledge of biology and ecology. (Standards 12A and 12 B)

- *Physical Sciences:* Understanding and applying knowledge of chemistry and physics. (Standards 12C and 12D)

- *Earth and Space Sciences:* Understanding and applying knowledge of geology, weather, renewable resources, astronomy, and space science. (Standards 12E and 12F)

- *Science, Technology, and Society:* Understanding and applying knowledge of safety, valid sources of data, and ethical practices. Understanding and applying knowledge of the history and sociology of science, ethics, environmental issues, and recycling. (Standards 13A and 13B)

A set of science pilot items and a set of health/physical development items used for conducting state studies bring the total number of items in each test to 80. The pilot items do not contribute to test scores.

The Productive Thinking Scale (PTS) is used to evaluate the quality of science items. It is hierarchical with respect to the production of knowledge and independent of an item's difficulty or grade. Four cognitive skills define the hierarchy of productive thinking in generating scientific knowledge. Each skill applies to both content (knowledge) and to process (research methods): (1) recall of conventions, whether names or norms; (2) reproduction of empirical facts or methodological tools and steps; (3) production of solutions to problems or research designs; and (4) creation of new theories and methods. The PTS subdivides reproduction and production into secondary processes. Hence, the PTS comprises six levels of productive thinking on a scale from low level (recall of conventional uses) to high level (creation of new theory).

Based on estimates of the thought processes which most students must use to answer an item, each item is ranked as to the level of conceptual skill it requires. Items that provide a rough balance across the middle ranks are selected, and items at the level of vocabulary or rote memory are usually omitted. Items are also examined to determine whether there is a reasonable distribution within tests of items among major learning areas: earth science, physical science, and life science.

Social Science

Social science provides students with an understanding of themselves and of society, prepares them for citizenship in a democracy, and gives them the basics for understanding the complexities of the world community. The study of social science helps people develop the ability to make informed and reasoned decisions for the public good as citizens of a culturally diverse, democratic society in an interdependent world.

The ISAT social science tests are designed to measure the following learning standards.

- Goal 14: Understand political systems, with an emphasis on the United States.
  - 14A. Understand and explain basic principles of the United States government.
  - 14B. Understand the structures and functions of the political systems of Illinois, the United States and other nations.
  - 14C. Understand election processes and responsibilities of citizens.
  - 14D. Understand the roles and influences of individuals and interest groups in the political systems of Illinois, the United States and other nations.
  - 14E. Understand United States foreign policy as it relates to other nations and international issues.
  - 14F. Understand the development of United States political ideas and traditions.

- Goal 15: Understand economic systems, with an emphasis on the United States.
  - 15A. Understand how different economic systems operate in the exchange, production, distribution and consumption of goods and services.
  - 15B. Understand that scarcity necessitates choices by consumers.
  - 15C. Understand that scarcity necessitates choices by producers.
  - 15D. Understand trade as an exchange of goods or services.
  - 15E. Understand the impact of government policies and decisions on production and consumption in the economy.

- Goal 16: Understand events, trends, individuals and movements shaping the history of Illinois, the United States and other nations.
  - 16A. Apply the skills of historical analysis and interpretation.
  - 16B. Understand the development of significant political events.
  - 16C. Understand the development of economic systems.
  - 16D. Understand Illinois, United States and world social history.
  - 16E. Understand Illinois, United States and world environmental history.

- Goal 17: Understand world geography and the effects of geography on society, with an emphasis on the United States.
  17A. Locate, describe and explain places, regions and features on the Earth.
  17B. Analyze and explain characteristics and interactions of the Earth's physical systems.
  17C. Understand relationships between geographic factors and society.
  17D. Understand the historical significance of geography.

- Goal 18: Understand social systems, with an emphasis on the United States.
  18A. Compare characteristics of culture as reflected in language, literature, the arts, traditions and institutions.
  18B. Understand the roles and interactions of individuals and groups in society.
  18C. Understand how social systems form and develop over time.

The social science assessment consists of single-right-answer, multiple-choice items. In addition to an overall score, results are reported in terms of the percent of items correctly answered within five standard sets. These scores are as follows:

- *Government:* Understanding and applying knowledge of political systems, including the basic principles and traditions of the U.S. government, the structure and functions of government, the election process, and foreign policy. (Standards 14A, 14B, 14D, 14F, and 18B)

- *Economics:* Understanding and applying knowledge of economic systems and the nature of the U.S. economy, including the choices people make in the production and distribution of goods and services and the relationship of governments to trade and economic practices. (Standards 15A, 15B, 15C, 15D, and 15E)

- *Geography:* Demonstrating the ability to locate places, regions, and features; to understand characteristics of Earth's physical system and the relationship between geographic factors and society; and to understand the historical significance of geography. (Standards 17A, 17B, 17C, and 17D)

- *United States History:* Understanding and analyzing the development of political events, economic systems, and social systems. (Standards 16A, 16B, 16C, 16D, 16E, 18A, 18B, and 18C)

- *Global Perspectives:* Understanding and applying knowledge of the political, economic, historical, social, and environmental events and conditions in the world beyond the United States. (Standards 14B, 14E, 16A, 16B, 16C, 16D, 16E, 18A, 18B, and 18C)

2. RELIABILITY

The reliability of a test reflects the degree to which scores are free from random errors of measurement. Test reliability indicates the extent to which differences in test scores reflect real differences in the ability being measured and, thus, the consistency of test scores across some change of condition, such as a change of test items or a change of time. Different reliability coefficients result from different changes in testing conditions. For example, test-retest reliability measures the extent to which scores remain constant over time. A low test-retest reliability coefficient means that a person's scores are likely to shift unpredictably from one time to another.

Internal Consistency of Overall Scores

Because the items used in achievement tests represent only a relatively small sample from a much larger domain of items, the consistency of test scores across items is of particular interest. That is, how precisely will tests rank students if different sets of items from the same domain were used? Unless the rankings are very similar, it is difficult or impossible to make educationally sound decisions on the basis of test scores. This characteristic of test scores is most commonly referred to as *internal consistency*. Table 2.1 presents internal consistency values (coefficient alpha) for each of the tests administered in the 2000 assessment.

Table 2.1
2000 Reliability Estimates

Grade	Reading	Mathematics	Writing	Science	Social Science
03	.94	.93	.86		
04				.92	.92
05	.91	.94	.89		
07				.92	.90
08	.91	.95	.90		
10	.89	.94	.90		

Note: Sample sizes on which these coefficients are based are as follows:

Reading: 3 (15,713), 5 (15,742), 8 (15,718), 10 (15,259)
Mathematics: 3 (15,832), 5 (15,804), 8 (15,716), 10 (15,175)
Writing: 3 (138,312), 5 (141,598), 8 (134,834), 10 (34,678)
Science: 4(15,907), 7(15,856)
Social Science: 4(15,935), 7(15,876)

The reliability coefficients reported in Table 2.1 are derived within the context of classical test theory (CTT) and provide a single measure of precision for the entire test. Within the context of item response theory (IRT), it is possible to measure the relative precision of the test at different points on the scale. Figure 2.1 presents the test information functions for the four ISAT reading tests; Figures 2.2, 2.3, and 2.4 present comparable information for

the ISAT mathematics tests, science tests, and social science tests, respectively. IRT scaling is not used with the writing test.

The amount of information at any point is directly related to the precision of the test. That is, precision is highest where information is highest. Conversely, where information is lowest, precision is lowest and ability is most poorly estimated.

Figure 2.1
2000 ISAT Reading Test Information Functions

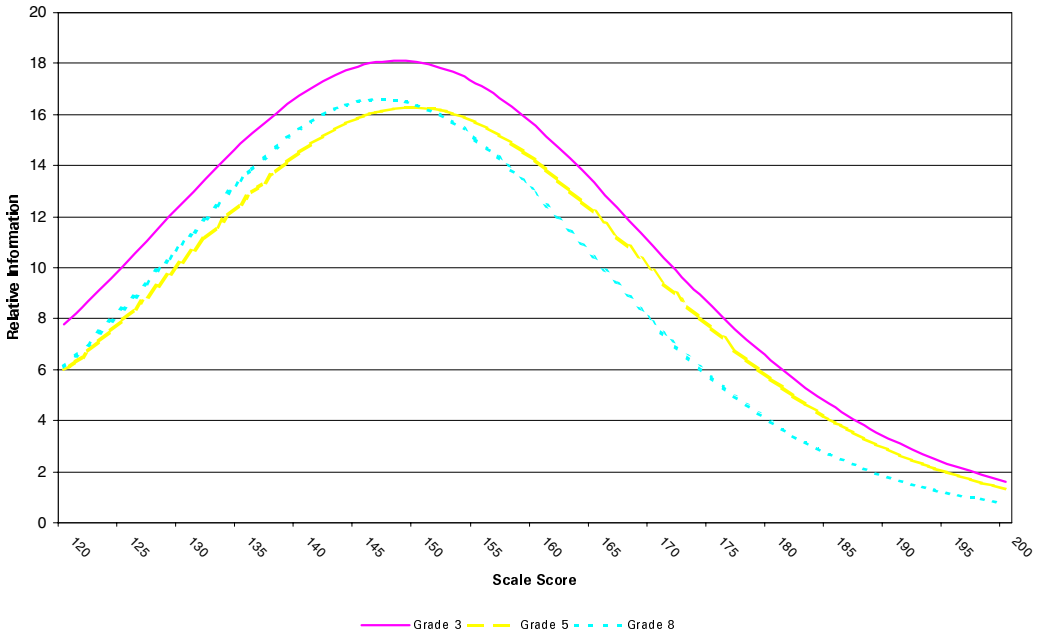


Figure 2.2
2000 ISAT Mathematics Test Information Functions

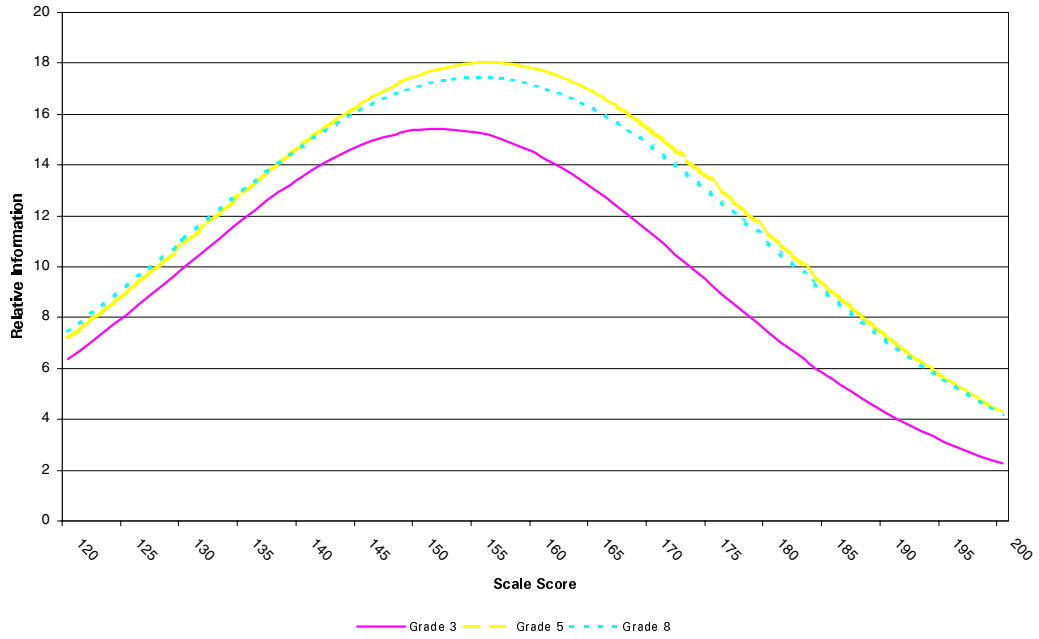


Figure 2.3
2000 ISAT ScienceTest Information Functions

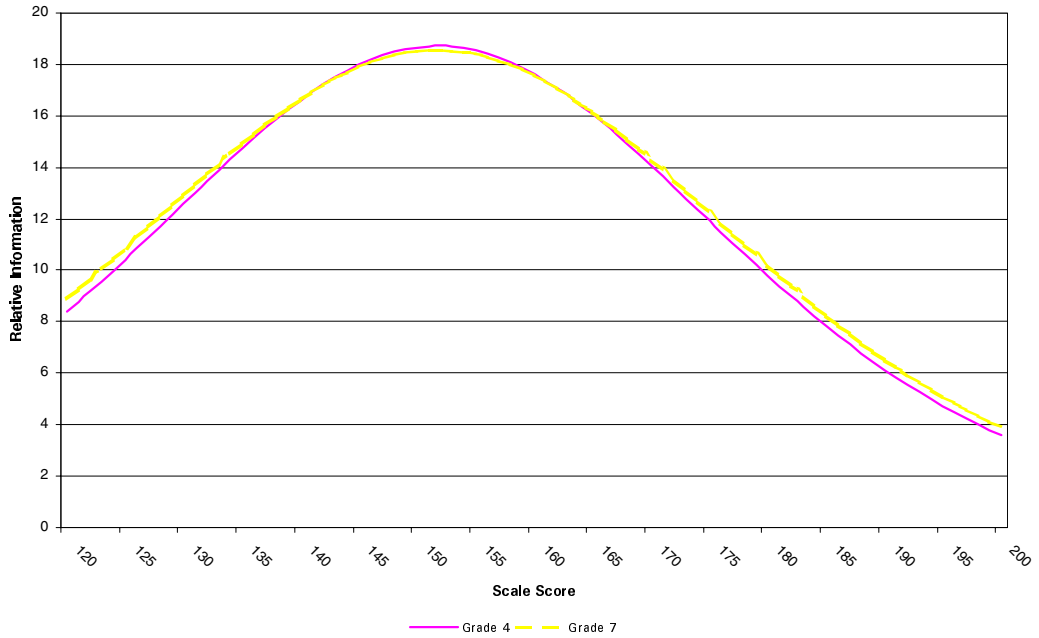
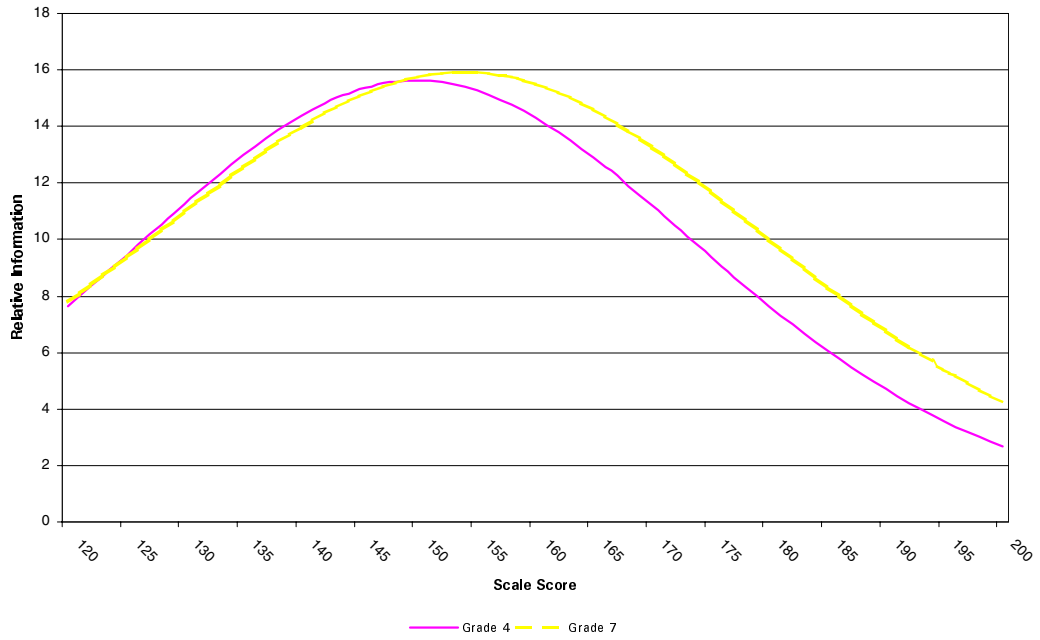


Figure 2.4
2000 ISAT Social ScienceTest Information Functions



A second way of evaluating precision from the IRT perspective is in terms of how well the test as a whole separates people. The ratio of the standard deviation of ability estimates, after subtracting from their observed variance the error variance attributable to their standard errors of measurement, to the root mean square standard error computed over persons provides this index (Wright & Stone, 1979). These values are reported in Table 2.2.

Table 2.2
Person Separation Values for the 2000 ISAT Tests

	Reading	Mathematics
Grade 3	3.35	3.48
Grade 5	2.99	3.81
Grade 8	2.80	4.05
Grade 10	2.37	3.88
	Science	Social Science
Grade 4	3.26	3.24
Grade 7	3.18	2.93

Reliability of the Writing Scores

Writing scores are affected by other sources of variance, particularly readers (raters), since different readers evaluate different students, and prompts. The effect attributable to

prompts is important for students at all grades. However, it can only be evaluated directly for 5th-, 8th- and 10th-grade students who wrote on two different prompts.

Interrater Agreement. Interrater agreement evaluates the consistency of scores assigned to the same essay by different readers. For the 2000 writing assessment, interrater agreement was monitored daily, and two readers independently scored 10% of the student essays across grades and prompts. The interrater agreement coefficients for all features and discourse modes are summarized in Table 2.3. The results for the interrater agreement on double-scored papers exceeded the minimum acceptable level of agreement (90% agreement within one point). Scores across raters agree within one point at least 92% of the time.

Table 2.3
Interrater Agreement for Writing Scores

Discourse Mode	Score	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent
Persuasive	Focus	59	36	95
	Support	53	42	95
(n = 23,365)	Organization	54	41	95
	Conventions	94	6	100
	Integration	57	40	96
Expository	Focus	58	36	94
	Support	54	42	95
(n = 15,879)	Organization	53	42	95
	Conventions	93	7	100
	Integration	56	40	96
Narrative	Focus	53	39	92
	Support	56	40	96
(n = 28,936)	Organization	53	39	92
	Conventions	92	8	100
	Integration	55	39	94

In addition to agreement across raters, writing scores are checked against a standard or "validation" set of papers. The Validation Committee assigns the scores for these papers. Essay packets, each containing 10 essays, were circulated among the readers. Essays for these check sets were chosen to represent a range of score points in all categories.

Readers encountered the validation packets at random intervals throughout the scoring, and some encountered several packets during the scoring process. Readers were unaware of the scores assigned to the papers by the committee. The extent of agreement between a reader's scores and the scores assigned to the papers was calculated every day during the scoring and shared with the readers. This process allowed for the monitoring of reader scoring. The results for all grades, features and discourse modes are summarized in Table 2.4. Again, the results exceeded the minimum acceptable level of agreement (90% agreement within one point). The agreement of readers with validation papers was higher than the interrater agreement. This is possibly attributable to the fact that the validation

papers are specifically selected to illustrate all points on the scoring scale. The papers that are selected for double scoring, on the other hand, represent a more nearly random selection of papers and scores. Consequently, they are likely to include proportionately fewer extreme scores (e.g., 1, 6), on which there is likely to be higher agreement between raters.

Table 2.4
Agreement with Validation Papers for Writing Scores

Discourse Mode	Score	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent
Grade 3				
Narrative	Focus	71	25	96
	Support	66	32	98
	Organization	67	31	98
(N = 810)	Conventions	94	6	100
	Integration	72	25	97
Persuasive	Focus	83	15	98
/Expository	Support	73	26	99
(N = 1,740)	Organization	69	30	99
	Conventions	99	1	100
	Integration	74	26	100
Grade 5				
Narrative	Focus	71	25	96
	Support	71	28	99
	Organization	69	29	98
(N = 1,900)	Conventions	95	5	100
	Integration	72	25	97
Persuasive	Focus	60	35	95
/Expository	Support	61	35	96
(N = 2,670)	Organization	53	40	93
	Conventions	90	10	100
	Integration	58	40	98

Table 2.4 (continued)

Discourse Mode	Score	% Exact Agreement	% Adjacent Agreement	% Exact + Adjacent
Grade 8				
Narrative	Focus	75	21	96
	Support	79	21	100
	Organization	75	23	98
(N = 1,700)	Conventions	93	7	100
	Integration	78	21	99
Persuasive	Focus	76	23	99
/Expository	Support	64	35	99
(N = 1,830)	Organization	66	33	99
	Conventions	92	8	100
	Integration	74	26	100
Grade 10				
Narrative	Focus	74	23	97
	Support	71	28	99
	Organization	71	26	97
(N = 1,150)	Conventions	98	2	100
	Integration	75	24	99
Persuasive	Focus	73	26	99
/Expository	Support	80	19	99
(N = 1,790)	Organization	77	23	100
	Conventions	91	9	100
	Integration	79	21	100

Reliability of the Performance Category Decisions

Students' ISAT scores are reported relative to four performance categories: Academic Warning, Below Standards, Meets Standards, and Exceeds Standards. Sets of score cutoffs were developed for each learning area and each grade. The development of the score cutoffs that define these categories is fully documented in separate publications available from ISBE (*Performance Levels for the Illinois Standards Achievement Tests: Reading, Mathematics, Writing* and *Performance Levels for the Illinois Standards Achievement Tests: Science, Social Science*). However, the process may be briefly described as follows.

Prior to the meetings of the standard-setting panels themselves, which took place during April 1999 (reading, mathematics, writing) and April 2000 (science, social science), ISBE convened committees of curriculum experts to develop concrete descriptions of student

knowledge and skill levels that define the specific performance categories. Educators throughout Illinois extensively reviewed these descriptions.

Panels of recognized subject matter experts convened in Springfield to translate the verbal descriptions into cut scores on the ISAT tests (i.e., scores that define the boundaries between categories). Panelists were drawn from a pool of educators who had specific knowledge of student performance at the grade levels being assessed by ISAT and experience in assessing students at those grade levels. Panelists were selected to be broadly representative of the geographic and ethnic diversity of Illinois' public school system. A total of 170 educators participated in the standard-setting process. The distribution of educators across learning areas was as follows: mathematics—56; writing—62; reading—52; science–30; social science–30.

A procedure originally proposed by Angoff is one of the most frequently used methods for determining cut scores when multiple-choice test scores are used. It can be most simply described as a focused, judgmental process by knowledgeable content experts. The basic Angoff procedure fit the format of the ISAT reading, mathematics, science, and social science tests. However, certain modifications of the basic procedure were developed to fit the format of the ISAT writing tests.

In the most frequent application of the Angoff method (e.g., to establish a pass-fail standard), panelists are asked to examine an item and decide what proportion of minimally competent individuals will answer the question correctly. With respect to the ISAT, however, instead of being asked about minimally competent students, panelists were asked to indicate what percentage of three groups of students—those who were just above the Academic Warning/Below Standards boundary, those who were just above the Below Standards/Meets Standards boundary, and those who were just above the Meets Standards/Exceeds Standards boundary—would answer the question correctly. The ratings were made sequentially rather than simultaneously (i.e., panelists made all judgments relative to one cut score before moving to the next cut score). Item performance statistics were provided to help panelists anchor their ratings. The cutoff scores that resulted are shown in Table 2.5. Results of applying these cutoffs to the 2000 test population are shown later in Section 5.

The reliabilities of such classifications, which are criterion-referenced, are related to the reliabilities of the tests on which they are based, but they are not equivalent to the test reliabilities, which are based on norm-referenced measurement. Glaser (1963) was among the first to draw attention to this distinction, and Feldt and Brennan (1989) extensively reviewed the topic.

Table 2.5
ISAT Cutoffs for Each Performance Level

READING	Academic Warning	Below Standards	Meets Standards	Exceeds Standards
03	120-137	138-155	156-173	174-200
05	120-129	130-155	156-170	171-200
08	120-128	129-151	152-172	173-200
10	120-135	136-152	153-174	175-200
MATHEMATICS	**Academic Warning**	**Below Standards**	**Meets Standards**	**Exceeds Standards**
03	120-141	142-152	153-172	173-200
05	120-137	138-157	158-190	191-200
08	120-137	138-161	162-184	185-200
10	120-138	139-157	158-187	188-200
WRITING	**Academic Warning**	**Below Standards**	**Meets Standards**	**Exceeds Standards**
03	6-13	14-21	22-29	30-32
05	6-13	14-20	21-27	28-32
08	6-14	15-20	21-27	28-32
10	6-14	15-20	21-27	28-32
SCIENCE	**Academic Warning**	**Below Standards**	**Meets Standards**	**Exceeds Standards**
04	120-129	130-153	154-178	179-200
07	120-141	142-150	151-174	175-200
SOCIAL SCIENCE	**Academic Warning**	**Below Standards**	**Meets Standards**	**Exceeds Standards**
04	120-141	142-156	157-183	184-200
07	120-132	133-156	157-178	179-200

As Feldt and Brennan (1989, p. 140) point out, approaches to the development of reliability coefficients for criterion-referenced interpretations of test scores have been based either on squared-error loss or threshold loss. It is threshold loss, which evaluates the consistency with which people are consistently classified with respect to a criterion, that is of greater concern here. Specifically, the issue is how consistently do tests classify students with respect to the performance standards?

Two threshold-loss coefficients have been developed: *p*, the proportion of persons consistently classified on two parallel tests, and *k* (kappa), which corrects *p* for the proportion of consistent classifications that would be expected by chance. Because scores on classically parallel tests are rarely available in practice, methods have been developed to estimate these values from a single test (Subkoviak, 1984). An approach proposed by Peng and Subkoviak (1980) was applied to the performance classifications made on the basis of the 2000 tests.

Table 2.6 presents the 2000 values for *p, k,* and $p_{miss,}$ the expected proportion of inconsistent decisions, which is simply (1 - *p*). In interpreting the first two indexes, Feldt and Brennan (1989) suggest that *p* reflects the *consistency of decisions* made about examinees, whereas *k*, since it is corrected for chance, reflects the *contribution of the test* to the consistency of the decision.

Overall, the values suggest that decisions made with respect to the student performance classifications would be very consistent. Note that the *p* and *k* values are calculated for the complete test population. Values for other test populations (e.g., IEP students alone, non-IEP students only) may differ.

Table 2.6
Reliability of Student Performance Decisions Based on 2000 Test Scores

		Academic Warning/Below Standards			Below Standards/Meets Standards			Meets Standards/Exceeds Standards		
Area	Grade	p	kappa	p_{miss}	p	kappa	p_{miss}	p	kappa	p_{miss}
Reading	3	.972	.731	.028	.906	.792	.094	.928	.785	.072
	5	.986	.526	.014	.864	.712	.136	.902	.691	.098
	8	.986	.454	.014	.880	.698	.120	.914	.678	.086
	10	.976	.613	.024	.878	.704	.122	.926	.662	.074
Mathematics	3	.958	.741	.042	.916	.789	.084	.924	.783	.076
	5	.970	.738	.030	.926	.779	.074	.922	.787	.078
	8	.964	.737	.036	.920	.782	.080	.928	.785	.072
	10	.972	.731	.028	.924	.783	.076	.978	.714	.022
Writing	3	.966	.510	.034	.824	.648	.176	.966	.510	.034
	5	.990	.578	.010	.880	.698	.120	.910	.683	.090
	8	.984	.611	.016	.880	.698	.120	.918	.673	.082
	10	.984	.611	.016	.884	.697	.116	.914	.678	.086
Science	4	.976	.466	.024	.872	.709	.128	.930	.656	.070
	7	.934	.652	.066	.884	.697	.116	.910	.683	.090
Social Science	4	.934	.652	.066	.862	.713	.138	.958	.633	.042
	7	.976	.613	.024	.862	.713	.138	.930	.656	.070
AVERAGE		.955	.596	.045	.870	.708	.130	.932	.657	.068

3. SCALING AND EQUATING PROCEDURES

ISAT reading, mathematics, science and social science scores are reported on a standard score scale. Individual student scores on this scale range between 120 and 200, regardless of the characteristics of the raw score distribution. Each scale is defined by letting 160 represent the average proficiency of the first-year test population. Every unit on the scale represents 1/15 of the standard deviation of proficiency scores for the first-year population. In other words, the first year mean and standard deviation of scale scores for each grade are 160 and 15. Results in subsequent years are equated to the base-year scale. The scaling constants used to transform the Rasch proficiency estimates to the reporting scale are shown in Table 3.1.

Table 3.1
ISAT Scaling Constants

Reading	Slope	Intercept
Grade 3	12.6428	146.2066
Grade 5	12.0100	144.7660
Grade 8	11.2280	141.7730
Grade 10	12.1470	140.1670
Mathematics		
Grade 3	13.5122	147.6910
Grade 5	14.9686	153.4644
Grade 8	14.7578	146.7806
Grade 10	13.0795	148.3945
Science		
Grade 4	15.3781	152.4255
Grade 7	15.9209	152.4527
Social Science		
Grade 4	14.6746	149.2394
Grade 7	16.6587	148.9095

Because test items change each year, raw scores (i.e., number or percent correct scores) will not always have the same meaning or represent the same level of proficiency. Without equating, each administration of a test with different items would lead to a new reporting scale, independent of that used previously. It would still be possible to measure relative performance, but it would not be possible to indicate growth across years for schools, districts, or the state. The equating process makes longitudinal comparisons possible.

The statistical fit of the one-parameter logistic (1PL) or Rasch model to the ISAT multiple-choice tests has been previously examined and found to be satisfactory. The 1PL model uses only the item difficulty and the person's proficiency level to describe the probability of a correct response to an item. The 1PL model is the simplest of currently available IRT models and is perhaps the one in widest use today.

The equating procedures may be summarized as follows. Each test contains a sufficient number of items that have been previously administered to provide a reliable and content-

representative equating link. During calibration of the new tests, item difficulties for these linking items are set to their historical values. By estimating values for the remaining items under this constraint, difficulty values for the remaining items are expressed on the existing scale. That is, the proficiency (theta) scale that results from the constrained calibration run is equated to the existing scale. The final step in the procedure is to apply equations that transform values on the proficiency scale to their corresponding ISAT scale score values. These equations were originally developed during the first year of equating and are then applied in each subsequent year of equating.

The logic of the equating procedure rests on certain assumptions. The most important is that the items used for linking stay the same in the two tests. During the assembly of tests, items that will be used for equating are placed exactly at or very near the location in the booklet where they previously appeared (i.e., item 23 in 1998 is also item 23 in 1999) to minimize effects from positional differences. Differences between the anchored difficulties and the best-fit values are examined to ensure that no unusually large differences exist that would strain the equivalence assumption.

The equating analyses are conducted on samples of approximately 16,000 drawn from the total test population. A 1/nth selection results in a sample that has characteristics essentially identical with that of the total population.

This approach places both sets of tests on a firm basis to meet future equating needs. Successive years' test forms, which will have different items, will be equated so that test scores will remain comparable across administrations. Each new test will contain a sufficient number of items that have been previously administered to provide a reliable and content-representative equating link. During calibration of the new tests, item difficulties for these linking items will be set to their historical values. By estimating values for the remaining items under this constraint, difficulty values for the remaining items will be automatically adjusted to the existing scale. The final step in the procedure is to apply equations that transform values on the proficiency scale to their corresponding scale score values. These equations are developed during the first year of testing and are then applied in each subsequent year.

ISAT also uses two forms of the reading test. At each grade, two passages (and their associated items) are identical across the two forms and one passage is different. Because the two tests are not exactly equal in difficulty, scores on the two forms are statistically equated using the one-parameter (Rasch) model. The two forms were jointly calibrated, which places the difficulty of both sets of items on the same scale and makes proficiency estimates equivalent across test forms. IRT scaling is also used with the ISAT mathematics tests.

Tables 3.2 through 3.5 show results of the Rasch calibration and equating procedures for reading. Column 1 of each table shows the Form in which the item appeared, A for Form A alone, B for Form B alone, AB for items that appeared in both forms. Column 2 of each table shows the item number within the test booklet. Column 3 shows the Rasch difficulties resulting from an anchored (constrained) calibration of the 2000 test. Column 4 shows the standard error of the difficulty estimate (S_{ed}). The next two columns present statistics designed to assess how well the test "fits" the IRT model. Both are standardized, mean-

square statistics with an expected value of 1.00 (indicating perfect fit). The first, “Infit,” is more sensitive to departures from model fit when item difficulty and person ability are close. The second, “Outfit,” is more sensitive to model fit when item difficulty and person ability are far apart. The last column shows the point-biserial correlation between the item and the rest of the items in the test.

Tables 3.6 through 3.9 show similar information for the mathematics tests. The information is organized in the same way as the earlier tables except no “form” designation is necessary. Tables 3.10 and 3.11 present information for the science tests, and Tables 3.12 and 3.13 present information for the social science tests.

Table 3.2
Results of the 2000 Equating Process–Reading Grade 3

Form	Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
AB	1	-1.23	.03	.97	.83	.34
AB	2	.57	.02	1.09	1.10	.36
AB	3	.60	.02	1.07	1.07	.38
AB	4	.15	.02	1.09	1.17	.33
AB	5	-.78	.02	.99	.94	.36
AB	6	-.10	.02	1.03	1.04	.38
AB	7	-.48	.02	1.00	.99	.38
AB	8	-.33	.02	.94	.90	.44
AB	9	-.92	.02	1.01	1.04	.32
AB	10	.66	.02	1.16	1.21	.31
AB	11	.90	.02	1.15	1.18	.32
AB	12	-.70	.02	.89	.79	.47
AB	13	-.09	.02	1.20	1.24	.23
AB	14	.33	.02	1.17	1.22	.28
A	15	-.12	.03	.91	.82	.47
A	16	.54	.03	1.12	1.16	.35
A	17	.20	.03	.94	.93	.49
A	18	-.07	.03	1.08	1.22	.35
A	19	-.78	.03	.88	.85	.45
A	20	-1.24	.04	.72	.43	.50
A	21	-1.05	.03	.95	.79	.44
A	22	-.53	.03	.90	.77	.50
A	23	-.34	.03	.96	.87	.47
A	24	.76	.03	.95	.92	.48
A	25	.91	.03	1.32	1.51	.19
A	26	.16	.03	.87	.76	.53
A	27	-.30	.03	.88	.78	.50
A	28	-.38	.03	.78	.62	.55
AB	29	-1.09	.02	.92	.84	.40
AB	30	.06	.02	.88	.79	.53
AB	31	.89	.02	1.02	1.00	.42
AB	32	.25	.02	.93	.86	.49
AB	33	-.06	.02	.99	.99	.42
AB	34	1.52	.02	1.22	1.36	.26
AB	35	1.15	.02	1.14	1.17	.32
AB	36	.71	.02	.98	.93	.45
AB	37	-1.78	.03	.85	.45	.43
AB	38	.29	.02	.88	.79	.53
AB	39	1.33	.02	.99	1.01	.42
AB	40	.38	.02	1.01	.99	.42
AB	41	-.36	.02	.86	.69	.53
AB	42	1.50	.02	1.03	1.10	.39
AB	43	1.27	.02	1.14	1.19	.32
AB	44	.50	.02	.85	.77	.56
AB	45	1.17	.02	.97	.96	.45
AB	46	1.00	.02	1.00	1.00	.44
AB	47	1.37	.02	1.23	1.34	.26
AB	48	.10	.02	.95	.86	.47
AB	49	-.50	.02	.99	1.07	.38

Table 3.2 (continued)

AB	50	.99	.02	1.00	1.00	.44
AB	51	-.81	.02	.87	.82	.47
AB	52	1.17	.02	1.36	1.52	.16
AB	53	-.74	.02	.86	.67	.50
AB	54	-.41	.02	.88	.73	.50
AB	55	.96	.02	1.12	1.16	.34
AB	56	.44	.02	.89	.80	.53
AB	57	.10	.02	.85	.72	.56
AB	58	.15	.02	.95	.92	.46
AB	59	.26	.02	1.02	1.14	.41
AB	60	.39	.02	.88	.79	.54
AB	61	.73	.02	1.08	1.08	.37
AB	62	1.22	.02	1.18	1.26	.29
AB	63	-.50	.02	.87	.72	.51
AB	64	1.29	.02	1.08	1.14	.35
AB	65	.12	.02	.89	.81	.52
AB	66	.08	.02	.89	.82	.52
AB	67	.47	.02	1.03	1.04	.41
B	15	-.25	.03	1.05	1.18	.34
B	16	-.71	.03	.91	.89	.43
B	17	.92	.03	.98	.98	.46
B	18	-.66	.03	.94	1.07	.41
B	19	.85	.03	.99	.98	.45
B	20	.63	.03	.97	.91	.46
B	21	-.62	.03	.99	1.13	.36
B	22	.42	.03	1.01	1.01	.42
B	23	1.33	.03	1.03	1.06	.40
B	24	.60	.03	.85	.78	.57
B	25	.12	.03	.95	.89	.46
B	26	-.31	.03	.88	.72	.51
B	27	.25	.03	.97	.89	.45
B	28	-.63	.03	.94	.91	.42

Table 3.3
Results of the 2000 Equating Process–Reading Grade 5

Form	Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
AB	1	-.11	.02	.99	.95	.37
AB	2	-.40	.02	1.16	1.39	.16
AB	3	2.08	.02	1.07	1.15	.31
AB	4	.99	.02	1.13	1.18	.29
AB	5	.42	.02	.85	.75	.54
AB	6	.39	.02	.94	.88	.45
AB	7	-.34	.02	.97	.93	.38
AB	8	.79	.02	1.06	1.05	.35
AB	9	.57	.02	.96	.88	.44
AB	10	.04	.02	.96	.89	.42
AB	11	1.09	.02	1.04	1.05	.37
AB	12	.64	.02	1.06	1.08	.34
AB	13	.08	.02	1.00	.94	.38
AB	14	1.26	.02	1.05	1.05	.36
AB	15	.86	.02	.94	.90	.46
AB	16	1.46	.02	.98	1.00	.40
AB	17	1.87	.02	1.11	1.22	.27
AB	18	.63	.02	.98	.96	.41
A	20	.81	.03	1.28	1.40	.14
A	21	.70	.03	.93	.89	.46
A	22	.49	.03	.95	.90	.45
A	23	.46	.03	.96	.93	.44
A	24	-.07	.03	.87	.78	.49
A	25	.96	.03	.96	.95	.45
A	26	-.19	.03	.93	.83	.43
A	27	.59	.03	.98	.99	.42
A	28	-.93	.03	.91	.79	.39
A	29	-.58	.03	.86	.74	.47
A	30	.68	.03	.92	.87	.48
A	31	.88	.03	1.16	1.29	.26
A	32	-.20	.03	.89	.82	.47
A	33	.16	.03	1.00	1.02	.37
A	34	-.16	.03	.88	.81	.48
A	35	-.07	.03	.92	.89	.44
A	36	1.10	.03	.91	.90	.48
A	37	2.09	.03	1.11	1.35	.24
AB	38	-1.47	.03	.91	.77	.35
AB	39	1.92	.02	.96	1.02	.41
AB	40	1.25	.02	1.09	1.13	.32
AB	41	-.75	.02	.91	.79	.41
AB	42	-.83	.02	.95	.86	.37
AB	43	1.26	.02	1.12	1.15	.30
AB	44	.09	.02	1.03	1.13	.34
AB	45	1.24	.02	.99	.99	.41
AB	46	1.01	.02	1.10	1.17	.32
AB	47	.67	.02	.97	.92	.43
AB	48	.47	.02	.95	.93	.44
AB	49	-.11	.02	.97	.91	.39
AB	50	.20	.02	.85	.76	.53

Table 3.3 (continued)

AB	51	.52	.02	.96	.95	.44
AB	52	.97	.02	1.27	1.45	.17
AB	53	1.60	.02	1.13	1.19	.29
AB	54	-.81	.02	.90	.76	.42
AB	55	.07	.02	1.00	1.02	.38
AB	56	.93	.02	1.11	1.14	.31
B	20	-.47	.03	1.11	1.19	.28
B	21	-.98	.04	.93	.85	.40
B	22	1.37	.03	1.05	1.07	.37
B	23	.94	.03	.95	.95	.44
B	24	-.27	.03	.88	.79	.44
B	25	.81	.03	.96	.94	.44
B	26	-.55	.03	.82	.64	.51
B	27	.02	.03	1.03	.97	.39
B	28	1.00	.03	1.01	1.00	.39
B	29	-.27	.03	.81	.67	.52
B	30	-1.08	.04	1.01	.88	.43
B	31	.42	.03	.91	.88	.48
B	32	.60	.03	1.11	1.17	.29
B	33	.29	.03	.97	.95	.39
B	34	.84	.03	1.16	1.24	.27
B	35	.58	.03	.98	.96	.41
B	36	.35	.03	.85	.78	.53
B	37	.05	.03	.98	.95	.36

Table 3.4
Results of the 2000 Equating Process–Reading Grade 8

Form	Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
AB	1	.87	.02	1.14	1.22	.27
AB	2	.59	.02	.96	.88	.44
AB	3	-.18	.02	.96	.99	.39
AB	4	1.71	.02	1.20	1.32	.22
AB	5	1.54	.02	1.08	1.11	.33
AB	6	.23	.02	.89	.78	.49
AB	7	1.27	.02	1.03	1.03	.38
AB	8	-.10	.02	1.01	1.07	.34
AB	9	-1.38	.03	.87	.60	.40
AB	10	-.02	.02	.95	.89	.41
AB	11	-.18	.02	.99	.98	.35
AB	12	.17	.02	1.01	1.07	.36
AB	13	.18	.02	.94	.90	.44
AB	14	.17	.02	.97	.91	.41
AB	15	1.17	.02	1.02	1.01	.38
AB	16	.34	.02	.94	.85	.45
AB	17	-.10	.02	1.09	1.26	.26
AB	18	1.14	.02	.96	.93	.43
A	20	-.70	.04	1.16	1.29	.23
A	21	.75	.03	1.02	1.00	.34
A	22	1.18	.03	1.07	1.10	.30
A	23	.24	.03	.90	.78	.43
A	24	.34	.03	.97	.91	.45
A	25	.97	.03	1.01	1.03	.40
A	26	.86	.03	.99	.99	.40
A	27	1.09	.03	1.13	1.15	.28
A	28	.96	.03	1.22	1.38	.20
A	29	1.81	.03	1.02	1.06	.37
A	30	.38	.03	.95	.91	.40
A	31	-.30	.03	.94	.83	.46
A	32	1.09	.03	.99	.95	.41
A	33	.67	.03	1.04	1.00	.39
A	34	.78	.03	.93	.89	.45
A	35	.00	.03	.90	.73	.56
A	36	-.77	.04	.90	.68	.48
A	37	.88	.03	.99	1.00	.43
AB	38	.98	.02	1.13	1.20	.29
AB	39	.17	.02	1.00	.92	.38
AB	40	1.00	.02	1.00	.98	.41
AB	41	-.07	.02	1.00	.97	.36
AB	42	1.16	.02	.92	.87	.47
AB	43	.50	.02	1.00	.97	.39
AB	44	.08	.02	.90	.81	.47
AB	45	1.47	.02	1.18	1.23	.24
AB	46	.71	.02	.96	.92	.43
AB	47	.00	.02	1.02	.97	.35
AB	48	.79	.02	1.05	1.06	.35
AB	49	.80	.02	1.04	1.05	.36
AB	50	.44	.02	.98	.92	.41

Table 3.4 (continued)

AB	51	-.32	.02	.92	.78	.44
AB	52	1.46	.02	1.09	1.11	.31
AB	53	.83	.02	1.07	1.08	.33
AB	54	-.23	.02	.90	.81	.45
AB	55	-.22	.02	.85	.69	.51
AB	56	.41	.02	1.01	.97	.38
B	20	1.50	.03	1.03	1.05	.37
B	21	-.13	.03	.97	.98	.39
B	22	-.23	.03	.92	.84	.44
B	23	-1.12	.04	.85	.67	.44
B	24	.70	.03	1.11	1.17	.30
B	25	-.53	.03	1.01	1.15	.31
B	26	1.03	.03	1.34	1.54	.11
B	27	1.06	.03	1.09	1.10	.32
B	28	.13	.03	.95	.87	.43
B	29	1.15	.03	1.07	1.10	.34
B	30	.56	.03	.95	1.02	.45
B	31	.38	.03	.88	.77	.51
B	32	2.04	.03	1.11	1.17	.29
B	33	.82	.03	.91	.83	.49
B	34	-.11	.03	.90	.79	.46
B	35	-.41	.03	.96	.87	.38
B	36	.67	.03	.84	.76	.55
B	37	.36	.03	.94	.85	.45

Table 3.5
Results of the 2000 Equating Process–Reading Grade 10

Form	Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
AB	1	1.39	.02	1.05	1.09	.30
AB	2	.37	.02	1.02	.98	.35
AB	3	-1.84	.04	.89	.57	.37
AB	4	-1.19	.03	.92	.73	.38
AB	5	.53	.02	1.04	1.00	.32
AB	6	-1.16	.03	.94	.78	.35
AB	7	1.35	.02	1.27	1.38	.11
AB	8	.83	.02	1.13	1.20	.24
AB	9	2.02	.02	.92	.95	.40
AB	10	-2.00	.04	.89	.62	.35
AB	11	-2.52	.05	.89	.51	.33
AB	12	-1.01	.03	.99	.98	.31
AB	13	.92	.02	1.03	1.03	.34
AB	14	.05	.02	1.01	1.02	.35
AB	15	.80	.02	1.00	.98	.37
AB	16	-1.11	.03	.97	.86	.33
AB	17	.98	.02	1.08	1.10	.29
AB	18	.34	.02	1.09	1.17	.27
AB	19	.89	.02	1.07	1.06	.30
A	20	-1.30	.05	.82	.54	.42
A	21	.36	.03	1.14	1.16	.28
A	22	.78	.03	1.01	.99	.39
A	23	.54	.03	1.05	1.10	.33
A	24	-.30	.03	1.11	1.11	.30
A	25	-1.12	.04	.97	.67	.47
A	26	.19	.03	.95	.87	.45
A	27	-.60	.04	.85	.61	.50
A	28	-.10	.03	1.01	.94	.42
A	29	-.15	.03	.93	.82	.43
A	30	-.13	.03	1.07	1.22	.34
A	31	.82	.03	.99	.98	.39
A	32	1.30	.03	.99	1.01	.36
A	33	-.54	.04	.87	.69	.48
A	34	-.17	.03	.82	.70	.47
A	35	-.14	.03	.79	.67	.47
A	36	.78	.03	.99	.96	.38
AB	38	-.61	.03	.92	.81	.41
AB	39	.79	.02	1.00	.99	.38
AB	40	-.34	.02	.89	.71	.47
AB	41	.50	.02	1.02	1.08	.35
AB	42	-.85	.03	.88	.70	.45
AB	43	-1.33	.03	.85	.58	.45
AB	44	.81	.02	1.13	1.19	.25
AB	45	-.54	.03	.90	.77	.44
AB	46	-.42	.03	.94	.92	.39
AB	47	1.18	.02	1.14	1.19	.23
AB	48	.53	.02	1.07	1.07	.30
AB	49	-.52	.03	.91	.81	.43
AB	50	1.01	.02	.98	.94	.39

Table 3.5 (continued)

AB	51	1.33	.02	1.00	1.01	.35
AB	52	-.26	.02	.89	.77	.46
AB	53	-.45	.03	.82	.59	.54
AB	54	1.00	.02	1.21	1.38	.18
AB	55	.41	.02	1.03	1.05	.34
B	20	-.18	.03	1.01	1.04	.32
B	21	.28	.03	1.09	1.23	.25
B	22	-1.73	.06	.94	.78	.31
B	23	.59	.03	1.24	1.41	.11
B	24	1.11	.03	1.03	1.03	.31
B	25	1.52	.03	1.13	1.20	.21
B	26	-.88	.04	.92	.83	.39
B	27	-.62	.04	.95	.88	.37
B	28	.29	.03	.96	.88	.40
B	29	-.14	.03	.98	.89	.36
B	30	1.35	.03	.96	.97	.38
B	31	-.30	.03	.92	.78	.43
B	32	-.18	.03	.96	.89	.38
B	33	1.88	.03	1.01	1.05	.33
B	34	-.83	.04	.95	.88	.36
B	35	-1.25	.05	.87	.59	.43
B	36	1.47	.03	.95	.96	.39

Table 3.6
Results of the 2000 Equating Process–Mathematics Grade 3

Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
1	-.44	.02	.91	.83	.41
2	-2.09	.03	.98	.94	.24
3	.77	.02	1.01	.99	.41
4	-.27	.02	1.00	.93	.40
5	-.90	.02	1.04	.92	.47
6	-.66	.02	1.03	1.15	.31
7	-.70	.02	.93	.84	.42
8	-.29	.02	.98	.96	.40
9	-.10	.02	.78	.67	.60
10	.76	.02	1.17	1.21	.27
11	-.36	.02	.83	.81	.53
12	1.17	.02	1.01	1.03	.40
13	.81	.02	1.26	1.39	.19
14	-.10	.02	.89	.84	.49
15	-.31	.02	.99	1.01	.36
16	.15	.02	1.25	1.43	.26
17	.10	.02	.99	.98	.38
18	.58	.02	.95	.92	.46
19	1.05	.02	1.00	1.02	.42
20	.63	.02	1.03	1.07	.39
21	-1.13	.02	.94	.92	.36
22	1.56	.02	.99	.99	.42
23	.33	.02	.95	.90	.46
24	.57	.02	1.08	1.12	.34
25	1.46	.02	.97	.98	.43
26	1.21	.02	.98	1.00	.43
27	1.33	.02	1.01	1.05	.39
28	-.59	.02	1.05	1.12	.32
29	1.42	.02	1.16	1.22	.28
30	.14	.02	.90	.85	.47
31	-.98	.02	.95	.90	.37
32	1.57	.02	.91	.95	.47
33	.33	.02	1.12	1.20	.38
34	.97	.02	1.09	1.11	.34
35	.97	.02	.96	.95	.46
36	.98	.02	.94	.92	.47
37	.58	.02	.94	.94	.47
38	-.69	.02	.85	.74	.45
39	.51	.02	.87	.82	.54
40	1.23	.02	1.14	1.19	.29
41	1.83	.02	.97	1.08	.39
42	1.30	.02	.99	1.01	.41
43	-.71	.02	.96	.90	.39
44	-1.02	.02	.90	.72	.43
45	.74	.02	1.02	1.03	.40
46	.39	.02	.87	.80	.52
47	.36	.02	.94	.95	.47
48	.31	.02	.85	.80	.55
49	1.29	.02	1.06	1.09	.36

Table 3.6 (continued)

50	-1.11	.02	.94	.84	.38
51	1.18	.02	1.11	1.14	.32
52	.35	.02	1.13	1.26	.35
53	-.24	.02	.98	.94	.40
54	.49	.02	1.00	1.00	.42
55	.75	.02	.94	.92	.46
56	1.08	.02	1.25	1.35	.21
57	.05	.02	.86	.78	.54
58	.42	.02	1.03	1.01	.40
59	.80	.02	1.09	1.12	.33
60	-.30	.02	1.09	1.17	.29
61	.95	.02	1.16	1.22	.28
62	.12	.02	.81	.79	.55
63	-.65	.02	.95	.90	.40
64	-.84	.02	.77	.61	.36
65	.13	.02	1.01	.98	.39
66	.33	.02	1.11	1.14	.27
67	.31	.02	1.18	1.31	.24
68	1.07	.02	.99	.98	.43
69	-.36	.02	.88	.80	.49
70	-.44	.02	1.02	1.00	.34

Table 3.7
Results of the 2000 Equating Process–Mathematics Grade 5

Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
1	-1.92	.03	.93	.84	.33
2	-.92	.02	.97	1.04	.34
3	-.38	.02	.96	1.00	.43
4	.37	.02	.96	.94	.45
5	1.05	.02	1.16	1.26	.23
6	.89	.02	.98	.98	.43
7	-1.10	.02	.94	.88	.38
8	1.20	.02	1.07	1.15	.34
9	.67	.02	1.00	1.02	.41
10	.63	.02	1.15	1.19	.28
11	-.59	.02	.98	.99	.39
12	.68	.02	.93	.91	.48
13	-.91	.02	1.08	1.23	.26
14	.10	.02	1.02	1.01	.40
15	.03	.02	1.13	1.21	.28
16	.08	.02	1.01	1.03	.40
17	-.32	.02	1.00	1.07	.38
18	-.14	.02	1.06	1.16	.35
19	.92	.02	1.03	1.09	.37
20	-.13	.02	.80	.75	.59
21	-.08	.02	.93	.88	.47
22	.00	.02	.99	.97	.41
23	.28	.02	.99	.99	.43
24	1.32	.02	1.10	1.18	.32
25	1.04	.02	.95	.96	.45
26	-1.35	.02	.92	.91	.36
27	-.61	.02	.91	.88	.44
28	.43	.02	.96	.94	.46
29	1.27	.02	1.03	1.08	.43
30	.36	.02	.93	.89	.48
31	.37	.02	1.15	1.18	.28
32	-1.72	.02	1.01	1.24	.25
33	.56	.02	1.00	1.00	.42
34	.09	.02	.88	.86	.52
35	.92	.02	.96	.96	.45
36	-.50	.02	.89	.78	.50
37	.13	.02	.98	1.00	.43
38	.27	.02	1.02	1.00	.39
39	.90	.02	1.12	1.19	.31
40	.63	.02	1.06	1.10	.36
41	.21	.02	1.18	1.27	.24
42	.52	.02	1.01	1.02	.40
43	.75	.02	1.09	1.13	.34
44	-.13	.02	1.03	1.03	.37
45	-1.01	.02	.84	.75	.50
46	.33	.02	1.02	1.01	.39
47	.85	.02	1.11	1.17	.31
48	1.22	.02	.86	.85	.53
49	-1.38	.02	.95	.95	.35

Table 3.7 (continued)

50	.48	.02	.99	.98	.43
51	-.70	.02	.99	.98	.37
52	-.72	.02	.82	.72	.54
53	-1.48	.02	.93	.83	.37
54	.18	.02	1.09	1.15	.32
55	-.63	.02	.94	.88	.43
57	-1.42	.02	.83	.73	.32
58	1.02	.02	.96	.99	.44
59	-1.02	.02	.80	.64	.54
60	-.79	.02	.88	.83	.47
61	.38	.02	1.08	1.15	.34
62	1.26	.02	.95	.98	.43
63	-.14	.02	.88	.83	.51
64	.25	.02	1.07	1.10	.35
65	.29	.02	1.15	1.22	.28
66	.81	.02	1.04	1.05	.38
67	.23	.02	1.12	1.15	.31
68	1.09	.02	.96	.98	.40
69	.87	.02	1.02	1.05	.39
70	-.63	.02	1.00	1.05	.37
71	.66	.02	.88	.86	.53
72	-.14	.02	1.00	1.02	.40
73	.53	.02	.88	.85	.53
74	.65	.02	1.01	1.02	.41
75	.65	.02	1.10	1.12	.32
76	1.23	.02	.82	.78	.54
77	.79	.02	1.08	1.10	.34
78	.32	.02	.98	.97	.43
79	-.08	.02	.93	.88	.47
80	-.09	.02	1.05	1.09	.35
81	-.27	.02	.97	1.00	.41

Table 3.8
Results of the 2000 Equating Process–Mathematics Grade 8

Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
1	-.63	.02	1.03	1.01	.37
2	-.38	.02	1.09	1.27	.29
3	.50	.02	1.04	1.00	.41
4	.89	.02	1.01	1.03	.44
5	-.47	.02	.83	.73	.50
6	1.14	.02	.94	.94	.50
7	.05	.02	.87	.78	.53
8	1.29	.02	1.05	1.08	.41
9	-.17	.02	.95	.88	.45
10	.28	.02	1.17	1.29	.27
11	1.34	.02	1.11	1.14	.36
12	.84	.02	1.06	1.09	.40
13	.24	.02	.99	.97	.43
14	1.13	.02	1.04	1.05	.42
15	1.71	.02	1.03	1.10	.40
16	.32	.02	.90	.86	.51
17	-1.10	.02	.82	.68	.39
18	.90	.02	.89	.85	.54
19	-.40	.02	.91	.85	.46
20	.93	.02	.99	.97	.46
21	1.59	.02	1.13	1.19	.34
22	.78	.02	.89	.83	.54
23	2.08	.02	.88	.84	.53
24	2.09	.02	1.06	1.22	.36
25	.78	.02	.98	.94	.46
26	1.04	.02	1.14	1.24	.33
27	.97	.02	1.02	1.03	.43
28	1.11	.02	.90	.86	.54
29	.32	.02	.92	.94	.49
30	1.49	.02	.94	.96	.47
31	1.21	.02	1.01	1.03	.44
32	-.95	.02	1.11	1.06	.44
33	-1.20	.02	.99	.96	.40
34	.14	.02	.98	1.00	.41
35	.61	.02	1.08	1.08	.38
36	1.54	.02	1.10	1.16	.37
37	1.23	.02	1.13	1.18	.34
38	-.40	.02	.99	.92	.39
39	-.20	.02	1.12	1.23	.28
40	.38	.02	1.01	1.06	.42
41	.55	.02	1.13	1.24	.33
42	1.43	.02	.96	.99	.48
43	-.55	.02	.87	.79	.46
44	-.37	.02	.83	.72	.52
45	1.16	.02	1.07	1.09	.38
46	1.19	.02	.89	.87	.53
47	-1.53	.03	.91	.81	.35
48	-1.10	.02	.94	1.09	.35
49	1.31	.02	1.06	1.09	.40

Table 3.8 (continued)

50	-1.46	.03	.89	.92	.37
51	1.34	.02	.93	.92	.50
52	.56	.02	.98	.94	.46
53	.18	.02	.92	.86	.49
54	1.50	.02	1.06	1.09	.40
55	-.01	.02	.91	.87	.48
57	.09	.02	.97	1.01	.44
58	1.98	.02	1.12	1.23	.34
59	2.05	.02	1.08	1.20	.36
60	.92	.02	.95	.93	.50
61	1.38	.02	.99	1.01	.45
62	.17	.02	1.01	1.11	.41
63	.14	.02	1.01	1.12	.41
64	.35	.02	1.10	1.18	.35
65	1.03	.02	1.15	1.28	.33
66	.77	.02	.98	.96	.47
67	.41	.02	.86	.82	.55
68	.68	.02	.99	1.03	.45
69	2.00	.02	1.11	1.29	.32
70	-.59	.02	.90	.78	.46
71	.86	.02	1.06	1.10	.40
72	.86	.02	1.16	1.27	.32
73	-.65	.02	.84	.69	.50
74	.35	.02	.93	.86	.50
75	-.76	.02	.92	.77	.43
76	.98	.02	1.08	1.10	.38
77	-.35	.02	.99	.96	.39
78	.82	.02	1.09	1.09	.37
79	.27	.02	.84	.75	.57
80	1.73	.02	1.08	1.15	.37
81	1.41	.02	1.07	1.10	.40

Table 3.9
Results of the 2000 Equating Process–Mathematics Grade 10

Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
1	-1.05	.02	1.09	1.56	.19
2	-.16	.02	1.00	.97	.39
3	-2.40	.04	.88	.52	.33
4	-1.02	.02	.98	.99	.32
5	.89	.02	.96	.92	.46
6	.97	.02	.99	.99	.44
7	.98	.02	1.22	1.28	.23
8	-.70	.02	.77	.76	.34
9	1.43	.02	1.01	1.05	.42
10	1.20	.02	.95	.95	.47
11	.70	.02	1.09	1.17	.35
12	.93	.02	.91	.88	.51
13	1.44	.02	.90	.90	.49
14	-.31	.02	1.12	1.41	.36
15	.32	.02	.93	.86	.48
16	.60	.02	1.12	1.22	.32
17	1.22	.02	1.04	1.05	.40
18	1.32	.02	.97	.99	.45
19	.26	.02	.84	.77	.56
20	-1.13	.02	1.00	1.13	.28
21	.05	.02	.94	.87	.51
22	1.09	.02	1.05	1.05	.38
23	1.82	.02	1.16	1.25	.28
24	.42	.02	1.12	1.24	.34
25	1.23	.02	1.05	1.07	.39
26	1.02	.02	.93	.90	.49
27	.42	.02	1.04	1.07	.38
28	-.62	.02	1.10	1.29	.37
29	-.46	.02	1.00	.96	.40
30	1.13	.02	1.05	1.08	.38
31	-1.03	.02	.92	.77	.39
32	.11	.02	.96	1.00	.43
33	-1.10	.02	.90	.79	.37
34	-.11	.02	1.05	1.19	.30
35	.88	.02	1.12	1.17	.32
36	.68	.02	1.08	1.13	.33
37	.56	.02	.97	.96	.45
38	1.05	.02	1.00	1.00	.43
39	.35	.02	1.07	1.13	.34
40	1.67	.02	.95	.98	.46
41	-.52	.02	.83	.65	.52
42	1.51	.02	1.23	1.33	.23
43	1.55	.02	1.03	1.09	.39
44	.77	.02	.90	.87	.51
45	.60	.02	.98	.96	.44
46	-.19	.02	1.03	1.10	.36
47	1.11	.02	1.24	1.33	.22
48	1.41	.02	1.15	1.19	.30
49	.26	.02	.91	.90	.49

Table 3.9 (continued)

50	1.76	.02	1.25	1.46	.20
51	.30	.02	.92	.90	.48
52	2.79	.02	1.07	1.65	.24
53	.06	.02	.88	.77	.52
54	-.06	.02	.86	.80	.52
55	-1.16	.02	.81	.67	.40
57	-.95	.02	.90	.83	.41
58	.16	.02	.94	.99	.35
59	.21	.02	.80	.72	.52
60	-.35	.02	1.06	1.21	.42
61	.21	.02	.96	.91	.48
62	.86	.02	.98	1.00	.45
63	.29	.02	.97	.98	.46
64	.04	.02	.98	.95	.42
65	.84	.02	1.16	1.22	.29
66	.38	.02	.93	.92	.48
67	1.66	.02	1.30	1.49	.16
68	.94	.02	.96	.98	.47
69	-.02	.02	.89	.83	.48
70	-.08	.02	1.05	1.18	.39
71	.90	.02	.85	.81	.56
72	1.46	.02	1.05	1.08	.37
73	-.04	.02	.87	.76	.52
74	.36	.02	.85	.77	.55
75	1.84	.02	.96	1.03	.38
76	.23	.02	.86	.79	.53
77	1.09	.02	.98	.96	.45
78	2.26	.02	1.04	1.23	.39
79	.54	.02	.87	.86	.52
80	1.07	.02	.95	.94	.47
81	.47	.02	1.00	.96	.44

Table 3.10
Results of the 2000 Scaling Process–Science Grade 4

Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
1	-.61	.02	.97	.95	.38
2	-.23	.02	1.03	1.05	.34
3	-.78	.02	1.05	1.10	.28
4	.71	.02	1.07	1.09	.31
5	-.09	.02	.93	.89	.45
6	1.98	.02	1.18	1.55	.10
7	-1.02	.02	1.00	.99	.32
8	-.30	.02	.96	.94	.41
9	.73	.02	1.18	1.22	.21
10	-.46	.02	.92	.86	.45
11	.10	.02	.97	.94	.41
12	-.64	.02	.89	.82	.47
13	1.22	.02	1.05	1.10	.30
14	.42	.02	1.11	1.12	.28
15	-.40	.02	1.06	1.09	.30
16	.82	.02	1.15	1.19	.23
17	.58	.02	1.11	1.14	.26
18	-.39	.02	.96	.93	.40
19	-.99	.02	.85	.73	.50
20	.67	.02	1.07	1.08	.30
21	-.90	.02	.97	1.01	.35
22	-.98	.02	.98	.94	.35
23	.26	.02	.96	.94	.42
24	.25	.02	1.10	1.11	.28
25	-.45	.02	.89	.83	.48
26	.74	.02	1.02	1.05	.35
27	.68	.02	1.08	1.10	.29
28	.11	.02	.89	.85	.50
29	.14	.02	.92	.89	.46
30	.06	.02	.98	.96	.40
31	-.03	.02	1.03	1.07	.34
32	-.18	.02	.84	.77	.54
33	-.50	.02	.95	.92	.41
34	.04	.02	1.03	1.07	.35
35	-.11	.02	.86	.81	.52
41	-.44	.02	1.04	1.07	.32
42	-1.25	.02	.91	.82	.41
43	1.56	.02	1.22	1.40	.12
44	.03	.02	1.06	1.09	.31
45	-1.05	.02	.87	.78	.46
46	.34	.02	.96	.94	.43
47	.26	.02	.97	.96	.41
48	.09	.02	.90	.86	.48
49	-.85	.02	.89	.82	.46
50	.45	.02	1.14	1.18	.24
51	-.32	.02	1.08	1.11	.28
52	-.49	.02	.89	.84	.47
53	.89	.02	1.02	1.04	.35
54	.14	.02	.99	.96	.39

Table 3.10 (continued)

55	-.30	.02	1.13	1.25	.22
56	.16	.02	1.05	1.06	.33
57	.22	.02	1.09	1.12	.29
58	-.71	.02	.88	.80	.47
59	-1.06	.02	1.04	1.16	.26
60	.25	.02	1.00	.99	.38
61	-.60	.02	.95	.96	.40
62	.02	.02	.96	.94	.42
63	.29	.02	1.00	.99	.38
64	-.10	.02	.90	.85	.48
65	.88	.02	1.16	1.22	.21
66	.04	.02	.93	.90	.45
67	-.13	.02	.92	.88	.46
68	.06	.02	1.02	1.04	.35
69	.37	.02	.91	.88	.47
70	-.01	.02	1.00	1.00	.38
71	.31	.02	1.06	1.05	.32
72	-.78	.02	.97	.95	.37
73	.63	.02	1.00	1.00	.38
74	.73	.02	1.07	1.10	.30
75	-.10	.02	.90	.86	.48

Table 3.11
Results of the 2000 Scaling Process–Science Grade 7

Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
1	-1.07	.02	.89	.75	.44
2	.03	.02	1.11	1.15	.24
3	-.73	.02	.95	.87	.39
4	.38	.02	.91	.90	.46
5	.44	.02	1.03	1.03	.34
6	.48	.02	1.02	1.03	.35
7	.78	.02	1.06	1.09	.30
8	-1.00	.02	.89	.82	.43
9	-1.59	.02	.97	.91	.30
10	-.69	.02	.88	.79	.47
11	-.54	.02	1.03	1.05	.30
12	.50	.02	.95	.94	.42
13	.83	.02	1.03	1.07	.32
14	-1.27	.02	.93	.86	.37
15	1.26	.02	1.17	1.26	.17
16	-.31	.02	.98	.98	.37
17	.93	.02	1.07	1.09	.30
18	-1.10	.02	.97	.98	.33
19	-.13	.02	1.11	1.14	.24
20	-1.01	.02	.88	.75	.45
21	1.36	.02	1.26	1.41	.08
22	.69	.02	1.00	1.00	.37
23	-.18	.02	.87	.82	.50
24	.63	.02	1.10	1.11	.26
25	-.95	.02	.98	.97	.34
26	.60	.02	1.02	1.03	.34
27	.56	.02	1.18	1.22	.18
28	-.28	.02	1.06	1.07	.29
29	-.01	.02	1.01	1.01	.35
30	-.03	.02	.91	.86	.46
31	-.37	.02	.98	1.07	.36
32	-.33	.02	.87	.79	.50
33	.74	.02	.97	.97	.40
34	.21	.02	.97	.96	.40
35	.20	.02	1.03	1.04	.34
41	-.36	.02	.92	.88	.44
42	-1.21	.02	.99	1.03	.30
43	-1.40	.02	.95	.98	.32
44	1.21	.02	1.07	1.12	.28
45	.16	.02	1.03	1.07	.34
46	.44	.02	.95	.93	.43
47	-.25	.02	.91	.86	.45
48	.01	.02	1.00	1.03	.36
49	-.61	.02	.88	.81	.47
50	.48	.02	.96	.96	.41
51	.30	.02	1.13	1.17	.24
52	.72	.02	1.01	1.02	.35
53	-.28	.02	1.05	1.09	.29
54	.09	.02	1.02	1.02	.34

Table 3.11 (continued)

55	-.39	.02	1.08	1.14	.26
56	.67	.02	1.00	1.01	.36
57	.04	.02	1.00	1.00	.36
58	.07	.02	1.20	1.28	.15
59	.35	.02	1.06	1.07	.31
60	-.29	.02	1.03	1.02	.32
61	-.16	.02	1.06	1.07	.29
62	-.27	.02	.96	.96	.40
63	.42	.02	.99	.99	.38
64	.61	.02	1.01	1.01	.36
65	.26	.02	.91	.88	.47
66	-.04	.02	.98	.96	.38
67	.32	.02	.96	.94	.41
68	.75	.02	1.06	1.09	.30
69	.44	.02	1.00	.99	.37
70	-.40	.02	.88	.82	.48
71	-.71	.02	.85	.72	.51
72	.28	.02	1.05	1.05	.31
73	.35	.02	.98	.96	.39
74	.10	.02	.93	.93	.44
75	-.72	.02	.95	.93	.38

Table 3.12
Results of the 2000 Scaling Process–Social Science Grade 4

Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
1	-.39	.02	1.06	1.08	.30
2	-.89	.02	.97	.98	.36
3	-1.01	.02	.95	.88	.38
4	-.69	.02	1.01	1.09	.33
5	-.51	.02	.94	.88	.43
6	-.46	.02	1.10	1.16	.26
7	-.63	.02	1.04	1.07	.31
8	-.75	.02	.95	.93	.39
9	-.02	.02	.97	.94	.41
10	-.86	.02	.96	.94	.38
11	-.53	.02	.91	.88	.45
12	.22	.02	.98	.96	.41
13	-.02	.02	.92	.85	.47
14	.40	.02	.99	.98	.39
15	.82	.02	.95	.94	.43
16	-1.10	.02	.91	.85	.41
17	-.38	.02	.92	.86	.45
18	-.48	.02	1.01	1.05	.34
19	-1.67	.03	.96	1.04	.30
20	-.86	.02	.88	.73	.48
21	-.50	.02	.87	.78	.50
22	.96	.02	1.23	1.31	.16
23	-.12	.02	.96	.91	.42
24	-.70	.02	1.07	1.22	.26
25	-.30	.02	1.06	1.09	.31
26	-.16	.02	.95	.91	.43
27	-.02	.02	1.14	1.18	.24
28	.15	.02	.91	.87	.47
29	.46	.02	1.13	1.14	.26
30	1.39	.02	1.14	1.27	.21
31	1.03	.02	.99	1.01	.39
32	1.40	.02	1.14	1.25	.22
33	-1.14	.02	.89	.80	.43
34	-.26	.02	.92	.87	.46
35	.23	.02	.86	.81	.53
36	-.27	.02	.83	.74	.55
37	.95	.02	1.03	1.04	.36
38	.82	.02	1.08	1.12	.30
39	-1.07	.02	.94	1.02	.37
40	.36	.02	.98	.97	.41
41	-.15	.02	.92	.87	.46
42	-.19	.02	.82	.74	.57
43	.55	.02	1.01	1.01	.38
44	1.01	.02	1.13	1.18	.26
45	-.26	.02	1.04	1.04	.33
46	-.33	.02	1.10	1.13	.26
47	.17	.02	1.14	1.17	.24
48	.32	.02	1.05	1.04	.34
49	.22	.02	1.15	1.22	.24

Table 3.12 (continued)

50	.98	.02	1.07	1.10	.32
51	.25	.02	1.15	1.19	.24
52	-.86	.02	.87	.75	.48
53	.37	.02	1.00	1.00	.39
54	-.12	.02	.89	.84	.50
55	-.14	.02	1.05	1.15	.32
56	-.29	.02	.97	.90	.40
57	1.33	.02	.96	1.00	.39
58	1.56	.02	1.14	1.30	.22
59	-.56	.02	.99	.97	.37
60	-.26	.02	.87	.79	.51
61	-.17	.02	1.01	1.04	.37
62	.04	.02	.94	.91	.45
63	.07	.02	.90	.86	.49
64	.38	.02	.97	.96	.42
65	-.36	.02	.88	.79	.50
66	-.29	.02	.91	.86	.46
67	.27	.02	1.02	1.05	.37
68	1.12	.02	1.10	1.15	.28
69	.88	.02	1.14	1.19	.24
70	1.06	.02	1.07	1.11	.31

Table 3.13
Results of the 2000 Scaling Process–Social Science Grade 7

Item	Difficulty	S_{ed}	Infit	Outfit	r_{pb}
1	-2.08	.03	.95	.89	.27
2	-1.22	.02	.97	.97	.29
3	-1.79	.03	.94	.84	.30
4	-1.26	.02	.97	.95	.30
5	1.01	.02	1.03	1.05	.30
6	.09	.02	1.00	.98	.34
7	.55	.02	1.02	1.02	.32
8	-.10	.02	1.09	1.15	.22
9	1.43	.02	1.11	1.20	.20
10	.16	.02	.95	.94	.40
11	.27	.02	.99	.97	.36
12	-.25	.02	.91	.85	.43
13	.65	.02	.96	.95	.38
14	-1.02	.02	.90	.85	.40
15	-.30	.02	.95	.95	.38
16	-.33	.02	1.02	1.03	.30
17	-.28	.02	1.04	1.06	.28
18	.78	.02	1.10	1.13	.23
19	.39	.02	1.09	1.11	.24
20	-.39	.02	.92	.85	.42
21	.56	.02	1.01	1.02	.34
22	.99	.02	1.10	1.15	.22
23	-1.27	.02	.89	.75	.40
24	-.07	.02	.97	.96	.37
25	.44	.02	1.11	1.17	.22
26	-.54	.02	1.01	1.05	.29
27	1.21	.02	1.03	1.07	.29
28	-1.19	.02	.93	.82	.36
29	-.44	.02	.85	.75	.51
30	1.32	.02	1.00	1.02	.33
31	.58	.02	1.01	1.02	.34
32	-.82	.02	1.06	1.04	.22
33	-.02	.02	.88	.82	.48
34	-1.31	.02	.92	.85	.36
35	.68	.02	1.02	1.02	.32
36	.35	.02	1.12	1.18	.20
37	.45	.02	.89	.87	.46
38	.87	.02	.92	.91	.43
39	.69	.02	.93	.91	.42
40	.50	.02	.93	.92	.42
41	-.20	.02	1.07	1.08	.25
42	-.03	.02	1.04	1.05	.29
43	.83	.02	1.02	1.04	.31
44	-.04	.02	1.01	.98	.33
45	-.39	.02	.91	.85	.43
46	.69	.02	.95	.95	.40
47	-1.25	.02	.94	.89	.33
48	-1.76	.03	.95	.81	.30
49	.44	.02	1.10	1.12	.23

Table 3.13 (continued)

50	.78	.02	1.02	1.03	.32
51	.13	.02	1.07	1.07	.26
52	1.22	.02	1.15	1.22	.17
53	-.18	.02	.91	.87	.43
54	.29	.02	.93	.90	.42
55	-.57	.02	.90	.83	.44
56	-.09	.02	.99	.98	.34
57	.98	.02	1.01	1.02	.33
58	.89	.02	1.09	1.13	.23
59	-.13	.02	1.10	1.17	.21
60	-.96	.02	1.06	1.29	.19
61	.18	.02	1.29	1.40	.01
62	.81	.02	.95	.94	.40
63	-.92	.02	.90	.78	.42
64	1.27	.02	1.13	1.18	.19
65	-.04	.02	.94	.90	.41
66	-.47	.02	.93	.92	.39
67	.42	.02	.96	.94	.39
68	.40	.02	1.09	1.13	.24
69	-1.44	.02	.91	.78	.37
70	-.15	.02	.91	.86	.43

The raw score that is initially derived from multiple-choice items has no particular meaning beyond the number of answers the student has answered correctly. Writing, on the other hand, uses criterion-referenced scales. Each point on these scales has a specific interpretation. For example, when readers evaluate the quality of a 3rd-grade persuasive essay's focus, they assign a score of 6 when the paper "sets its purpose in an introduction through either a general thematic introduction or a specific preview, maintains the position or logic throughout, addresses any previewed points, and provides an effective closing." They assign a score of 3 when the paper "lacks clarity, provides multiple positions with a unifying umbrella statement, contains responses that do not serve a persuasive purpose, or lacks sufficiency to demonstrate a developed focus." Transforming writing scores to another scale would lose the specific meanings attached to each score point. For this reason, the ISAT writing score is a simple summation of the features. Because of the importance of Integration, it is given double weight in the summation. This leads to a writing score that ranges from 6 to 32.

4. NATIONAL NORM COMPARISONS

The legislation that authorized the development of ISAT required that reports provide national comparative data as a secondary reference point for evaluating school improvement efforts. Since the costs of obtaining nationally representative samples of students for each test would be prohibitively expensive, that mandate has been met by administering a nationally standardized achievement test along with ISAT to a sample of Illinois students. The two score distributions are then compared to identify points on the ISAT scale that correspond to the 25th, 50th, and 75th percentile performance levels for the national sample. For 2000, this process was conducted for science and social science tests.

ISAT uses the Ninth Edition of the Stanford Achievement Tests (SAT9) for purposes of determining Illinois students' relative standing within the national population. The specific levels/norms of each SAT9 test used were as follows:

Grade 4: Intermediate 1

Grade 7: Advanced 1

Equipercentile methodology was used to equate scores on the two tests. In equipercentile equating, scores on two tests are assumed to be equivalent if they have the same percentile rank. For example, the SAT9 score that cuts off 10% of the equating sample is assumed to represent a level of proficiency equal to the ISAT score that cuts off 10% of the equating sample, even though the scores themselves may be quite different numerically.

In order to conduct the equating process, ISAT and SAT9 results were matched by name to create a set of records in which each student had ISAT results and a corresponding SAT9 score. Frequency distributions of ISAT and SAT9 scale scores were then compiled. Each scale score on the ISAT was matched to the corresponding scale score on the Stanford test, based on the cumulative mid-percentile interval associated with each score.

Table 4.1 summarizes results of these studies. It shows the sample sizes that were used for the equating, the average SAT9 national percentile for the samples, and the correlations between the two instruments. Table 4.2 presents the ISAT scale score cutoffs that define the *upper limits* of national quartile categories 1, 2, and 3. These are shown as score ranges for each national quarter. For example, scale scores of 120 to 145 on the 4th-grade science test define Q1, the quartile that represents the lowest 25% of student performance nationally. Note that although the scale score cutoffs remain the same from year to year, the percentage of students in each category need not remain constant. Table 4.2 also presents cutoffs for reading and mathematics, which were established by a parallel process in 1999, the first year in which those tests were administered.

Table 4.1
Summary of 2000 ISAT–SAT9 National Norm Studies

SCIENCE	Grade 4	Grade 7
Sample Size	1062	1128
Mean SAT9 NCE	51.4	56.8
r between ISAT and SAT9 Scores	.720	.818
SOCIAL SCIENCE	Grade 4	Grade 7
Sample Size	1077	570
Mean SAT9 NCE	60.0	54.7
r between ISAT and SAT9 Scores	.744	.781

Table 4.2
ISAT National Quarter Scale Score Cutoffs

READING	Q1	Q2	Q3	Q4
03	120-147	148-157	158-167	168-200
05	120-147	148-157	158-168	169-200
08	120-144	145-154	155-165	166-200
10	120-153	154-162	163-171	172-200
MATHEMATICS	Q1	Q2	Q3	Q4
03	120-145	146-155	156-166	167-200
05	120-146	147-156	157-166	167-200
08	120-144	145-154	155-164	165-200
10	120-146	147-154	155-164	165-200
SCIENCE	Q1	Q2	Q3	Q4
04	120-145	146-157	158-168	169-200
07	120-142	143-154	155-163	164-200
SOCIAL SCIENCE	Q1	Q2	Q3	Q4
04	120-144	145-155	156-166	167-200
07	120-145	146-154	155-165	166-200

The results of applying these cutoffs to the 2000 assessment data are shown in Chapter 5.

5. RESULTS

Performance Relative to the Illinois Learning Standards

Table 5.1 shows the percentages of students by performance level and by grade for reading. The percentage of students falling into the Exceeds category is highest at 3rd grade. The percentage of students not meeting standards is highest at 5th grade. Overall, the percentage of students meeting (or exceeding) standards is highest at 8th grade.

Table 5.1
Percentages of Students by Grade Falling into Each Performance Level for ISAT Reading: 1999-2000

Grade	Academic Warning	Below Standards	Meets Standards	Exceeds Standards
3				
1999	8	31	44	17
2000	6	32	41	21
5				
1999	1	38	37	24
2000	0	41	39	20
8				
1999	1	27	54	18
2000	0	28	56	16

Note: Because of rounding, the percentages in each row may not total exactly to 100%.

Table 5.2 provides additional information with respect to the reading test. It presents the average percent of items students answered correctly with respect to the standards sets that were previously described.

Table 5.2
Reading Average Percent Correct by Standards Sets: 2000

	Set					
Grade	1	2	3	4	5	6
03	67	60	68	61	68	72
05	63	63	66	63	61	–
08	68	69	66	70	72	–

Table 5.3 shows the percentages of students by performance level and by grade for mathematics. Generally, the percentage of students meeting state standards is lower for mathematics than for reading. Grade 3 is an exception to that rule. The percentage of students falling into the Exceeds category is highest at 3rd grade and lowest at 5th grade.

Table 5.3
Percentages of Students by Grade Falling into Each Performance Level for ISAT Mathematics: 1999-2000

Grade	Academic Warning	Below Standards	Meets Standards	Exceeds Standards
3				
1999	12	20	47	21
2000	10	21	46	23
5				
1999	6	39	53	3
2000	6	37	52	5
8				
1999	5	52	36	7
2000	8	46	35	12

Note: Because of rounding, the percentages in each row may not total exactly to 100%.

Table 5.4 presents the average percent of items students answered correctly with respect to the mathematics standards sets that were previously described.

Table 5.4
Mathematics Average Percent Correct by Standards Sets: 2000

	Set							
Grade	1	2	3	4	5	6	7	8
03	60	61	52	62	74	57	73	62
05	54	55	52	61	65	50	57	61
08	57	55	52	56	48	49	62	63

Table 5.5 shows results for writing. A greater percentage of 5th and 8th grade students meet standards with respect to writing as compared to 3rd graders.

Table 5.6 summarizes results with respect to writing feature scores. Note that Conventions is scored on a two-point scale while all other features are scored on a six-point scale.

Table 5.5
Percentages of Students by Grade Falling into Each Performance Level for ISAT Writing: 2000

Grade	Academic Warning	Below Standards	Meets Standards	Exceeds Standards
3				
1999	9	35	50	6
2000	6	38	53	2
5				
1999	2	23	52	23
2000	3	26	57	14
8				
1999	5	36	56	3
2000	3	27	59	11

Note: Because of rounding, the percentages in each row may not total exactly to 100%.

Table 5.6
Mean Writing Feature Scores of Students by Prompt: 2000

Grade	Type	F	S	O	C	I
03	P	4.52	3.67	3.69	1.94	3.89
03	E	4.49	3.72	3.69	1.94	3.90
03	N	3.92	3.93	3.77	1.90	3.87
05	P	4.64	4.20	4.22	1.94	4.31
05	E	4.64	4.15	4.16	1.95	4.27
05	N	4.29	4.38	4.21	1.95	4.28
08	P	3.93	3.74	3.72	1.90	3.77
08	E	3.95	3.74	3.69	1.90	3.76
08	N	4.03	4.11	3.93	1.87	4.03

Note: Prompt type: P = Persuasive; E = Expository; N = Narrative

Table 5.8 shows the percentages of students by performance level and by grade for science.

Table 5.8
Percentages of Students by Grade Falling into Each Performance Level for ISAT Science: 2000

Grade	Academic Warning	Below Standards	Meets Standards	Exceeds Standards
4				
2000	1	35	51	13
7				
2000	12	16	54	18

Note: Because of rounding, the percentages in each row may not total exactly to 100%.

Table 5.9 presents the average percent of items students answered correctly with respect to the science standards sets that were previously described.

Table 5.9
Science Average Percent Correct by Standards Sets: 2000

	Set				
Grade	1	2	3	4	5
04	60	57	62	59	61
07	59	61	57	57	60

Table 5.10 shows the percentages of students by performance level and by grade for social science.

Table 5.10
Percentages of Students by Grade Falling into Each Performance Level for ISAT Social Science: 2000

Grade	Academic Warning	Below Standards	Meets Standards	Exceeds Standards
04				
2000	11	30	53	6
07				
2000	3	39	46	12

Note: Because of rounding, the percentages in each row may not total exactly to 100%.

Table 5.11 presents the average percent of items students answered correctly with respect to the social science standards sets that were previously described.

Table 5.11
Social Science Average Percent Correct by Standards Sets: 2000

	Set				
Grade	1	2	3	4	5
04	66	67	65	58	62
07	70	61	60	61	62

Performance Relative to National Quarters

National quarters are reported for each subject area in Table 5.12. As noted earlier, results in writing are not reported relative to national quarters.

Table 5.12
Percentages of Students by Grade and Learning Area Falling into Each National Quartile: 1999-2000

READING	Q1	Q2	Q3	Q4
Grade/Year				
3				
1999	22	22	25	32
2000	21	21	25	33
5				
1999	21	23	27	28
2000	21	26	28	25
8				
1999	15	22	30	33
2000	13	24	33	30
MATHEMATICS	Q1	Q2	Q3	Q4
Grade/Year				
3				
1999	19	21	28	32
2000	18	21	26	36
5				
1999	20	22	24	33
2000	19	22	21	38
8				
1999	15	25	25	35
2000	18	20	21	41

Table 5.12 (continued)

SCIENCE	Q1	Q2	Q3	Q4
Grade/Year				
4				
2000	18	26	25	31
7				
2000	14	24	22	41
SOCIAL SCIENCE	Q1	Q2	Q3	Q4
Grade/Year				
4				
2000	17	21	29	33
7				
2000	17	19	29	35

Note: Because of rounding, the percentages in each row may not total exactly to 100%.

References

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement (3rd Edition)* (pp. 105-146). New York: Macmillan.

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist, 18,* 519-521.

Peng, C-Y, J., & Subkoviak, M. J. (1980). A note on Huynh's normal approximation procedure for estimating criterion-referenced reliability. *Journal of Educational Measurement, 17,* 359-368.

Subkoviak, M. J. (1984). Estimating the reliability of mastery-non-mastery classifications. In R. A. Berk (Ed.), *A guide to criterion-referenced test construction* (pp. 267-291). Baltimore: Johns Hopkins Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design: Rasch measurement*. Chicago: Mesa.

APPENDIX A. SUPPLEMENTARY TABLES

Tables A.1 through A.5 present correlations among the various standards sets, goal, or feature scores presented in student, school, and district reports. The sample sizes for the various analyses are summarized below. For writing at grades 5, 8, and 10, the sample size refers to the number of papers, not the number of students.

Reading: Grade 3	141,129
Reading: Grade 5	143,532
Reading: Grade 8	136,617
Reading: Grade 10	70,125
Mathematics: Grade 3	142,398
Mathematics: Grade 5	144,110
Mathematics: Grade 8	136,512
Mathematics: Grade 10	69,763
Writing: Persuasive Prompt: Grade 3	46,073
Writing: Expository Prompt: Grade 3	46,373
Writing: Narrative Prompt: Grade 3	45,845
Writing: Persuasive Prompt: Grade 5	67,553
Writing: Expository Prompt: Grade 5	75,176
Writing: Narrative Prompt: Grade 5	142,931
Writing: Persuasive Prompt: Grade 8	95,141
Writing: Expository Prompt: Grade 8	40,982
Writing: Narrative Prompt: Grade 8	136,202
Writing: Persuasive Prompt: Grade 10	70,841
Writing: Expository Prompt: Grade 10	46,841
Writing: Narrative Prompt: Grade 10	23,603
Science: Grade 4	148,003
Science: Grade 7	148,213
Social Sciences: Grade 4	138,035
Social Sciences: Grade 7	138,278

Table A.1
Intercorrelations among Reading Standards Sets

Grade 3	S1	S2	S3	S4	S5	S6
S1	1.000	.787	.803	.942	.863	.647
S2	.792	1.000	.826	.865	.775	.613
S3	.884	.872	1.00	.736	.620	.584
S4	.942	.873	.831	1.000	.794	.623
S5	.866	.795	.767	.782	1.000	.612
S6	.649	.613	.623	.628	.624	1.000
Grade 5	S1	S2	S3	S4	S5	
S1	1.000	.782	.806	.866	.789	
S2	.767	1.000	.828	.935	.880	
S3	.814	.851	1.000	.756	.712	
S4	.861	.940	.780	1.000	.772	
S5	.784	.839	.722	.769	1.000	
Grade 8	S1	S2	S3	S4	S5	
S1	1.000	.781	.869	.757	.840	
S2	.741	1.000	.909	.889	.876	
S3	.706	.913	1.000	.754	.771	
S4	.919	.871	.745	1.000	.748	
S5	.731	.875	.711	.762	1.000	
Grade 10	S1	S2	S3	S4	S5	
S1	1.000	.746	.772	.951	.811	
S2	.714	1.000	.767	.832	.808	
S3	.743	.646	1.000	.711	.643	
S4	.942	.847	.656	1.000	.743	
S5	.779	.708	.559	.693	1.000	

Note: Values for Form A are presented above the principal diagonal, and values for Form B are presented below the principal diagonal.

Table A.2
Intercorrelations among Mathematics Standards Sets

Grade 3	S1	S2	S3	S4	S5	S6	S7	S8
S1	1.000	.901	.355	.790	.616	.857	.831	.811
S2	.901	1.000	.380	.685	.558	.741	.720	.715
S3	.355	.380	1.000	.343	.289	.330	.331	.322
S4	.790	.685	.343	1.000	.712	.778	.668	.730
S5	.616	.558	.289	.712	1.000	.589	.558	.565
S6	.857	.741	.330	.778	.589	1.000	.684	.717
S7	.831	.720	.331	.668	.558	.684	1.000	.683
S8	.811	.715	.322	.730	.565	.717	.683	1.000
Grade 5	S1	S2	S3	S4	S5	S6	S7	S8
S1	1.000	.931	.878	.798	.740	.859	.886	.773
S2	.931	1.000	.845	.739	.719	.741	.767	.681
S3	.878	.845	1.000	.691	.627	.719	.745	.642
S4	.798	.739	.691	1.000	.738	.694	.710	.658
S5	.740	.719	.627	.738	1.000	.633	.645	.596
S6	.859	.741	.719	.694	.633	1.000	.700	.621
S7	.886	.767	.745	.710	.645	.700	1.000	.675
S8	.773	.681	.642	.658	.596	.621	.675	1.000
Grade 8	S1	S2	S3	S4	S5	S6	S7	S8
S1	1.000	.958	.742	.824	.800	.898	.873	.830
S2	.958	1.000	.767	.787	.809	.865	.805	.782
S3	.742	.767	1.000	.669	.671	.662	.661	.635
S4	.824	.787	.669	1.000	.755	.727	.704	.666
S5	.800	.809	.671	.755	1.000	.732	.681	.649
S6	.898	.865	.662	.727	.732	1.000	.732	.701
S7	.873	.805	.661	.704	.681	.732	1.000	.755
S8	.830	.782	.635	.666	.649	.701	.755	1.000
Grade 10	S1	S2	S3	S4	S5	S6	S7	S8
S1	1.000	.971	.774	.917	.723	.808	.876	.723
S2	.971	1.000	.777	.914	.747	.763	.805	.650
S3	.774	.777	1.000	.734	.569	.647	.699	.541
S4	.917	.914	.734	1.000	.638	.732	.761	.612
S5	.723	.747	.569	.638	1.000	.548	.563	.436
S6	.808	.763	.647	.732	.548	1.000	.686	.557
S7	.876	.805	.699	.761	.563	.686	1.000	.697
S8	.723	.650	.541	.612	.436	.557	.697	1.000

Table A.3
Intercorrelations among Writing Feature Scores

Persuasive Prompt: Grade 3	F	S	O	C	I
F	1.000	.669	.698	.431	.714
S	.669	1.000	.846	.373	.931
O	.698	.846	1.000	.395	.920
C	.431	.373	.395	1.000	.412
I	.714	.931	.920	.412	1.000
Expository Prompt: Grade 3	F	S	O	C	I
F	1.000	.657	.692	.408	.706
S	.657	1.000	.837	.357	.927
O	.692	.837	1.000	.380	.914
C	.408	.357	.380	1.000	.403
I	.706	.927	.914	.403	1.000
Narrative Prompt: Grade 3	F	S	O	C	I
F	1.000	.754	.915	.271	.964
S	.754	1.000	.768	.308	.794
O	.915	.768	1.000	.290	.945
C	.271	.308	.290	1.000	.278
I	.964	.794	.945	.278	1.000
Persuasive Prompt: Grade 5	F	S	O	C	I
F	1.000	.710	.724	.307	.751
S	.710	1.000	.895	.261	.947
O	.724	.895	1.000	.286	.951
C	.307	.261	.286	1.000	.287
I	.751	.947	.951	.287	1.000
Expository Prompt: Grade 5	F	S	O	C	I
F	1.000	.726	.744	.325	.766
S	.726	1.000	.890	.278	.946
O	.744	.890	1.000	.307	.947
C	.325	.278	.307	1.000	.307
I	.766	.946	.947	.307	1.000

Table A.3 (continued)

Narrative Prompt: Grade 5	F	S	O	C	I
F	1.000	.825	.915	.236	.945
S	.825	1.000	.839	.267	.891
O	.915	.839	1.000	.250	.950
C	.236	.267	.250	1.000	.252
I	.945	.891	.950	.252	1.000
Persuasive Prompt: Grade 8	**F**	**S**	**O**	**C**	**I**
F	1.000	.711	.761	.271	.825
S	.711	1.000	.815	.257	.876
O	.761	.815	1.000	.309	.933
C	.271	.257	.309	1.000	.300
I	.825	.876	.933	.300	1.000
Expository Prompt: Grade 8	**F**	**S**	**O**	**C**	**I**
F	1.000	.731	.780	.320	.841
S	.731	1.000	.836	.293	.890
O	.780	.836	1.000	.338	.940
C	.320	.293	.338	1.000	.336
I	.841	.890	.940	.336	1.000
Narrative Prompt: Grade 8	**F**	**S**	**O**	**C**	**I**
F	1.000	.822	.897	.210	.923
S	.822	1.000	.862	.220	.920
O	.897	.862	1.000	.237	.950
C	.210	.220	.237	1.000	.229
I	.923	.920	.950	.229	1.000
Persuasive Prompt: Grade 10	**F**	**S**	**O**	**C**	**I**
F	1.000	.732	.748	.294	.782
S	.732	1.000	.921	.283	.952
O	.748	.921	1.000	.293	.964
C	.294	.283	.293	1.000	.295
I	.782	.952	.964	.295	1.000

Table A.3 (continued)

Expository Prompt: Grade 10	F	S	O	C	I
F	1.000	.734	.752	.296	.787
S	.734	1.000	.916	.275	.949
O	.752	.916	1.000	.286	.963
C	.296	.275	.286	1.000	.290
I	.787	.949	.963	.290	1.000
Narrative Prompt: Grade 10	**F**	**S**	**O**	**C**	**I**
F	1.000	.825	.893	.222	.945
S	.825	1.000	.792	.247	.894
O	.893	.792	1.000	.219	.929
C	.222	.247	.219	1.000	.244
I	.945	.894	.929	.244	1.000

Table A.4
Intercorrelations among Science Standards Sets

Grade 4	S1	S2	S3	S4	S5
S1	1.000	.691	.645	.676	.691
S2	.691	1.000	.659	.696	.697
S3	.645	.659	1.000	.659	.667
S4	.676	.696	.659	1.000	.699
S5	.691	.697	.667	.699	1.000
Grade 7	S1	S2	S3	S4	S5
S1	1.000	.711	.622	.671	.708
S2	.711	1.000	.632	.682	.710
S3	.622	.632	1.000	.619	.609
S4	.671	.682	.619	1.000	.687
S5	.708	.710	.609	.687	1.000

Table A.5
Intercorrelations among Social Sciences Standards Sets

Grade 4	S1	S2	S3	S4	S5
S1	1.000	.690	.630	.603	.667
S2	.690	1.000	.709	.652	.731
S3	.630	.709	1.000	.633	.695
S4	.603	.652	.633	1.000	.654
S5	.667	.731	.695	.654	1.000
Grade 7	S1	S2	S3	S4	S5
S1	1.000	.582	.590	.594	.633
S2	.582	1.000	.574	.581	.621
S3	.590	.574	1.000	.589	.659
S4	.594	.581	.589	1.000	.639
S5	.633	.621	.659	.639	1.000

APPENDIX B. FACTOR ANALYSES OF MULTIPLE-CHOICE AND OPEN-ENDED COMPONENTS OF THE ISAT LANGUAGE ARTS AND MATHEMATICS TESTS

A series of factor analyses were conducted to examine the relationship between the multiple-choice and open-ended components of the ISAT language arts (reading, writing) and mathematics tests.

Factor analysis is a well-regarded statistical technique that examines the covariation among a set of variables and identifies a generally smaller set of constructs (factors) that may be thought to explain these relationships.

Separate analyses were conducted at the four grade levels at which the ISAT reading, writing, and mathematics tests are administered. Each analysis began with a set of variables designed to represent both types of item formats–multiple-choice and open-ended–used in the tests. These variables are identified in Table B.1. Because of differences in the format of the 3rd-grade test, not all variables were available for this analysis.

Table B.1
Description of Variables Used in the Factor Analyses

READMC1	Number correct on reading passage 1 multiple-choice items
READMC2	Number correct on reading passage 2 multiple-choice items
READMC3	Number correct on reading passage 3 multiple-choice items
READSA1	Score on first reading short answer question
READSA2	Score on second reading short answer question
MATHSA1K	Knowledge score on first mathematics short answer question
MATHSA1S	Strategy score on first mathematics short answer question
MATHSA1E	Explanation score on first mathematics short answer question
MATHSA2K	Knowledge score on second mathematics short answer question
MATHSA2S	Strategy score on second mathematics short answer question
MATHSA2E	Explanation score on second mathematics short answer question
MATHMC1	Number correct on mathematics session 1 multiple-choice items
MATHMC2	Number correct on mathematics session 2 multiple-choice items
MATHMC3*	Number correct on mathematics session 3 multiple-choice items
F1	Focus score on writing sample 1
S1	Support score on writing sample 1
O1	Organization score on writing sample 1
C1	Conventions score on writing sample 1
F2*	Focus score on writing sample 2
S2*	Support score on writing sample 2
O2*	Organization score on writing sample 2
C2*	Conventions score on writing sample 2

* indicates a variable not available for the grade 3 factor analysis

The same procedures were used at all grades. An inspection of the roots of the unreduced correlation matrix led to a decision to retain those greater than 1.0. (At 3rd grade, a decision was made to include one additional factor in the analysis for consistency with the pattern of relationships seen at all other grades.) Communalities were obtained through an interative principle-axis procedure. The resulting factor matrix was rotated to oblique simple structure using the Oblimin criterion.

Results for each grade are presented in a series of five tables. The first table shows the eigenvalues of the unreduced correlation matrix. The ratio of successive eigenvalues to the sum of all eigenvalues is an index of the relative amount of variance explained by the associated factor. For example, at 3rd grade, the first factor explains 44.96% of the variance (7.64/17). Together, the first five factors explain slightly more than 75% of the total variance in the set of 17 variables.

The second table shows the communalities of the variables. The column labeled "final" presents the values that actually lead to the subsequent solution.

The third table shows the factor loadings, technically pattern coefficients. These may be thought of as regression coefficients which would be applied to factor scores to produce the measured variables. The fourth table shows the factor structure matrix. Values in this matrix represent the correlations between the measured variables and the pure factors. The fifth table in each series shows the correlations among the factors.

Results for grade three are shown in Tables B.2 through B.6. These are followed by results for grade 5 (Tables B.7 through B.11), grade 8 (Tables B.12 through B16) and grade 10 (Tables B.17 through B.21).

Table B.2
Total Variance Explained: Grade 3

Factor	Initial Eigenvalues	% of Variance	Cumulative %
1	7.64	44.96	44.96
2	1.72	10.12	55.08
3	1.30	7.66	62.74
4	1.18	6.95	69.68
5	0.97	5.69	75.38
6	0.76	4.47	79.84
7	0.57	3.37	83.22
8	0.46	2.71	85.93
9	0.45	2.63	88.56
10	0.42	2.50	91.06
11	0.33	1.94	93.00
12	0.26	1.52	94.52
13	0.26	1.51	96.03
14	0.22	1.31	97.33
15	0.17	0.99	98.32
16	0.14	0.84	99.17
17	0.14	0.83	100.00

Table B.3
Communalities: Grade 3

	Initial	Extraction
READMC1	0.69	0.74
READMC2	0.67	0.71
READMC3	0.69	0.72
READSA1	0.42	0.58
READSA2	0.34	0.52
MATHSA1K	0.74	0.80
MATHSA1S	0.76	0.90
MATHSA1E	0.51	0.50
MATHSA2K	0.58	0.68
MATHSA2S	0.61	0.77
MATHSA2E	0.50	0.52
MATHMC1	0.77	0.75
MATHMC2	0.77	0.75
F1	0.48	0.50
S1	0.67	0.69
O1	0.73	0.94
C1	0.19	0.19

Table B.4
Pattern Matrix: Grade 3

	Factor				
	1	2	3	4	5
READMC1	0.82				
READMC2	0.84				
READMC3	0.84				
READSA1					0.68
READSA2					0.71
MATHSA1K				0.91	
MATHSA1S				0.98	
MATHSA1E				0.57	
MATHSA2K			0.84		
MATHSA2S			0.91		
MATHSA2E			0.57		
MATHMC1	0.74				
MATHMC2	0.74				
F1		0.64			
S1		0.80			
O1		1.02			
C1		0.28			

Table B.5
Structure Matrix: Grade 3

	Factor				
	1	2	3	4	5
READMC1	0.86	0.53	0.44	0.49	0.54
READMC2	0.84	0.46	0.41	0.45	0.54
READMC3	0.85	0.48	0.41	0.45	0.55
READSA1	0.53	0.39	0.34	0.35	0.75
READSA2	0.42	0.37	0.28	0.29	0.72
MATHSA1K	0.50	0.32	0.44	0.89	0.33
MATHSA1S	0.51	0.34	0.46	0.95	0.35
MATHSA1E	0.46	0.36	0.47	0.68	0.39
MATHSA2K	0.41	0.28	0.83	0.41	0.29
MATHSA2S	0.42	0.31	0.88	0.43	0.31
MATHSA2E	0.46	0.38	0.69	0.47	0.41
MATHMC1	0.85	0.52	0.54	0.57	0.51
MATHMC2	0.85	0.52	0.54	0.57	0.51
F1	0.46	0.70	0.31	0.32	0.38
S1	0.48	0.83	0.32	0.33	0.45
O1	0.49	0.97	0.33	0.34	0.44
C1	0.37	0.40			

Table B.6
Factor Correlation Matrix: Grade 3

Factor	1	2	3	4	5
1	1.00	0.57	0.51	0.56	0.60
2	0.57	1.00	0.38	0.38	0.49
3	0.51	0.38	1.00	0.52	0.38
4	0.56	0.38	0.52	1.00	0.39
5	0.60	0.49	0.38	0.39	1.00

Table B.7
Total Variance Explained: Grade 5

Factor	Initial Eigenvalues	% of Variance	Cumulative %
1	9.29	42.22	42.22
2	2.20	10.01	52.24
3	1.60	7.29	59.53
4	1.31	5.95	65.47
5	1.27	5.79	71.27
6	1.03	4.69	75.96
7	0.89	4.04	80.00
8	0.69	3.14	83.14
9	0.60	2.74	85.87
10	0.50	2.26	88.13
11	0.49	2.22	90.35
12	0.38	1.75	92.10
13	0.33	1.48	93.59
14	0.29	1.31	94.90
15	0.27	1.24	96.13
16	0.19	0.88	97.01
17	0.18	0.81	97.82
18	0.17	0.77	98.58
19	0.11	0.49	99.07
20	0.09	0.39	99.46
21	0.07	0.33	99.78
22	0.05	0.22	100.00

Table B.8
Communalities: Grade 5

	Initial	Extraction
READMC1	0.63	0.67
READMC2	0.63	0.66
READMC3	0.64	0.67
READSA1	0.39	0.49
READSA2	0.35	0.47
MATHSA1K	0.90	0.90
MATHSA1S	0.91	0.98
MATHSA1E	0.44	0.43
MATHSA2K	0.86	0.87
MATHSA2S	0.87	0.96
MATHSA2E	0.44	0.44
MATHMC1	0.77	0.79
MATHMC2	0.76	0.78
MATHMC3	0.76	0.78
F1	0.85	0.89
S1	0.74	0.77
O1	0.86	0.93
C1	0.15	0.11
F2	0.58	0.62
S2	0.81	0.85
O2	0.82	0.90
C2	0.17	0.13

Table B.9
Pattern Matrix: Grade 5

	Factor					
	1	2	3	4	5	6
READMC1	0.76					
READMC2	0.73					
READMC3	0.73					
READSA1						0.49
READSA2						0.53
MATHSA1K				0.97		
MATHSA1S				1.02		
MATHSA1E				0.50		
MATHSA2K			0.93			
MATHSA2S			0.98			
MATHSA2E			0.55			
MATHMC1	0.78					
MATHMC2	0.81					
MATHMC3	0.81					
F1		0.98				
S1		0.82				
O1		1.00				
C1						
F2					0.72	
S2					0.94	
O2					0.98	
C2						

Table B.10
Structure Matrix: Grade 5

	Factor					
	1	2	3	4	5	6
READMC1	0.81	0.40	0.36	0.41	0.42	0.48
READMC2	0.80	0.42	0.33	0.40	0.45	0.49
READMC3	0.80	0.43	0.34	0.40	0.45	0.51
READSA1	0.52	0.39	0.34	0.36	0.41	0.65
READSA2	0.44	0.39	0.30	0.32	0.41	0.65
MATHSA1K	0.47	0.29	0.40	0.95	0.31	
MATHSA1S	0.48	0.30	0.41	0.99	0.33	0.26
MATHSA1E	0.39	0.31	0.41	0.62	0.34	0.36
MATHSA2K	0.42	0.27	0.93	0.41	0.28	0.26
MATHSA2S	0.43	0.28	0.98	0.43	0.30	0.26
MATHSA2E	0.30	0.28	0.63	0.38	0.30	0.34
MATHMC1	0.88	0.47	0.46	0.53	0.51	0.38
MATHMC2	0.87	0.45	0.44	0.52	0.50	0.36
MATHMC3	0.87	0.45	0.45	0.51	0.50	0.36
F1	0.43	0.94	0.26	0.29	0.50	0.34
S1	0.45	0.88	0.28	0.31	0.55	0.38
O1	0.44	0.96	0.27	0.30	0.50	0.34
C1	0.28	0.26			0.28	
F2	0.47	0.48	0.28	0.33	0.78	0.37
S2	0.44	0.51	0.29	0.32	0.92	0.38
O2	0.47	0.51	0.30	0.33	0.95	0.38
C2	0.30				0.33	

Table B.11
Factor Correlation Matrix: Grade 5

Factor	1	2	3	4	5	6
1	1.00	0.48	0.41	0.50	0.53	0.46
2	0.48	1.00	0.29	0.33	0.56	0.39
3	0.41	0.29	1.00	0.45	0.31	0.30
4	0.50	0.33	0.45	1.00	0.36	0.30
5	0.53	0.56	0.31	0.36	1.00	0.42
6	0.46	0.39	0.30	0.30	0.42	1.00

Table B.12
Total Variance Explained: Grade 8

Factor	Initial Eigenvalues	% of Variance	Cumulative %
1	9.71	44.13	44.13
2	2.19	9.97	54.10
3	1.43	6.49	60.59
4	1.37	6.25	66.84
5	1.18	5.38	72.21
6	1.06	4.81	77.02
7	0.92	4.20	81.23
8	0.74	3.37	84.60
9	0.60	2.74	87.33
10	0.37	1.67	89.00
11	0.31	1.42	90.42
12	0.30	1.37	91.79
13	0.30	1.35	93.14
14	0.29	1.30	94.44
15	0.26	1.20	95.64
16	0.18	0.84	96.47
17	0.17	0.78	97.25
18	0.16	0.75	98.00
19	0.15	0.69	98.69
20	0.10	0.45	99.14
21	0.10	0.44	99.58
22	0.09	0.42	100.00

Table B.13
Communalities: Grade 8

	Initial	Extraction
READMC1	0.61	0.63
READMC2	0.61	0.63
READMC3	0.62	0.65
READSA1	0.47	0.63
READSA2	0.45	0.61
MATHSA1K	0.80	0.82
MATHSA1S	0.85	0.94
MATHSA1E	0.66	0.67
MATHSA2K	0.62	0.63
MATHSA2S	0.85	0.97
MATHSA2E	0.82	0.84
MATHMC1	0.79	0.78
MATHMC2	0.77	0.75
MATHMC3	0.77	0.75
F1	0.81	0.85
S1	0.76	0.79
O1	0.85	0.94
C1	0.11	0.09
F2	0.62	0.67
S2	0.70	0.75
O2	0.75	0.88
C2	0.15	0.13

Table B.14
Pattern Matrix: Grade 8

	Factor					
	1	2	3	4	5	6
READMC1	0.77					
READMC2	0.77					
READMC3	0.73					
READSA1						0.69
READSA2						0.71
MATHSA1K				0.88		
MATHSA1S				0.95		
MATHSA1E				0.69		
MATHSA2K			0.81			
MATHSA2S			1.01			
MATHSA2E			0.87			
MATHMC1	0.70					
MATHMC2	0.72					
MATHMC3	0.72					
F1		0.94				
S1		0.85				
O1		1.00				
C1						
F2					0.77	
S2					0.83	
O2					0.95	
C2						

Table B.15
Structure Matrix: Grade 8

	Factor					
	1	2	3	4	5	6
READMC1	0.79	0.39	0.43	0.36	0.46	0.41
READMC2	0.79	0.40	0.44	0.38	0.46	0.41
READMC3	0.79	0.43	0.45	0.38	0.49	0.47
READSA1	0.46	0.42	0.37	0.32	0.44	0.78
READSA2	0.40	0.41	0.35	0.31	0.42	0.78
MATHSA1K	0.48	0.29	0.43	0.90	0.31	0.30
MATHSA1S	0.51	0.32	0.46	0.97	0.34	0.32
MATHSA1E	0.50	0.36	0.48	0.80	0.38	0.39
MATHSA2K	0.43	0.28	0.79	0.34	0.32	0.27
MATHSA2S	0.51	0.35	0.98	0.43	0.38	0.33
MATHSA2E	0.52	0.37	0.91	0.47	0.41	0.38
MATHMC1	0.86	0.45	0.57	0.60	0.50	0.37
MATHMC2	0.85	0.44	0.55	0.57	0.50	0.36
MATHMC3	0.85	0.44	0.55	0.57	0.49	0.35
F1	0.43	0.92	0.34	0.27	0.51	0.41
S1	0.44	0.89	0.35	0.28	0.55	0.43
O1	0.45	0.97	0.35	0.29	0.52	0.41
C1	0.27					
F2	0.47	0.49	0.37	0.29	0.81	0.41
S2	0.48	0.52	0.37	0.30	0.86	0.45
O2	0.51	0.52	0.39	0.32	0.94	0.43
C2	0.30				0.33	

Table B.16
Factor Correlation Matrix: Grade 8

Factor	1	2	3	4	5	6
1	1.00	0.48	0.55	0.49	0.56	0.43
2	0.48	1.00	0.38	0.30	0.57	0.44
3	0.55	0.38	1.00	0.46	0.42	0.36
4	0.49	0.30	0.46	1.00	0.31	0.31
5	0.56	0.57	0.42	0.31	1.00	0.45
6	0.43	0.44	0.36	0.31	0.45	1.00

Table B.17
Total Variance Explained: Grade 10

Factor	Initial Eigenvalues	% of Variance	Cumulative %
1	9.52	43.28	43.28
2	2.43	11.04	54.32
3	1.49	6.76	61.08
4	1.17	5.32	66.40
5	1.08	4.93	71.33
6	1.05	4.77	76.10
7	0.95	4.33	80.43
8	0.68	3.10	83.53
9	0.67	3.03	86.56
10	0.40	1.83	88.40
11	0.38	1.75	90.14
12	0.35	1.60	91.74
13	0.33	1.49	93.23
14	0.31	1.43	94.66
15	0.26	1.20	95.86
16	0.25	1.14	97.00
17	0.17	0.79	97.80
18	0.16	0.73	98.52
19	0.13	0.58	99.11
20	0.08	0.37	99.47
21	0.06	0.28	99.75
22	0.05	0.25	100.00

Table B.18
Communalities: Grade 10

	Initial	Extraction
READMC1	0.54	0.56
READMC2	0.57	0.60
READMC3	0.56	0.59
READSA1	0.45	0.54
READSA2	0.49	0.65
MATHSA1K	0.88	0.88
MATHSA1S	0.89	0.96
MATHSA1E	0.59	0.59
MATHSA2K	0.89	0.89
MATHSA2S	0.90	0.95
MATHSA2E	0.61	0.61
MATHMC1	0.77	0.76
MATHMC2	0.76	0.75
MATHMC3	0.78	0.76
F1	0.58	0.61
S1	0.85	0.89
O1	0.86	0.93
C1	0.15	0.11
F2	0.67	0.71
S2	0.77	0.81
O2	0.80	0.92
C2	0.14	0.10

Table B.19
Pattern Matrix: Grade 10

	Factor					
	1	2	3	4	5	6
READMC1	0.64					
READMC2	0.72					
READMC3	0.73					
READSA1						0.61
READSA2						0.67
MATHSA1K			0.94			
MATHSA1S			1.00			
MATHSA1E			0.71			
MATHSA2K				0.93		
MATHSA2S				0.96		
MATHSA2E				0.63		
MATHMC1	0.66					
MATHMC2	0.70					
MATHMC3	0.67					
F1					0.69	
S1					0.95	
O1					1.00	
C1						
F2		0.83				
S2		0.85				
O2		1.01				
C2						

Table B.20
Structure Matrix: Grade 10

	Factor					
	1	2	3	4	5	6
READMC1	0.73	0.44	0.31	0.36	0.46	0.44
READMC2	0.77	0.42	0.32	0.37	0.45	0.42
READMC3	0.75	0.40	0.28	0.33	0.42	0.44
READSA1	0.40	0.44	0.30	0.30	0.44	0.71
READSA2	0.49	0.46	0.33	0.34	0.45	0.78
MATHSA1K	0.43	0.31	0.94	0.49	0.30	
MATHSA1S	0.43	0.32	0.98	0.50	0.31	
MATHSA1E	0.36	0.30	0.76	0.47	0.29	0.26
MATHSA2K	0.50	0.33	0.50	0.94	0.33	0.26
MATHSA2S	0.51	0.34	0.51	0.97	0.34	0.26
MATHSA2E	0.45	0.36	0.53	0.75	0.35	0.32
MATHMC1	0.82	0.46	0.56	0.60	0.46	0.27
MATHMC2	0.82	0.45	0.53	0.58	0.45	0.26
MATHMC3	0.82	0.47	0.55	0.60	0.47	0.28
F1	0.44	0.55	0.29	0.29	0.77	0.39
S1	0.47	0.60	0.31	0.32	0.94	0.43
O1	0.48	0.59	0.32	0.32	0.96	0.43
C1					0.32	
F2	0.42	0.84	0.28	0.28	0.55	0.39
S2	0.45	0.90	0.31	0.31	0.60	0.44
O2	0.45	0.96	0.30	0.31	0.58	0.41
C2		0.29			0.26	

Table B.21
Factor Correlation Matrix: Grade 10

Factor	1	2	3	4	5	6
1	1.00	0.50	0.43	0.48	0.52	0.41
2	0.50	1.00	0.31	0.31	0.65	0.45
3	0.43	0.31	1.00	0.55	0.30	0.23
4	0.48	0.31	0.55	1.00	0.30	0.23
5	0.52	0.65	0.30	0.30	1.00	0.45
6	0.41	0.45	0.23	0.23	0.45	1.00

The agreement among all four solutions is excellent. In each case, the reading and mathematics short answer items identify separate that are factorially distinct from factors that identify the multiple-choice components of the test. The writing features identify two factors at grades 5, 8, and 10, where there are two writing samples, and one factor at grade 3, where students write only one prompt. Separate structural analyses of the writing scores typically reflect differences along a narrative versus persuasive/expository axis, which, in turn, reflects substantive differences in rubrics used to score the different types of writing. The mathematics open-ended scores appear to be best resolved in terms of task-specific factors, each loading the mathematics knowledge, strategy, and explanation scores associated with the task.

APPENDIX C. PREDICTING ISAT PERFORMANCE IN SCIENCE AND SOCIAL SCIENCE FROM ISAT READING, MATHEMATICS, AND WRITING SCORES

In 2000, 4th-grade students took ISAT tests in science and social science for the first time. One year earlier, many of these students had participated in the first administration of ISAT tests of reading, mathematics, and writing.

Although Illinois does not use a common identification number to uniquely identify students in the state, it is possible to match results from one year to the next using local identification codes. A match of the 1999 and 2000 data sets by this criterion identified 32,057 students from 80 districts across the state for whom it was possible to reliably match complete grade 3 results with complete grade 4 results.

Table C.1
Correlations Among ISAT Tests

	Reading	Mathematics	Writing	Social Science	Science
Reading	1.000	.747	.512	.732	.729
Mathematics	.747	1.000	.492	.705	.692
Writing	.512	.492	1.000	.459	.448
Social Science	.732	.705	.459	1.000	.876
Science	.729	.692	.448	.876	1.000

Correlations among the tests are shown in Table C.1. The best single predictor of performance on the social science and science tests is reading, and writing, although positive correlated, has the lowest correlation with 4th-grade performance. When the 3rd-grade tests are combined the multiple correlation is .771 for social science and .764 for science. In both cases, mathematics and writing each make significant increases to the prediction of 4th-grade outcomes.

It is also possible to examine the relationship between categorical performance on the tests. For example, the data in Table C.2 shows the relationship between students' classifications on the reading test and the social science test. Rows of the table represent the 4th-grade outcome, and columns represent the 3rd-grade outcome. The first set of numbers shows the actual count of students. The second set of numbers shows the percentage of the total sample falling into each cell.

For example, 1,466 students (5% of the total data set) were classified as Academic Warning in both reading and social science. Overall, 40.3% of students in the study met (or exceeded) both standards, and 38.1% failed to meet both standards. A total of 13.7% met or exceeded the reading standard but subsequently failed to meet the social science standard, and 7.9% met or exceeded the social science standard but had failed to meet the reading standard.

Table C.2
Relationship Between Performance Classification in Reading and Social Science

	Reading Classification (N)			
Social Science Classification	Warning	Below	Meet	Exceed
Warning	1466	3141	414	33
Below	701	6907	3752	203
Meet	106	2375	8654	2861
Exceed	7	48	408	981

	Reading Classification (%)			
Social Science Classification	Warning	Below	Meet	Exceed
Warning	4.6%	9.8%	1.3%	0.1%
Below	2.2%	21.5%	11.7%	0.6%
Meet	0.3%	7.4%	27.0%	8.9%
Exceed	0.0%	0.1%	1.3%	3.1%

Table C.3 presents parallel information on the relationship between reading and science performance classifications. The outcomes are very similar. Overall, 43.5% of students in the study met (or exceeded) both standards, and 36.0% failed to meet both standards. A total of 10.5% met or exceeded the reading standard but subsequently failed to meet the science standard, and 10.1% met or exceeded the science standard but had failed to meet the reading standard.

Table C.3
Relationship Between Performance Classification in Reading and Science

	Reading Classification (N)			
Science Classification	Warning	Below	Meet	Exceed
Warning	157	189	10	3
Below	1966	9213	3193	170
Meet	137	2934	8953	2132
Exceed	20	135	1072	1773

	Reading Classification (%)			
Science Classification	Warning	Below	Meet	Exceed
Warning	0.5%	0.6%	0.0%	0.0%
Below	6.1%	28.7%	10.0%	0.5%
Meet	0.4%	9.2%	27.9%	6.7%
Exceed	0.1%	0.4%	3.3%	5.5%

Figures C.1 and C.2 present a second way of looking at the data of Tables C.2 and C.3. These figures show the probability of meeting the 4th-grade standard associated with each reading scale score value. The four arrows across the bottom of the chart delineate the four performance categories for reading. The vertical axis shows the probability of meeting (or exceeding) the 4th-grade standard for each reading score. With respect to social science, for example, a reading score of approximately 142 is associated with a probability of only .10 of meeting the 4th-grade standard. At the other end of the scale, students who exceed the

reading standard have a very high probability of meeting or exceeding the 4th-grade standard. The dark arrows toward the center of the graph indicate the point at which a student has a .50 probability of meeting the 4th-grade standard. Figure C.2 provides similar information with respect to science.

Figure C.1
Relationship Between ISAT Reading and Social Science

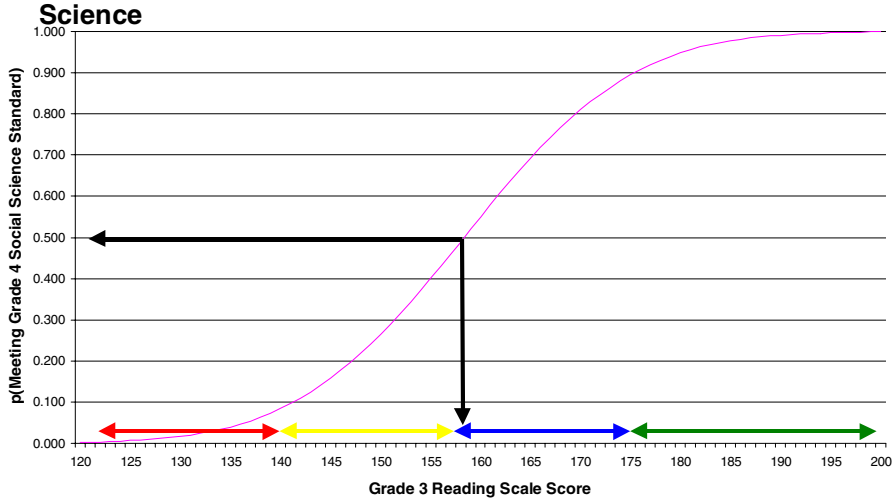


Figure C.2
Relationship Between ISAT Reading and Science