

**Kansas Assessments  
in Reading and Mathematics**

**2006**

**TECHNICAL MANUAL**

for the

Kansas Assessments of Modified Measures  
(KAMM)

*Prepared by:*

Patrick M. Irwin, Andrew J. Poggio, Xiangdong Yang,  
Douglas R. Glasnapp, and John P. Poggio

Edited by: Brooke Nash and Patrick Irwin

**Center for Educational Testing and Evaluation**  
The University of Kansas

**September 2008**

**\*This report was prepared to be included in materials submitted by the Kansas State Department of Education to satisfy requirements of the Federal No Child Left Behind Peer Review.**

## Table of Contents

Purpose of the Technical Report.....	2
Introduction and Orientation.....	3
Test Development and Content Representation.....	5
Standard Setting.....	9
Bookmark.....	9
Review Committee.....	15
Performance Level Cut scores .....	17
Reliability Analyses .....	18
Validity Analyses.....	21
References.....	39

# **The Kansas Assessments in Reading and Mathematics**

## **PURPOSE OF THE TECHNICAL REPORT**

The *Standards for Educational and Psychological Testing* (AERA/APA/NCME, 1999) requires that test developers and publishers produce a technical manual that provides information documenting the technical quality of an assessment, including evidence for the reliability and validity of test scores. This report contains the technical information for the 2006 Kansas Assessments of Modified Measures (KAMM) for grades 3-8 and high school. The information included in this report is intended for use by those who evaluate tests, interpret scores, or use test results in making educational decisions. It is assumed that the reader has some technical knowledge of test construction and measurement procedures.

Information is provided addressing the technical quality of assessments developed to measure reading and mathematics learning outcomes specific to one distinct population of Kansas students, namely students with moderate disabilities. The Kansas Assessments of Modified Measures (KAMM) are intended for administration to students with such disabilities. Thus, the body of this report addresses technical aspects focusing on scores from the KAMM tests.

## Section 1

### INTRODUCTION AND ORIENTATION

#### KAMM Test Technical Characteristics

The Kansas Assessment of Modified Measures (KAMM) is a state assessment with modified achievement standards based on grade level content standards. A student with a disability whose Individual Educational Plan (IEP) team used the KAMM eligibility criteria (available from the Kansas State Department of Education) and determined the KAMM/modified assessment is the appropriate assessment for the student may take the KAMM.

The basis for the KAMM is less complexity of test items. KAMM compares to the general assessments in reading and mathematics in that the same assessed indicators are used; however, the number of indicators assessed is reduced. Fewer multiple-choice items appear on the KAMM than on the general assessment.

In reading, the number of total items across grade level tests range from 30 (grade 3) to 49 (grade 11) items. At all levels, there are fewer passages to read. At grades 3 and 4, there are two narrative and two expository passages. At grades 5, 6, and 7, there are two narrative, two expository, and one technical passage. At grades 8 and HS, there are two narrative, two expository, one technical, and one persuasive passage.

In mathematics, all grade level KAMM forms have 40 items, and calculator use was allowed for all portions of the assessment. Students were allowed to take the KAMM over as many days as necessary.

Accommodations are allowed on the KAMM as IEP teams make decisions about accommodations for the KAMM the same as they do for the general assessment. Additionally, the KAMM is primarily a computer-delivered assessment. For students who cannot complete the KAMM assessment online, a paper-and-pencil accommodation can be made.

The population of students taking the KAMM did not result in a sufficient number of students within subgroups to conduct empirical Differential Item Function (DIF) analyses. However, items were reviewed by KSDE panels for bias, offensiveness, and insensitivity.

As background, new KAMM assessments in reading and mathematics were planned and developed, then administered for the first time in Spring 2006. The Kansas State Department of Education (KSDE) designed and developed the KAMM. Students with moderate disabilities in grades 3-8 (reading and mathematics), 10 (mathematics), and 11 (reading) participated in the assessments.

The Spring 2006 administration of the KAMM serves as the baseline for the new cycle of state assessments. The assessments administered were all newly developed to measure the new targeted indicators (learning outcomes) in the most recent editions of the Kansas Curricular Standards for the content areas. These documents should be referenced when examining and evaluating any of the information resulting from the state assessment

programs. The Curricular Standards serve as the basis for what is assessed by the tests, and any interpretation and subsequent action based on student or group performance on these tests must focus on the assessed standards, benchmarks, and indicators. Copies of the Kansas Curricular Standards in the content areas are available from the KSDE website at [www.ksde.org](http://www.ksde.org).

As the baseline year of the new round of assessments, the Spring 2006 administration incorporated important changes from prior KAMM assessments administered in the 2000 – 2005 testing cycle. Curriculum standards and targets for the assessments were changed, test specifications revised, and assessed grade levels expanded to include students in grades 3-8 and one grade level in high school. In effect, no comparison to past student, building, district, or state performance should be made.

KAMM assessments were developed by the state and made available in lieu of the general assessments for administration to students with moderate disabilities. Single forms measuring targeted learning outcomes in reading and mathematics were developed for each grade level, 3 – 8 and high school. The item format used in these assessments was multiple-choice with three response options.

The KAMM assessments are planned and created to reflect and otherwise operationalize certain grade level learning outcomes that should serve as curricular and instructional targets in Kansas K-12 schools. As in previous years, the assessments contribute information for ongoing school accreditation status determinations, and results from the reading and mathematics assessments have a primary role in monitoring annual yearly progress (AYP) as part of the federally mandated No Child Left Behind assessment requirements. As related to the accountability demands, cutscores on each test were determined to classify students into one of five performance categories (Exemplary, Exceeds Standard, Meets Standard, Approaches Standard, and Academic Warning). The proportion of students classified in these categories becomes a primary source of information in determining AYP for schools, districts, and state. Section 3 provides additional details about procedures for setting the specific test score criteria used to classify students into one of the five performance categories.

## Section 2

### TEST DEVELOPMENT AND CONTENT REPRESENTATION

The content of the KAMM is derived from the Kansas Curricular Standards (see <http://www.ksde.org/Default.aspx?tabid=1678> for the Curricular Standards in all subject areas). These Curricular Standards define, for Kansas schools, what students should know and be able to do in the respective content domains at each grade level. The 2006 KAMM measured targeted indicators in the Curricular Standards for reading and mathematics in grades 3-8 and high school (grade 10 for mathematics and grade 11 for reading).

#### Test Specifications

Test specifications provide the blueprint to be followed when writing items and constructing test forms. KSDE developed and provided the test specifications that guided all item and test development efforts. Test specifications were provided in matrix form that identified, by cognitive complexity level and targeted indicators (skill) to be assessed, the number and distribution of items to be on each test form at a grade level. These grade level and content area specifications guided the construction of operational forms development, but the order and manner in which items were placed throughout the forms was left to the KSDE content specialists. The most recent versions of the test specifications can be obtained through the KSDE website at <http://www.ksde.org/Default.aspx?tabid=2371>.

#### Item Type

The multiple choice item type is utilized exclusively on KAMM in reading and mathematics. For all multiple choice items appearing on any KAMM form, students select the one best answer from among three choices provided.

#### Item Development

Beginning with the 2006 assessment cycle, KSDE contracted with WestEd, to supply reading and mathematics items that were aligned with content area Curricular Standards.

The final rounds of item pool reviews involved content review and fairness review committees comprised of Kansas educators. Along with KSDE specialists, the content committees reviewed each item, focusing on its alignment to the table of specifications and to the Kansas Curricular Standards, as well as the appropriateness of item content, ensuring that each item accurately reflected what was intended to be taught in Kansas schools. The fairness review committees focused on language and content that might be inappropriate, offensive, or insensitive to students, parents, or communities, making sure that no individual or group would be unfairly favored or disadvantaged due to the content of the items. With both review committees, each item was accepted, edited, or rejected from its respective item pools.

### KAMM Reading Summary Statistics

Table 2.1 reports summary findings for the KAMM Reading Assessments by grade level, identifying the number of items per KAMM form, the number of students administered at a particular grade level KAMM, reliability coefficients, and descriptive statistics in terms of raw total scores and percent correct total scores. All of the reliability coefficients are greater than 0.85. The percent mean correct ranged from 57.92% (SD = 17.398) at grade 7 to 76.73% (SD = 18.147) at grade 3. Table 2.2 combines the 2006 and 2007 data files in order to have stable estimates and sufficient sample sizes to reports reliability coefficients for the KAMM Reading Assessment by demographic (specifically by race), students categorized as English language learners (ELL), and students on free and reduced lunch programs.

Table 2.1  
*KAMM Reading Summary Statistics*

Grade	Test ID	# of Items	N	Reliability ( $\alpha$ )	Mean Raw Score	SD of Mean Score	Mean Percent Correct	SD of Mean Percent Correct
3	745	30	731	0.86	23.02	5.44	76.73	18.15
4	747	37	882	0.88	26.48	6.92	71.58	18.70
5	748	45	952	0.89	31.24	8.33	69.43	18.52
6	749	47	1053	0.86	30.22	7.79	64.30	16.58
7	750	46	976	0.86	26.64	8.00	57.92	17.40
8	751	48	1024	0.86	30.02	8.56	62.54	17.84
11	752	49	730	0.90	33.55	9.04	68.48	18.45

Table 2.2  
*Reliability Coefficients for 2006-2007 Reading KAMM Assessments Based on Demographic*

Grade	N of Students	N of Items	Reliability ( $\alpha$ )	Race						Free/Reduced Lunch			
				Hispanic		African-American		Caucasian		ELL		N	$\alpha$
				N	$\alpha$	N	$\alpha$	N	$\alpha$	N	$\alpha$	N	$\alpha$
3	1,776	30	0.86	235	0.85	273	0.87	1,135	0.85	165	0.82	1,205	0.85
4	1,921	37	0.87	271	0.86	327	0.86	1,188	0.86	203	0.86	1,314	0.86
5	2,105	45	0.89	348	0.88	336	0.88	1,249	0.89	255	0.88	1,423	0.89
6	2,152	47	0.86	348	0.85	377	0.84	1,280	0.86	228	0.84	1,419	0.86
7	2,118	46	0.86	347	0.83	387	0.83	1,267	0.87	208	0.83	1,418	0.85
8	2,149	48	0.87	308	0.87	398	0.85	1,302	0.88	155	0.85	1,389	0.86
10	1,471	49	0.90	149	0.88	239	0.88	1,002	0.91	40	N/A	869	0.90

### KAMM Mathematics Summary Statistics

Table 2.3 reports summary findings for the KAMM Mathematics Assessments by grade level, identifying the number of items per KAMM form, the number of students administered a particular grade level KAMM, reliability coefficients, and descriptive statistics in terms of raw total scores and percent correct total scores. The reliability coefficients for grades 3 through 6 are above .85 and for grades 7 and 8, above .80. The 10<sup>th</sup> grade KAMM Mathematics Assessment has the lowest reliability coefficient ( $\alpha = 0.750$ ). The percent mean correct ranges from 47.42% (SD = 14.722) at grade 10 to 69.65% (SD = 18.224) at grade 3. Table 2.4 combines the 2006 and 2007 data files in order to have stable estimates and sufficient sample sizes to reports reliability coefficients for the KAMM Mathematics Assessment by demographic (specifically by race), students categorized as English language learners (ELL), and students on free and reduced lunch programs.

Table 2.3  
*KAMM Mathematics Summary Statistics*

Grade	Test_ID	# of Items	N	Reliability ( $\alpha$ )	Mean Raw Score	SD of Mean Score	Mean Percent Correct	SD of Mean Percent Correct
3	736	40	541	0.872	27.86	7.290	69.65	18.224
4	735	40	740	0.865	26.35	7.370	65.88	18.425
5	729	40	825	0.865	26.29	7.228	65.73	18.070
6	728	40	965	0.861	24.93	7.390	62.33	18.476
7	738	40	998	0.818	23.00	6.629	57.51	16.572
8	737	40	1060	0.809	22.00	6.724	64.17	18.057
10	730	40	843	0.750	18.97	5.889	47.42	14.722

Table 2.4  
*Reliability Coefficients for 2006-2007 Mathematics KAMM Assessments Based on Demographic*

Grade	N of Students	N of Items	Reliability ( $\alpha$ )	Race						Free/Reduced Lunch			
				Hispanic		African-American		Caucasian		ELL		N	$\alpha$
				N	$\alpha$	N	$\alpha$	N	$\alpha$	N	$\alpha$	N	$\alpha$
3	1,316	40	0.87	174	0.87	228	0.87	813	0.87	127	0.85	915	0.87
4	1,568	40	0.85	210	0.83	302	0.83	938	0.85	148	0.85	1,090	0.85
5	1,833	40	0.85	283	0.84	312	0.86	1,101	0.85	197	0.85	1,225	0.85
6	1,931	40	0.86	249	0.86	396	0.81	1,160	0.87	155	0.88	1,262	0.85
7	2,064	40	0.81	271	0.82	421	0.81	1,267	0.81	168	0.80	1,380	0.81
8	2,203	40	0.81	272	0.81	421	0.75	1,364	0.82	137	0.81	1,384	0.80
10	1,209	40	0.74	113	0.54	217	0.61	805	0.75	29	N/A	705	0.73



## **Test Form Development**

KSDE was responsible for the development of the KAMM assessments. Items for the KAMM were selected or modified based on cognitive load. For reading, newly developed passages and items independent of those appearing on general reading assessment forms were selected. For mathematics, the Mathematics Revision Committee and the Special Education Assessment Advisory Committee, in cooperation with KSDE, were given the task of identifying the indicators to be included on the KAMM. In addition to defining which indicators would appear on the grade level tests, it was necessary to determine the number and kind of items to be included from the general mathematics assessment. Four items from the general mathematics assessments were selected and modified per grade level indicator to be assessed on KAMM (10 assessed indicators per KAMM mathematics form).

For the Spring 2006 administration, all operational test forms were administered via KCA. Following the administration of the first operational forms of the KAMM Reading and Mathematics Assessments in Spring 2006, analysis work commenced, employing classical and IRT methods. Traditional item analysis studies were conducted on each test form to reconfirm the pilot test results that items selected for operational use were functioning adequately and as expected. Before Annual Yearly Progress (AYP) reporting could occur, standard setting activities needed to be implemented to establish score ranges on the tests that would define levels of test score performance needed for students to be classified into one of the five performance level categories established by the state (Exemplary, Exceeds Standard, Meets Standard, Approaches Standard, and Academic Warning). See Section 3 for descriptions of the standard setting activities implemented. Based on the standards recommended by KSDE and approved by the Kansas State Board of Education, final results for the KAMM Reading and Mathematics Assessments were reported in September 2006.

## **Section 3**

### **STANDARD SETTING**

Standard Setting for the 2006 KAMM was conducted in late spring through summer 2006. Standard Setting for the 2006 Kansas General Reading and Mathematics Assessments was also conducted at this time. CETE implemented various standard setting methodologies to recommend performance level standards for the KAMM Reading and Mathematics Assessments. For the KAMM tests, the Bookmark standard setting procedure was used to recommend cut scores.

This section documents the standard setting processes conducted for each KAMM assessments. The first part provides an overview of the Bookmark standard setting workshop for reading and mathematics, including a description of the participants, the process and procedures, and evidence of the successfulness of the activity. The second part details the review committee panels that synthesized the judgments made by educators in the standard setting processes and made final recommendations to KSDE.

#### **Bookmark Standard Setting**

The Bookmark standard setting procedure (Mitzel, Lewis, Patz, & Green, 1996) was implemented to recommend performance level cut scores for the KAMM assessments for the content area tests for grades 3-8 (mathematics and reading), grade 10 (mathematics), and grade 11 (reading). The Bookmark procedure consisted of training, orientation, and three rounds of judgments by participants. The Bookmark standard setting for the KAMM Mathematics Assessments was held June 19 and 20, 2006, in Salina, Kansas. The standard setting for the KAMM Reading Assessments was held June 20 and 21, 2006, in Salina, Kansas. Each content area standard setting lasted two days, with one-half day devoted to table leader training and approximately one-and-one-half days devoted to the actual procedure.

The Bookmark standard setting was conducted to recommend cut scores on each grade-level KAMM test that separate students into five performance levels defined by the state of Kansas: Unsatisfactory, Basic, Proficient, Advanced, and Exemplary (performance levels were later renamed: Academic Warning, Approaches Standard, Meets Standard, Exceeds Standard, and Exemplary).

#### **Bookmark Standard Setting Roles**

##### **CETE Staff**

CETE worked with staff from KSDE to design and organize the standard setting. The CETE standard setting team was comprised of John Poggio, Ph.D., Doug Glasnapp, Ph.D., Patrick Irwin, Ph.D., and Andrew Poggio. Dr. Poggio and Dr. Glasnapp are co-directors of CETE. Dr. Irwin and Andrew Poggio are research assistants for CETE. During the standard setting, John Poggio, Patrick Irwin, and Andrew Poggio were responsible for facilitating the standard setting meeting, training

participants, and monitoring the participant results database. Prior to the meeting, the entire CETE group prepared all materials related to the standard setting.

Additionally, four graduate students working for CETE on assistantships attended the meeting and served as technical staff, working mainly in the operations room and taking responsibility for the entry of the judgment data into the optical scanner. These graduate students also implemented quality assurance procedures, such as double-checking the data. CETE support staff also played an integral role in planning and facilitating the standard setting meeting. Amy Tackett, Ronda Consolver, and Wendy Coonrod were primary points of contact with KSDE and assumed a variety of roles before and during the standard setting meeting, including creating the registration process and conducting it on-site.

### **Table Leaders**

In total, there were 14 tables of KAMM Bookmark standard setting panelists across the two content areas (one table per grade level). Every table had a table leader who was chosen prior to the standard setting. Table leaders were chosen based on years of experience as grade-level educators as well as by information received through the participant nomination process. Table leaders were active voting participants in the standard setting process. They received additional training on the Bookmark method in a session prior to the arrival of the rest of the participants (duration: approximately four hours). The primary role of the table leaders was to lead table discussions, monitor group discourse, facilitate discussion, collect materials, and maintain the schedule.

### **Participants**

Participants in the KAMM Bookmark standard setting meeting were selected by CETE from a pool of nominees. These nominations were made by Kansas educators (e.g., administrators, school faculty, and district coordinators). The participants were selected based on factors such as grade and content area of primary instruction, geographic location, school size, and years of experience teaching in Kansas. Every effort was made to select a representative group of participants. There was one table per content area, which was balanced as well as possible in terms of relevant demographic characteristics. There were forty four KAMM reading participants and thirty nine KAMM mathematics participants. Participants completed an evaluation following the standard setting. Using these evaluations, participants' demographic information was summarized, specifically the academic position (administrator, coordinator, or classroom teacher), teaching experience, district location, and building status (in terms of SES) of the panelists.

## Bookmark Materials

### Ordered Item Booklets

The Ordered Item Booklets (OIBs) for each grade level and content area were comprised of items from an operational form of the respective assessment administered for the first time in Spring 2006. For the general reading and mathematics assessments, the form used to create the OIBs was the base form; that is, the form that was available for computerized and paper-and-pencil delivery. For the KAMM forms, the single operational form in existence was used to create the OIBs. The items for each grade level assessment form utilized were ordered in terms of difficulty using a 2-parameter logistic (2PL) Item Response Theory (IRT) model. The OIBs were ordered from the easiest to the hardest item. Table 3.1 summarizes the number of score points in each OIB by test and content area.

Table 3.1  
*Number of Score Points in Ordered Item Booklets by Test Form and Content Area*

Level	Test/Content Area	Number of Score Points in OIB
<b>Grade 3</b>	KAMM Reading	30
	KAMM Mathematics	40
<b>Grade 4</b>	KAMM Reading	37
	KAMM Mathematics	40
<b>Grade 5</b>	KAMM Reading	45
	KAMM Mathematics	40
<b>Grade 6</b>	KAMM Reading	47
	KAMM Mathematics	40
<b>Grade 7</b>	KAMM Reading	46
	KAMM Mathematics	40
<b>Grade 8</b>	KAMM Reading	48
	KAMM Mathematics	40
<b>Grade 10</b>	KAMM Mathematics	40
<b>Grade 11</b>	KAMM Reading	49

### Standard Setting Day 1, Mathematics: Table Leader Training and Orientation

Table leaders were trained on the morning of the first day of the Kansas Bookmark standard setting for the general and KAMM Mathematics assessments. During this session, which lasted approximately four hours, table leaders were given an overview of the standard setting process and were trained specifically on the Bookmark method. The table leader training stressed the type of behavior expected of table leaders during the Bookmark standard setting.

### Standard Setting Day 1, Mathematics: Participant Training and Orientation

CETE welcomed the panelists to the KAMM assessments standard setting in the afternoon of Day 1. All participants were checked in and given a packet of materials that included the PowerPoint slides used during the training session. The meeting took place in a large conference space in Salina, Kansas, a central location in the state that allowed for participation from educators from areas across the state. All of the tables were located in one large room. Upon registration and

check-in, participants made their way to their tables, and the training session commenced. John Poggio of CETE ran the training session, which lasted approximately three-and-one-half hours. The orientation and training consisted of a brief overview of the Kansas testing program followed by an overview of the standard setting process. A description of the review procedures that would follow the Bookmark meeting was also provided to panelists. Participants were trained on the use of the OIBs during the orientation session as well.

A majority of the orientation session was devoted to training on the Bookmark method, specifically about how to place a bookmark. In the training materials provided, several explanations of bookmark placement were described. It was explained that for the Kansas assessments, four cut scores or bookmarks would need to be set to provide for the five performance level categories: Unsatisfactory, Basic, Proficient, Advanced, and Exemplary. The mechanics of bookmark placement were then described, with an explanation that all items preceding the bookmark define the knowledge, skills, and abilities that an Advanced student, for example, is expected to have. Participants were instructed to examine each item in terms of its content and the knowledge and skill requirements of the item and make a judgment about the type of content a student would need to know in order to be considered barely Advanced.

The final topic covered in the training session was using a response probability to determine an item's association with a given performance category. This explanation linked the use of the  $2/3$  or 0.67 mastery probability to bookmark placement.

## **Round 1**

All participants were mailed materials to be reviewed prior to the Bookmark standard setting meeting. Panelists were instructed to review and consider Kansas' Performance Level Descriptors (PLDs) that define the five performance categories. Round 1 began with an opportunity for tables to review and discuss the PLDs. Then, participants spent approximately one-and-a-half hours taking the operational grade level test in mathematics via the OIB. Panelists were asked to familiarize themselves with the test, specifically studying each item in terms of what each item measured; the knowledge, skills, and abilities required by the item; and why each item is more difficult than the items preceding it. The panelists were directed to place their bookmarks, starting with Unsatisfactory, keeping in mind the PLDs. Participants were instructed during training and reminded by the table leaders that bookmark placement is an individual activity. A dual-mode strategy for bookmark recording was utilized. For efficient and accurate database entry, a scannable recording form was developed by CETE and used for this meeting. Panelists were asked to bubble the bookmark placement for the appropriate performance category on the rating form. Also, a traditional hand-recording form was collected from each panelist as they recorded their judgments round by round. This form was a backup that was also used for accuracy checks. After the appropriate Round 1 recordings were made, the meeting adjourned for the day.

## **Standard Setting Day 2, Mathematics: Rounds 2 and 3**

### **Round 2**

Round 2 commenced with Round 1 results returned to each table. Each participant was provided with the minimum, maximum, and median cut score recommendation for each performance category for the table. Panelists were asked to discuss those items that fell between the first and last placed bookmark per performance category; that is, they were instructed to discuss those items for which there was disagreement among the group. After discussion, the participants were asked to place their Round 2 bookmarks. Again, participants were reminded that bookmark placement is an individual activity.

### **Round 3**

At the beginning of Round 3, descriptive results (minimum, maximum, and median) regarding cut score recommendations for each performance category were provided to each table. Relevant cumulative frequency distributions were administered to each table as well. John Poggio of CETE addressed the participants, presenting a grade-level sample of aggregated impact data based on the Round 2 bookmarks. It was emphasized to participants that the impact data was being presented as a reality check. The table leaders then facilitated discussion among the panelists on their bookmark placements. CETE staff answered process-related questions, while KSDE staff was present and available for any policy-related questions concerning the impact data that arose. After discussion, panelists placed their final bookmarks. Participants were again reminded that bookmark placement is an individual activity. Upon completion of the Round 3 bookmark placements, the participation of the panelists in the Bookmark standard setting for the KAMM Mathematics assessments had concluded. Prior to leaving the meeting, participants were asked to complete an evaluation of the standard setting meeting.

The Bookmark standard setting meeting for the KAMM Mathematics and Reading Assessments took place over three days, in June 2006 (June 19-21), with mathematics occurring on June 19 and 20 and reading on June 20 and 21. In essence, as the standard setting for mathematics was finishing, the participants for the reading standard setting were arriving. The methodology and procedures followed for both content areas were almost exclusively the same. Unless noted, the process for the Bookmark standard setting for reading was the same as the process for mathematics. Additional detail will be provided below only where there were deviations between the two implementations.

## **Standard Setting Day 1, Reading: Table Leader Training and Orientation**

### **Training**

Table leaders for reading were trained on the morning of the first day of the reading standard setting meeting, which coincided with the second day of the mathematics standard setting meeting. Andrew Poggio of CETE led the Table Leader training session. The presentation was identical to the orientation for the mathematics table leaders and lasted approximately four hours during which

the participants were given an overview of the standard setting process and trained specifically on the Bookmark method. They were given a synopsis of the meeting activities as well as their responsibilities and tasks. The table leader training stressed the type of behavior expected of table leaders during the Bookmark standard setting.

## **Standard Setting Day 1, Reading: Participant Training and Orientation**

### **Orientation**

Participants were welcomed to the Kansas reading general and KAMM assessments standard setting in the afternoon of day 1. The process and procedures for registration, orientation, and training for reading panelists were identical to those for the mathematics participants described above.

### **Round 1**

The same procedures were implemented for Round 1 of reading as Round 1 of mathematics. The PLDs for reading were sent to all participants prior to the standard setting meeting as with mathematics. The PLDs were reviewed and discussed by participants prior to taking the grade-level reading assessment. Participants were reminded that bookmark placement is an individual activity, and they were then asked to place their Round 1 bookmarks. The standard setting meeting adjourned for the day after this activity was completed.

## **Standard Setting Day 2, Reading: Rounds 2 and 3**

### **Round 2**

The same process was followed for Round 2 of the reading standard setting as the process for Round 2 of mathematics.

### **Round 3**

The same process was followed for Round 3 of the reading standard setting as the process for Round 3 of mathematics. Upon completion of the Round 3 bookmark placements, the participation of the panelists in the Bookmark standard setting for the Kansas Reading Assessments concluded. Prior to leaving the meeting, participants were asked to complete an evaluation of the standard setting meeting.

## **Evaluations of the Kansas Bookmark Standard Settings**

### **Quality Control Procedures**

CETE adhered to several quality control procedures to foster accuracy of the standard setting materials and the results presented during the standard setting. Prior to the meeting, the ordering of the items in the OIBs was checked, along with the accuracy and completeness of test information, training materials, and impact data tables. During the meeting, data were collected by a

dual-mode recording strategy. Data were entered via the scannable form and then checked for accuracy by CETE staff. Results from each round computed by SPSS were checked for accuracy before being returned to tables. Any results that appeared to be questionable were investigated further.

### **Effectiveness of Training**

At the end of the respective standard setting activities, participants were asked to respond to an evaluation of the standard setting meeting. Evidence of the successfulness of the standard setting may be found in the participants' self-reported understanding of training, perceived validity of the activity, and confidence in their bookmark placements. During the KAMM assessments standard setting, participants were trained during the orientation session before placing bookmarks. Table leaders spent an extra half-day in training before the start of the workshop.

The majority of participants indicated that the training was adequate in preparation for the judgment tasks. The majority of participants also agreed that the training was useful in the standard setting.

### **Perceived Validity**

Another indication of the successfulness of the standard setting may be found in the participants' perceived validity of the meeting. A majority of the participants understood the purpose of the standard setting and were confident that the process and methods used in the standard setting can contribute to valid and reliable cut score recommendations.

## **Standard Setting Review Committee Meeting**

A panel of participants was convened to examine all of the recommended cut scores for the KAMM assessments at each grade level and to synthesize the information in an attempt to make a final recommendation to KSDE. The KAMM review committees were presented the cut score recommendations from the Bookmark standard setting. The review committee was also presented with an explanation of the method(s) implemented, recommended cut scores for the assessments from the various methods used, student performance or impact data, and, where available, participant information and standard setting evaluation results.

The KAMM mathematics and reading cut score recommendations were each reviewed by a unique review committee. For the KAMM mathematics and reading, the respective review committees reviewed the cut score recommendations at all grade levels. The KAMM review committees met on August 2, 2006, in Lawrence, Kansas.



## **Review Committee Meetings**

The review committee meetings were conducted by CETE with representatives from KSDE in attendance. Dr. John Poggio and Dr. Doug Glasnapp of CETE directed the meetings with Dr. Patrick Irwin, Andrew Poggio, and Dr. Jonathan Templin, CETE research staff, assisting.

The review committee meetings began with an orientation that introduced the task at hand and the goals of the meetings. Andrew Poggio conducted the orientation for the KAMM assessment review committee panels. Participants were provided with an explanation of the standard setting methods used to collect the cut score recommendations they would be reviewing, information on the panelists that provided the scores and evaluation results where available, as well as the Kansas Performance Level Descriptors (PLDs).

After the orientation, the review committee panels reviewed all of the information provided for a given assessment. For the KAMM assessments, the specific assessment committees convened after the group orientation. After reviewing the cut score recommendations from the various standard setting methods previously implemented, considering impact data, reviewing any other information provided for a given assessment, and engaging in a group discussion of the information provided, the panels were asked to make individual cut score recommendations for the performance categories on a given test at each grade level being considered. Each review committee member's individual cut score recommendations were visually displayed to the panel. The range of the individual cut score recommendations served as the basis for a subsequent round of discussion where the goal was to determine a single cut score recommendation at each performance category, or to at least reduce the range at each performance level. In most cases, after group discussion, a single score point was recommended, while in a few cases, a very small range of scores was selected. This was done for each of the performance categories within a grade level. The panel members then considered the next grade level and repeated the process of reviewing materials, making individual recommendations, engaging in group discussion, and then coming to group consensus. After reviewing each of their assigned assessments, the review committee members were presented each recommended cut score across all grade levels visually and were asked to consider and discuss these scores further, if necessary, and make any final adjustments. The final cut score recommendations from the review committees were given to KSDE along with impact data. KSDE made the cut score recommendations to the Kansas State Board of Education and were subsequently approved. The cut scores are provided below.

**KAMM ASSESSMENTS  
 PERFORMANCE LEVEL CUTSCORES**

**Kansas Assessment of Modified Measures (KAMM) Mathematics  
 % correct**

<b>Grade</b>	<b>Academic Warning</b>	<b>Approaches Standard</b>	<b>Meets Standard</b>	<b>Exceeds Standard</b>	<b>Exemplary</b>
3 <sup>rd</sup>	0-34	35-56	57-81	82-94	95-100
4 <sup>th</sup>	0-34	35-54	55-81	82-94	95-100
5 <sup>th</sup>	0-34	35-56	57-81	82-91	92-100
6 <sup>th</sup>	0-34	35-54	55-79	80-91	92-100
7 <sup>th</sup>	0-34	35-51	52-74	75-86	87-100
8 <sup>th</sup>	0-34	35-49	50-74	75-84	85-100
High School	0-31	32-44	45-64	65-79	80-100

**Kansas Assessment of Modified Measures (KAMM) Reading  
 % correct**

<b>Grade</b>	<b>Academic Warning</b>	<b>Approaches Standard</b>	<b>Meets Standard</b>	<b>Exceeds Standard</b>	<b>Exemplary</b>
3 <sup>rd</sup>	0-39	40-57	58-88	89-93	94-100
4 <sup>th</sup>	0-38	39-54	55-84	85-90	91-100
5 <sup>th</sup>	0-37	38-51	52-82	83-90	91-100
6 <sup>th</sup>	0-38	39-47	48-75	76-83	84-100
7 <sup>th</sup>	0-34	35-40	41-71	72-82	83-100
8 <sup>th</sup>	0-35	36-46	47-76	77-85	86-100
High School	0-37	38-48	49-81	82-87	88-100

## Section 4

### RELIABILITY ANALYSES

#### Score Reliability

#### KAMM Performance Classification Reliability

Score ranges on the KAMM assessments were established through standard setting procedures to classify students into five performance level categories (Academic Warning, Approaches Standard, Meets Standard, Exceeds Standard, and Exemplary). Procedures to estimate classification consistency and accuracy for the KAMM assessments mirrored those used for the general assessment test forms. Tables 4.1 and 4.2 present summaries of the classification consistency and classification accuracy indices for the KAMM test forms across testing grade and content area. Included in the tables is information on the overall test classification reliability and related information for dichotomous decisions at three cut-points, one for Academic Warning versus all levels above, one for Exemplary versus all levels below, and one for the most important AYP reporting decision, i.e., the bottom two categories versus the upper three categories, the latter three categories defining performance judged to be acceptable.

For reading (Table 4.1), overall test classification consistency values across all categories range from 0.56 (grade 3) to 0.64 (grade 5). Classification accuracy, on the other hand, is consistently higher than classification consistency yet maintains the same pattern of variability across grades. Classification accuracy coefficients range from 0.69 (grade 5) to 0.80 (grade 10). For mathematics (Table 4.2), classification consistency values range from 0.45 (grade 10) to 0.63 (grade 4). As in reading, classification accuracy is consistently higher than classification consistency, yet maintains the same pattern of variability across grades. Classification accuracy coefficients range from 0.56 to 0.73. These coefficients are slightly lower than those reported for the general assessments, likely due to the fewer number of items on the test forms.

For both mathematics and reading, the reliabilities of classification at a given cut point are generally high whereas probabilities of misclassifications are low. For reading, all classification accuracy coefficients for the important AYP decision (levels 12 versus 345) are equal to or above 0.90. For mathematics, these coefficients are all equal to or above 0.85. The classification consistency values are slightly lower. These values support the adequacy of the KAMM assessments for making the major decision associated with AYP reporting.

Table 4.1  
*KAMM Classification Reliability Indices by Cut Points in Reading*

Grade	Cut Point*	Classification Accuracy	Classification Consistency	False Positive	False Negative
3	Overall	0.66	0.56		
	1 / 2345 <sup>†</sup>	0.97	0.96	0.01	0.02
	12 / 345	0.93	0.90	0.03	0.04
	123 / 45	0.87	0.82	0.07	0.06
	1234 / 5	0.87	0.84	0.08	0.08
4	Overall	0.71	0.62		
	1 / 2345	0.96	0.94	0.01	0.03
	12 / 345	0.92	0.89	0.03	0.05
	123 / 45	0.90	0.86	0.05	0.05
	1234 / 5	0.91	0.87	0.06	0.03
5	Overall	0.73	0.64		
	1 / 2345	0.96	0.95	0.02	0.02
	12 / 345	0.93	0.90	0.03	0.05
	123 / 45	0.91	0.87	0.04	0.05
	1234 / 5	0.92	0.89	0.05	0.02
6	Overall	0.7	0.61		
	1 / 2345	0.95	0.93	0.01	0.04
	12 / 345	0.92	0.88	0.03	0.05
	123 / 45	0.9	0.85	0.06	0.04
	1234 / 5	0.93	0.90	0.04	0.03
7	Overall	0.71	0.63		
	1 / 2345	0.93	0.91	0.02	0.04
	12 / 345	0.9	0.86	0.04	0.06
	123 / 45	0.91	0.87	0.06	0.03
	1234 / 5	0.94	0.92	0.04	0.01
8	Overall	0.71	0.62		
	1 / 2345	0.95	0.92	0.01	0.04
	12 / 345	0.92	0.88	0.03	0.05
	123 / 45	0.9	0.86	0.06	0.04
	1234 / 5	0.94	0.91	0.03	0.03
11	Overall	0.71	0.62		
	1 / 2345	0.95	0.92	0.01	0.04
	12 / 345	0.92	0.88	0.03	0.05
	123 / 45	0.9	0.86	0.06	0.04
	1234 / 5	0.94	0.91	0.03	0.03

\* 1 = Academic Warning, 2 = Approaches Standard, 3 = Meets Standard, 4 = Exceeds Standard, 5 = Exemplary

Table 4.2  
*KAMM Classification Reliability Indices by Cut Points in Mathematics*

Grade	Cut Point*	Classification Accuracy	Classification Consistency	False Positive	False Negative
3	Overall	0.72	0.62		
	1 / 2345 <sup>†</sup>	0.97	0.96	0.01	0.02
	12 / 345	0.91	0.88	0.04	0.05
	123 / 45	0.9	0.86	0.06	0.04
	1234 / 5	0.94	0.92	0.05	0.01
4	Overall	0.73	0.63		
	1 / 2345	0.96	0.94	0.01	0.03
	12 / 345	0.9	0.86	0.05	0.05
	123 / 45	0.91	0.88	0.05	0.03
	1234 / 5	0.96	0.94	0.04	0.00
5	Overall	0.72	0.61		
	1 / 2345	0.96	0.95	0.01	0.03
	12 / 345	0.9	0.85	0.05	0.05
	123 / 45	0.91	0.97	0.06	0.06
	1234 / 5	0.95	0.93	0.04	0.04
6	Overall	0.71	0.61		
	1 / 2345	0.94	0.92	0.02	0.04
	12 / 345	0.89	0.84	0.06	0.05
	123 / 45	0.92	0.88	0.06	0.03
	1234 / 5	0.96	0.95	0.03	0.01
7	Overall	0.69	0.57		
	1 / 2345	0.95	0.91	0.00	0.05
	12 / 345	0.85	0.79	0.09	0.06
	123 / 45	0.93	0.90	0.05	0.02
	1234 / 5	0.97	0.95	0.02	0.01
8	Overall	0.67	0.56		
	1 / 2345	0.91	0.87	0.00	0.08
	12 / 345	0.85	0.79	0.09	0.06
	123 / 45	0.93	0.91	0.05	0.02
	1234 / 5	0.96	0.95	0.03	0.01
10	Overall	0.56	0.45		
	1 / 2345	0.84	0.78	0.00	0.16
	12 / 345	0.79	0.71	0.15	0.06
	123 / 45	0.94	0.91	0.04	0.02
	1234 / 5	0.98	0.97	0.02	0.01

\* 1 = Academic Warning, 2 = Approaches Standard, 3 = Meets Standard, 4 = Exceeds Standard, 5 = Exemplary

## **Section 5**

### **EVIDENCE FOR THE VALIDITY OF INFERENCES FROM TEST SCORES**

#### **INTERNAL EVIDENCE FOR THE VALIDITY OF KANSAS ASSESSMENT OF MODIFIED MEASURES (KAMM) ASSESSMENT SCORES**

##### **Indicator Intercorrelations**

Validity is one of the most important attributes of assessment quality. It refers to the appropriateness or correctness of inferences, decisions, or descriptions made from test results about what students know and can do and is one of the fundamental considerations in developing and evaluating tests (AERA/APA/NCME, 1999). It is a multidimensional construct that resides not in tests, but in the relationships between any test score and its context (including the instructional practices and the examinee), the knowledge and skills it is to represent, the intended interpretations and uses, and the consequences of its interpretation and use. Therefore, validity is not based on a single study or type of study but instead should be considered an ongoing process of gathering evidence supporting every intended interpretation and use of the scores resulting from a measurement instrument. As validity is not a property of a test or a test score or even of an interpretation, inference, or use of a test score, it cannot be captured conclusively. Rather, a judgment must be made regarding whether a body of evidence supports specific test claims and uses. This process begins with the test design and continues throughout the entire assessment process, including design, content specifications, item development, psychometric quality, and inferences made from the results.

The analysis of test structure based on examinee response data is an important part of test development and evaluation. A simplistic approach to conducting such an analysis is to determine the relationships among subsections of a test. Items on the KAMM Assessments are based on indicators (learning outcomes); there are 10 indicators measured on each mathematics assessment with exactly 4 items per indicator, and there are typically 7 to 11 indicators measured on each reading assessment and between 3 and 6 items per indicator. Table 5.1. provides the number of items per indicator by grade level. Results from such analyses provide empirical evidence for understanding the basic internal structure of the domain being measured. Evaluating the associations across these indicators, which are defined by groups of related items, is one method for providing evidence for construct validity. It is expected that these relationships would be low to moderate because while they are all intended to measure the same construct (i.e, mathematics or reading ability), they are simultaneously measuring different aspects of that construct. The size of the correlation coefficient between these scores will indicate the strength of the relationship between the criteria. Generally, the higher the correlation coefficient between the scores, the more valid the test is at assessing the criterion (Carmines & Zeller, 1979, p. 18).

Table 5.1  
*Number of Items for Each Reading Indicator*

Grade	1.3.1	1.3.2	1.3.3	1.3.4	1.3.5	1.4.2	1.4.7	1.4.8	1.4.9	1.4.10	1.4.11	1.4.14	1.4.15	2.1.1	2.1.2	2.1.3	TOTAL
3	0	4	0	0	0	5	0	4	4	0	5	0	0	4	4	0	30
4	4	0	0	0	0	4	4	6	0	7	4	0	0	4	4	0	37
5	4	0	0	0	0	4	4	5	0	5	4	0	5	4	6	4	45
6	4	0	0	0	0	4	5	5	0	6	5	0	5	5	4	4	47
7	3	0	0	4	0	6	5	5	0	7	0	0	4	4	4	4	46
8	4	0	0	5	0	5	4	5	0	5	0	4	4	4	4	4	48
11	4	0	0	4	0	6	5	4	0	5	0	4	5	4	4	4	49

## Results

Inter-indicator correlations were calculated for the KAMM form at each grade in both mathematics and reading and are presented in Tables 5.2 to 5.15. The Pearson Product-Moment Correlation was used. As provided in Table 5.16, the mathematics median inter-indicator correlations ranged from between 0.21 and 0.33. However, at grade 11, there were two correlation coefficients that were not significant which could offer some explanation as to why that grade level had the lowest median inter-indicator correlation. In reading, the median inter-indicator correlations ranged from 0.34 to 0.46. It may be that curriculum differences between how indicators are taught across mathematics and across reading could account for the differences observed in these correlation coefficients.

Table 5.2  
*Correlation Coefficients for 2007 Grade 3 Mathematics KAMM Indicators*

<b>N = 779</b>	<b>3.1.1.K2</b>	<b>3.1.1.K3</b>	<b>3.1.1.K4</b>	<b>3.1.4.A1</b>	<b>3.1.4.K7</b>	<b>3.2.1.A2</b>	<b>3.3.1.K4</b>	<b>3.3.2.A1</b>	<b>3.3.2.K2</b>
<b>3.1.1.K3</b>	0.41								
<b>3.1.1.K4</b>	0.41	0.46							
<b>3.1.4.A1</b>	0.31	0.38	0.33						
<b>3.1.4.K7</b>	0.38	0.32	0.29	0.33					
<b>3.2.1.A2</b>	0.35	0.35	0.37	0.31	0.41				
<b>3.3.1.K4</b>	0.21	0.25	0.25	0.28	0.32	0.28			
<b>3.3.2.A1</b>	0.32	0.35	0.31	0.34	0.32	0.32	0.29		
<b>3.3.2.K2</b>	0.30	0.29	0.28	0.27	0.35	0.31	0.25	0.32	
<b>3.4.1.K2</b>	0.25	0.30	0.23	0.35	0.33	0.30	0.35	0.41	0.27

*Note:* All correlations are significant at the 0.01 level (2-tailed).

- 3.1.1.K2:** Compares and orders whole numbers from 0 through 10,000 with and without the use of concrete objects.
- 3.1.1.K3:** Knows, explains, and uses equivalent representations including the use of mathematical models for: addition/subtraction of whole numbers from 0 through 1,000; multiplication facts of the 5s and the 10s; and addition/subtraction of money.
- 3.1.1.K4:** Determines the value of mixed coins and bills with a total value of \$50 or less.
- 3.1.4.A1:** One-step real-world addition/subtraction problems with whole numbers from 0 through 10,000 and monetary amounts using dollar and cents notation through \$500.00.
- 3.1.4.K7:** Identifies multiplication/division fact families through the 5s and the 10s.
- 3.2.1.A2:** Multiple representations of the same pattern.
- 3.3.1.K4:** Recognizes and describes the square, triangle, rhombus, hexagon, parallelogram, and trapezoid from a pattern block set.
- 3.3.2.A1:** Solves real-world problems by applying appropriate measurements: length to the nearest inch, foot, or yard; length to the nearest centimeter or meter; and number of days in a week.
- 3.3.2.K2:** Reads and tells time to the minute using analog and digital clocks.
- 3.4.1.K2:** Lists some of the possible outcomes of a simple event in an experiment or simulation including the use of concrete objects.



Table 5.3  
*Correlation Coefficients for 2007 Grade 4 Mathematics KAMM Indicators*

<b>N = 829</b>	<b>4.1.2.K5</b>	<b>4.1.4.A1</b>	<b>4.1.4.K6</b>	<b>4.2.2.K2</b>	<b>4.2.3.A1</b>	<b>4.3.1.A2</b>	<b>4.3.2.K2</b>	<b>4.3.3.K2</b>	<b>4.3.4.K3</b>
<b>4.1.4.A1</b>	0.28								
<b>4.1.4.K6</b>	0.33	0.41							
<b>4.2.2.K2</b>	0.30	0.43	0.39						
<b>4.2.3.A1</b>	0.27	0.32	0.31	0.35					
<b>4.3.1.A2</b>	0.21	0.21	0.24	0.25	0.27				
<b>4.3.2.K2</b>	0.24	0.26	0.29	0.30	0.29	0.28			
<b>4.3.3.K2</b>	0.21	0.21	0.25	0.29	0.23	0.27	0.34		
<b>4.3.4.K3</b>	0.18	0.22	0.22	0.31	0.25	0.27	0.29	0.33	
<b>4.4.2.K1</b>	0.23	0.30	0.27	0.33	0.30	0.30	0.33	0.37	0.32

Note: All correlations are significant at the 0.01 level (2-tailed).

- 4.1.2.K5:** Uses the concepts of properties (i.e., addition/subtraction, additive identity, multiplicative identity, symmetric property of equality) with the whole number system and demonstrates their meaning including the use of concrete objects.
- 4.1.4.A1:** Solves one-and two-step real-world problems with one or two operations using these computational procedures: a) adds/subtracts whole numbers from 0 through 10,000 and when used as monetary amounts; b) multiplies through a two-digit whole number by a two-digit whole number; c) multiplies monetary amounts less than \$100 by whole numbers less than ten; and e) figures correct change through \$20.00.
- 4.1.4.K6:** Shows the relationship between operations (i.e, addition/subtraction, addition/multiplication, multiplication/division, subtraction/division) with the basic fact families (addition facts with sums from 0 through 20 and corresponding subtraction facts, multiplication facts from 1 x 1 through 12 x 12 and corresponding division facts) including the use of mathematical models.
- 4.2.2.K2:** Solves one-step equations using whole numbers with one variable and a whole number solution that: a) finds the unknown in a multiplicaiton/division equation based on the multiplication facts from 1 x 1 through 12 x 12 and corresponding division facts; b) finds the unknown in a money equation using multiplication/division based upon the facts and addition/subtraction with values through \$10; and c) finds the unknown in a time equation involving whole minutes, hours, days, and weeks with values through 200.
- 4.2.3.A1:** Represents and describes mathematical relationships between whole numbers from 0 through 1,000 using concrete objects, pictures, written descriptions, symbols, equations, tables, and graphs.
- 4.3.1.A2:** Identifies the plane figures (circles, squares, rectangles, triangles, ellipses, rhombi, octagons, hexagons, pentagons, trapezoids) used to form a composite figure.
- 4.3.2.K2:** Selects; explains the selection of; and uses measurement tools, units of measure, and degree of accuracy appropriate for a given situation to measure: a) length, width, and height to the nearest fourth of an inch or to the nearest centimeter; b) volume to the nearest cup, pint, quart, or gallon; to the nearest liter; or to the nearest whole unit of a nonstandard unit; c) weight to the nearest -ounce or pound or to the nearest whole unit of a nonstandard unit of measure; d) temperature to the nearest degree; and e) time including elapsed time.
- 4.3.3.K2:** Recognizes, performs, and describes one transformation (reflection/flip, rotation/turn, translation/slide) on a two-dimensional figure or concrete object.
- 4.3.4.K3:** Identifies and plots points as whole number ordered pairs in the first quadrant of a coordinate plane (coordinate grid).
- 4.4.2.K1:** Organizes, displays, and reads numerical (quantitative) and non-numerical (qualitative) data in a clear, organized, and accurate manner including title, labels, categories, and whole number intervals using these data displays: graphs using concrete objects; pictographs with a symbol or picture representing one, two, five, ten, twenty-five, or one-hundred including partial symbols when the symbol represents an even amount; and frequency tables (tally marks).

Table 5.4  
*Correlation Coefficients for 2007 Grade 5 Mathematics KAMM Indicators*

<b>N = 684</b>	<b>5.1.1.K1</b>	<b>5.1.3.A4</b>	<b>5.1.4.A1</b>	<b>5.2.2.K2</b>	<b>5.2.3.K4</b>	<b>5.3.1.A1</b>	<b>5.3.2.A1</b>	<b>5.3.2.K4</b>	<b>5.4.2.A1</b>
<b>5.1.3.A4</b>	0.25								
<b>5.1.4.A1</b>	0.37	0.35							
<b>5.2.2.K2</b>	0.36	0.19	0.34						
<b>5.2.3.K4</b>	0.26	0.15	0.31	0.33					
<b>5.3.1.A1</b>	0.24	0.20	0.28	0.31	0.32				
<b>5.3.2.A1</b>	0.26	0.25	0.34	0.28	0.20	0.19			
<b>5.3.2.K4</b>	0.24	0.27	0.25	0.21	0.24	0.26	0.28		
<b>5.4.2.A1</b>	0.30	0.28	0.37	0.33	0.29	0.37	0.35	0.31	
<b>5.4.2.K3</b>	0.27	0.23	0.28	0.25	0.21	0.19	0.26	0.22	0.29

Note: All correlations are significant at the 0.01 level (2-tailed).

- 5.1.1.K1:** Knows, explains, and uses equivalent representations for: whole numbers from 0 through 1,000,000; fractions greater than or equal to zero (including mixed numbers); decimals greater than or equal to zero through hundredths place and when used as monetary amounts.
- 5.1.3.A4:** Determines if a real-world problem calls for an exact or approximate answer using whole numbers from 0 through 100,000 and performs the appropriate computation using various computational methods including mental math, paper and pencil, concrete materials, and appropriate technology.
- 5.1.4.A1:** Solves one-and two-step real-world problems using these computational procedures; multiplies and divides by 10, 100, and 1,000 and single digit multiples of each (10, 20, 30, ...; 100, 200, 300, ...; 1,000; 2,000; 3,000; ...).
- 5.2.2.K2:** Solves one-step linear equations with one variable and a whole number solution using addition and subtraction with whole numbers from 0 through 100 and multiplication with the basic facts.
- 5.2.3.K4:** Uses a function table (input/output machine, T-table) to identify, plot, and label whole number ordered pairs in the first quadrant of a coordinate plane.
- 5.3.1.A1:** Solves real-world problems by applying the properties of plane figures (circles, squares, rectangles, triangles, ellipses, rhombi, parallelograms, hexagons, pentagons) and the line(s) of symmetry.
- 5.3.2.A1:** Solves real-world problems by applying appropriate measurements and measurement formulas: length to the nearest eighth of an inch or to the nearest centimeter; weight to the nearest whole unit (pounds, grams, nonstandard units); months in a year and minutes in an hour; perimeter of squares, rectangles, and triangles; and area of squares and rectangles.
- 5.3.2.K4:** Converts within the customary system: inches and feet, feet and yards, inches and yards, cups and pints, pints and quarts, quarts and gallons, and pounds and ounces.
- 5.4.2.A1:** Interprets and uses data to make reasonable inferences, predictions, and decisions, and to develop convincing arguments from these data displays: graphs using concrete materials, pictographs, frequency tables, bar and line graphs, Venn diagrams and other pictorial displays, line plots, charts and tables, and circle graphs.
- 5.4.2.K3:** Identifies, explains, and calculates or finds these statistical measures of a whole number data set of up to twenty whole number data points from 0 through 1,000: minimum and maximum values, range, mode (no-, uni-, bi-), median (including answers expressed as a decimal or a fraction without reducing to simplest form), and mean (including answers expressed as a decimal or a fraction without reducing to simplest form).

Table 5.5  
*Correlation Coefficients for 2007 Grade 6 Mathematics KAMM Indicators*

<b>N = 968</b>	<b>6.1.1.K4</b>	<b>6.1.3.A2</b>	<b>6.1.4.A1</b>	<b>6.2.1.K4</b>	<b>6.2.2.A1</b>	<b>6.3.1.K7</b>	<b>6.3.2.A1</b>	<b>6.3.3.K1</b>	<b>6.3.4.K3</b>
<b>6.1.3.A2</b>	0.32								
<b>6.1.4.A1</b>	0.31	0.44							
<b>6.2.1.K4</b>	0.29	0.38	0.49						
<b>6.2.2.A1</b>	0.24	0.31	0.39	0.43					
<b>6.3.1.K7</b>	0.25	0.28	0.32	0.34	0.32				
<b>6.3.2.A1</b>	0.29	0.36	0.44	0.44	0.33	0.39			
<b>6.3.3.K1</b>	0.18	0.31	0.35	0.37	0.27	0.29	0.35		
<b>6.3.4.K3</b>	0.27	0.26	0.31	0.35	0.28	0.34	0.35	0.27	
<b>6.4.1.K4</b>	0.24	0.31	0.38	0.39	0.35	0.31	0.38	0.34	0.39

Note: All correlations are significant at the 0.01 level (2-tailed).

- 6.1.1.K4:** Knows and explains numerical relationships between percents, decimals, and fractions between 0 and 1.
- 6.1.3.A2:** Estimates to check whether or not the result of a real-world problem using rational numbers and/or the irrational number pi is reasonable and makes predictions based on the information.
- 6.1.4.A1:** Generates and/or solves one-and two-step real-world problems with rational numbers using these computational procedures: addition, subtraction, multiplication, and division of decimals through hundredths place.
- 6.2.1.K4:** States the rule to find the next number of a pattern with one operational change (addition, subtraction, multiplication, division) to move between consecutive terms.
- 6.2.2.A1:** Represents real-world problems using variables and symbols to write and/or solve one-step equations (addition, subtraction, multiplication, and division).
- 6.3.1.K7:** Classifies angles as right, obtuse, acute, or straight as well as triangles as right, obtuse, acute, scalene, isosceles, or equilateral.
- 6.3.2.A1:** Solves real-world problems by applying these measurement formulas: perimeter of polygons using the same unit of measurement; and area of squares, rectangles, and triangles using the same unit of measurement.
- 6.3.3.K1:** Identifies, describes, and performs one or two transformations (reflection, rotation, translation) on a two-dimensional figure.
- 6.3.4.K3:** Uses all four quadrants of the coordinate plane to: identify the ordered pairs of integer values on a given graph; plot the ordered pairs of integer values.
- 6.4.1.K4:** Represents the probability of a simple event in an experiment or simulation using fractions and decimals.

Table 5.6  
*Correlation Coefficients for 2007 Grade 7 Mathematics KAMM Indicators*

<b>N = 723</b>	<b>7.1.1.A1</b>	<b>7.1.4.K2</b>	<b>7.1.4.K5</b>	<b>7.2.1.K1</b>	<b>7.2.2.K7</b>	<b>7.2.2.K8</b>	<b>7.3.2.A1</b>	<b>7.3.2.K6</b>	<b>7.3.3.A3</b>
<b>7.1.4.K2</b>	0.22								
<b>7.1.4.K5</b>	0.21	0.21							
<b>7.2.1.K1</b>	0.33	0.21	0.21						
<b>7.2.2.K7</b>	0.27	0.20	0.23	0.30					
<b>7.2.2.K8</b>	0.24	0.19	0.15	0.22	0.27				
<b>7.3.2.A1</b>	0.24	0.12	0.11	0.22	0.23	0.16			
<b>7.3.2.K6</b>	0.24	0.15	0.17	0.24	0.29	0.35	0.16		
<b>7.3.3.A3</b>	0.32	0.25	0.19	0.26	0.26	0.26	0.21	0.41	
<b>7.4.2.K1</b>	0.24	0.22	0.21	0.26	0.29	0.29	0.17	0.27	0.35

Note: All correlations are significant at the 0.01 level (2-tailed).

- 7.1.1.A1:** Generates and/or solves real-world problems using equivalent representations of rational numbers and simple algebraic expressions.
- 7.1.4.K2:** Performs and explains these computational procedures: adds and subtracts decimals from ten millions place through hundred thousandths place; multiplies and divides a four-digit number by a two-digit number using numbers from thousands place through thousandths place; multiplies and divides using numbers from thousands place through thousandths place by 10, 100, 1,000, 0.1, 0.01, 0.001, or single-digit multiples of each; adds, subtracts, multiplies, and divides fractions and expresses answers in simplest form.
- 7.1.4.K5:** Finds percentages of rational numbers.
- 7.2.1.K1:** Identifies, states, and continues a pattern presented in various formats including numeric (list or table), algebraic (symbolic notation), visual (picture, table, or graph), verbal (oral description), kinesthetic (action), and written using these attributes: counting numbers including perfect squares, cubes, and factors and multiples (number theory); positive rational numbers including arithmetic and geometric sequences (arithmetic: sequence of numbers in which the difference of two consecutive numbers is the same, geometric: a sequence of numbers in which each succeeding term is obtained by multiplying the preceding term by the same number).
- 7.2.2.K7:** Knows the mathematical relationship between ratios, proportions, and percents and how to solve for a missing term in a proportion with positive rational number solutions and monomials.
- 7.2.2.K8:** Evaluates simple algebraic expressions using positive rational numbers.
- 7.3.2.A1:** Solves real-world problems by finding perimeter and area of two-dimensional composite figures of squares, rectangles, and triangles.
- 7.3.2.K6:** Uses given measurement formulas to find surface area of cubes and the volume of rectangular prisms.
- 7.3.3.A3:** Determines the actual dimensions and/or measurements of a two-dimensional figure represented in a scale drawing.
- 7.4.2.K1:** Organizes, displays, and reads quantitative (numerical) and qualitative (non-numerical) data in a clear, organized, and accurate manner including title, labels, categories, and rational number intervals using these data displays: frequency tables; bar, line, and circle graphs; Venn diagrams or other pictorial displays; charts and tables; stem-and-leaf plots (single); scatter plots; box-and-whiskers plots.

Table 5.7  
*Correlation Coefficients for 2007 Grade 8 Mathematics KAMM Indicators*

<b>N = 778</b>	<b>8.1.1.K5</b>	<b>8.1.2.A1</b>	<b>8.1.4.A1</b>	<b>8.1.4.K2</b>	<b>8.2.2.K3</b>	<b>8.2.3.A3</b>	<b>8.3.1.A1</b>	<b>8.3.4.K1</b>	<b>8.4.1.A4</b>
<b>8.1.2.A1</b>	0.24								
<b>8.1.4.A1</b>	0.26	0.22							
<b>8.1.4.K2</b>	0.31	0.24	0.28						
<b>8.2.2.K3</b>	0.32	0.19	0.34	0.27					
<b>8.2.3.A3</b>	0.23	0.08*	0.28	0.16	0.32				
<b>8.3.1.A1</b>	0.28	0.24	0.41	0.25	0.39	0.28			
<b>8.3.4.K1</b>	0.25	0.16	0.27	0.23	0.30	0.28	0.34		
<b>8.4.1.A4</b>	0.26	0.20	0.29	0.26	0.30	0.24	0.38	0.31	
<b>8.4.2.K3</b>	0.28	0.22	0.31	0.31	0.31	0.26	0.35	0.29	0.31

Note: Unless indicated otherwise, all correlations are significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

- 8.1.1.K5:** Knows and explains what happens to the product or quotient when: a positive number is multiplied or divided by a rational number greater than zero and less than one; a positive number is multiplied or divided by a rational number greater than one; and a nonzero real number is multiplied or divided by zero.
- 8.1.2.A1:** Generates and/or solves real-world problems with rational numbers using the concepts of these properties to explain reasoning: commutative, associative, distributive, and substitution properties; identity and inverse properties of addition and multiplication.
- 8.1.4.A1:** Generates and/or solves one- and two-step real-world problems using computational procedures and mathematical concepts with rational numbers, the irrational number pi as an approximation, and applications of percents.
- 8.1.4.K2:** Performs and explains these computational procedures with rational numbers: addition, subtraction, multiplication, and division of integers; and order of operations (evaluates within grouping symbols, evaluates powers to the second or third power, multiplies or divides in order from left to right, then adds or subtracts in order from left to right).
- 8.2.2.K3:** Solves one- and two-step linear equations with one variable with rational number coefficients and constants intuitively and/or analytically.
- 8.2.3.A3:** Translates between the numerical, tabular, graphical, and symbolic representations of linear relationships with integer coefficients and constants.
- 8.3.1.A1:** Solves real-world problems by using the properties of corresponding parts of similar and congruent figures.
- 8.3.4.K1:** Uses the coordinate plane to: list several ordered pairs on the graph of a line and find the slope of the line; recognize that ordered pairs that lie on the graph of an equation are solutions to that equation; recognize that points that do not lie on the graph of an equation are not solutions to that equation; determine the length of a side of a figure drawn on a coordinate plane with vertices having the same x- or y-coordinates.
- 8.4.1.A4:** Makes predictions based on the theoretical probability of a simple event in an experiment or simulation.
- 8.4.2.K3:** Determines and explains the measures of central tendency (mode, median, mean) for a rational number data set.

Table 5.8  
*Correlation Coefficients for 2007 Grade 10 Mathematics KAMM Indicators*

<b>N = 378</b>	<b>10.1.3.A1</b>	<b>10.2.2.A2</b>	<b>10.2.2.K3</b>	<b>10.2.3.A2</b>	<b>10.2.3.K6</b>	<b>10.3.1.A1</b>	<b>10.3.3.A1</b>	<b>10.4.2.A1</b>	<b>10.4.2.K4</b>
<b>10.2.2.A2</b>	0.30								
<b>10.2.2.K3</b>	0.20	0.21							
<b>10.2.3.A2</b>	0.16	0.34	0.18						
<b>10.2.3.K6</b>	0.03*	0.24	0.26	0.24					
<b>10.3.1.A1</b>	0.18	0.23	0.21	0.20	0.20				
<b>10.3.3.A1</b>	0.12*	0.13*	0.09*	0.15	0.15	0.17			
<b>10.4.2.A1</b>	0.22	0.24	0.17	0.23	0.16	0.26	0.19		
<b>10.4.2.K4</b>	0.15	0.32	0.20	0.21	0.20	0.22	0.17	0.32	
<b>10.4.2.K5</b>	0.25	0.36	0.23	0.24	0.18	0.30	0.20	0.37	0.34

Note: Unless indicated otherwise, all correlations are significant at the 0.01 level (2-tailed).

\* Correlation is significant at the 0.05 level (2-tailed).

- 10.1.3.A1:** Adjusts original rational number estimate of a real-world problem based on additional information (a frame of reference).
- 10.2.2.A2:** Represents and/or solves real-world problems with linear equations and inequalities both analytically and graphically.
- 10.2.2.K3:** Solves systems of linear equations with two unknowns using integer coefficients and constants.
- 10.2.3.A2:** Interprets the meaning of the x- and y- intercepts, slope, and/or points on and off the line on a graph in the context of a real-world situation.
- 10.2.3.K6:** Recognizes how changes in the constant and/or slope within a linear function changes the appearance of a graph.
- 10.3.1.A1:** Solves real-world problems by applying the Pythagorean Theorem.
- 10.3.3.A1:** Analyzes the impact of transformations on the perimeter and area of circles, rectangles, and triangles and volume of rectangular prisms and cylinders.
- 10.4.2.A1:** Uses data analysis (mean, median, mode, range, quartile, interquartile range) in real-world problems with rational number data sets to compare and contrast two sets of data, to make accurate inferences and predictions, to analyze decisions, and to develop convincing arguments from these data displays: frequency tables; bar, line, and circle graphs; Venn diagrams or other pictorial displays; charts and tables; stem-and-leaf plots (single); scatter plots; box-and-whiskers plots; and histograms.
- 10.4.2.K4:** Explains the effects of outliers on the measures of central tendency (mean, median, mode) and range and inter-quartile range of a real number data set.
- 10.4.2.K5:** Approximates a line of best fit given a scatter plot and makes predictions using the equation of that line.

Table 5.9  
*Correlation Coefficients for 2007 Grade 3 Reading KAMM Indicators*

<b>N = 1,050</b>	<b>3.1.3.2</b>	<b>3.1.4.2</b>	<b>3.1.4.8</b>	<b>3.1.4.9</b>	<b>3.1.4.11</b>	<b>3.2.1.1</b>
<b>3.1.4.2</b>	0.44					
<b>3.1.4.8</b>	0.50	0.46				
<b>3.1.4.9</b>	0.44	0.44	0.49			
<b>3.1.4.11</b>	0.46	0.49	0.47	0.54		
<b>3.2.1.1</b>	0.41	0.46	0.47	0.44	0.49	
<b>3.2.1.2</b>	0.39	0.42	0.42	0.47	0.48	0.50

*Note:* All correlations are significant at the 0.01 level (2-tailed).

- 3.1.3.2:** Determines the meaning of unknown words or phrases using context clues (e.g., definitions, restatements, examples, descriptions) from sentences or paragraphs.
- 3.1.4.2:** Understands the purpose of text features (e.g., title, graphs and charts, table of contents, pictures/illustrations, boldface type, italics, glossary, index) and uses such features to locate information in and to gain meaning from appropriate-level texts.
- 3.1.4.8:** Compares and contrasts information (e.g., topics, characters) in a text.
- 3.1.4.9:** Links causes and effects in appropriate-level narrative and expository texts.
- 3.1.4.11:** Identifies the topic, main idea(s), and supporting details in appropriate-level texts.
- 3.2.1.1:** Identifies and describes characters' physical traits, basic personality traits, and actions.
- 3.2.1.2:** Identifies and describes the setting (e.g., environment, time of day or year, historical period, situation, place) of the story or literary text.

Table 5.10  
*Correlation Coefficients for 2007 Grade 4 Reading KAMM Indicators*

<b>N = 1,041</b>	<b>4.1.3.1</b>	<b>4.1.4.2</b>	<b>4.1.4.7</b>	<b>4.1.4.8</b>	<b>4.1.4.10</b>	<b>4.1.4.11</b>	<b>4.2.1.1</b>
<b>4.1.4.2</b>	0.46						
<b>4.1.4.7</b>	0.32	0.36					
<b>4.1.4.8</b>	0.47	0.46	0.44				
<b>4.1.4.10</b>	0.48	0.48	0.44	0.54			
<b>4.1.4.11</b>	0.34	0.41	0.31	0.39	0.38		
<b>4.2.1.1</b>	0.40	0.40	0.39	0.50	0.43	0.32	
<b>4.2.1.2</b>	0.33	0.35	0.39	0.48	0.39	0.32	0.50

*Note:* All correlations are significant at the 0.01 level (2-tailed).

- 4.1.3.1:** Determines the meaning of words or phrases by using context clues (e.g., definitions, restatements, examples, descriptions) from sentences or paragraphs.
- 4.1.4.2:** Understands the purpose of text features (e.g., title, graphs/charts and maps, table of contents, pictures/illustrations, boldface type, italics, glossary, index, headings, subheadings, topic and summary sentences, captions) and uses such features to locate information in and to gain meaning from appropriate-level texts.
- 4.1.4.7:** Compares and contrasts information (e.g., topics, characters' traits, themes, problem-solution, cause-effect relationships) in one or more appropriate-level text(s) and identifies compare/contrast signal words.
- 4.1.4.8:** Links causes and effects in appropriate-level narrative and expository texts.
- 4.1.4.10:** Identifies the topic, main idea(s), and supporting details in appropriate-level texts.
- 4.1.4.11:** Identifies the author's purpose (e.g., to persuade, to entertain, to inform).
- 4.2.1.1:** Identifies and describes characters' physical traits, personality traits, and feelings, and explains reasons for characters' actions and the consequences of those actions.
- 4.2.1.2:** Identifies and describes the setting (e.g., environment, time of day or year, historical period, situation, place) of the story or literary text.



Table 5.11  
*Correlation Coefficients for 2007 Grade 5 Reading KAMM Indicators*

<b>N = 1,155</b>	<b>5.1.3.1</b>	<b>5.1.4.2</b>	<b>5.1.4.7</b>	<b>5.1.4.8</b>	<b>5.1.4.10</b>	<b>5.1.4.11</b>	<b>5.1.4.15</b>	<b>5.2.1.1</b>	<b>5.2.1.2</b>
<b>5.1.4.2</b>	0.43								
<b>5.1.4.7</b>	0.50	0.43							
<b>5.1.4.8</b>	0.40	0.44	0.42						
<b>5.1.4.10</b>	0.45	0.44	0.45	0.44					
<b>5.1.4.11</b>	0.43	0.43	0.44	0.40	0.43				
<b>5.1.4.15</b>	0.40	0.39	0.38	0.34	0.38	0.39			
<b>5.2.1.1</b>	0.43	0.40	0.45	0.46	0.43	0.44	0.33		
<b>5.2.1.2</b>	0.41	0.45	0.44	0.49	0.44	0.43	0.35	0.55	
<b>5.2.1.3</b>	0.42	0.35	0.45	0.38	0.41	0.38	0.33	0.45	0.46

*Note:* All correlations are significant at the 0.01 level (2-tailed).

- 5.1.3.1:** Determines the meaning of words or phrases by using context clues (e.g., definitions, restatements, examples, descriptions) from sentences or paragraphs.
- 5.1.4.2:** Understands the purpose of text features (e.g., title, graphs/charts and maps, table of contents, pictures/illustrations, boldface type, italics, glossary, index, headings, subheadings, topic and summary sentences, captions) and uses such features to locate information in and to gain meaning from appropriate-level texts.
- 5.1.4.7:** Compares and contrasts varying aspects (e.g., topics, characters' traits, themes, problem-solution, cause-effect relationships) in one or more appropriate-level texts.
- 5.1.4.8:** Links causes and effects in appropriate-level narrative, expository, and technical texts, and identifies signal words related to cause-effect relationships.
- 5.1.4.10:** Identifies the topic, main idea(s), supporting details, and theme(s) in appropriate-level texts.
- 5.1.4.11:** Identifies the author's purpose (e.g., to persuade, to entertain, to inform).
- 5.1.4.15:** Distinguishes between fact and opinion and recognizes propaganda (e.g., advertising, media) in various types of appropriate-level texts.
- 5.2.1.1:** Identifies and describes characters' physical traits, personality traits, and feelings and explains reasons for characters' actions and the consequences of those actions.
- 5.2.1.2:** Identifies and describes the setting (e.g., environment, time of day or year, historical period, situation, place) and explains the importance of the setting to the story or literary text.
- 5.2.1.3:** Identifies and describes the major conflict in a story and major events related to the conflict (e.g., problem or conflict, climax, resolution).

Table 5.12  
*Correlation Coefficients for 2007 Grade 6 Reading KAMM Indicators*

<b>N = 1,103</b>	<b>6.1.3.1</b>	<b>6.1.4.2</b>	<b>6.1.4.7</b>	<b>6.1.4.8</b>	<b>6.1.4.10</b>	<b>6.1.4.11</b>	<b>6.1.4.15</b>	<b>6.2.1.1</b>	<b>6.2.1.2</b>
<b>6.1.4.2</b>	0.28								
<b>6.1.4.7</b>	0.39	0.38							
<b>6.1.4.8</b>	0.36	0.43	0.48						
<b>6.1.4.10</b>	0.38	0.42	0.47	0.48					
<b>6.1.4.11</b>	0.29	0.32	0.38	0.36	0.42				
<b>6.1.4.15</b>	0.36	0.37	0.39	0.37	0.40	0.33			
<b>6.2.1.1</b>	0.31	0.35	0.46	0.43	0.47	0.32	0.27		
<b>6.2.1.2</b>	0.28	0.34	0.41	0.46	0.43	0.29	0.26	0.49	
<b>6.2.1.3</b>	0.25	0.33	0.38	0.39	0.38	0.30	0.29	0.39	0.35

Note: All correlations are significant at the 0.01 level (2-tailed).

- 6.1.3.1:** Determines the meaning of words or phrases using context clues (e.g., definitions, restatements, examples, descriptions, comparison-contrast, clue words) from sentences or paragraphs.
- 6.1.4.2:** Understands the purpose of text features (e.g., title, graphs/charts and maps, table of contents, pictures/illustrations, boldface type, italics, glossary, index, headings, subheadings, topic and summary sentences, captions, sidebars, underlining, numbered or bulleted lists) and uses such features to locate information in and to gain meaning from appropriate-level texts.
- 6.1.4.7:** Compares and contrasts varying aspects (e.g., characters' traits and motives, themes, problem-solution, cause-effect relationships, ideas and concepts, procedures, viewpoints) in one or more appropriate-level texts.
- 6.1.4.8:** Explains cause-effect relationships in appropriate-level narrative, expository, technical, and persuasive texts.
- 6.1.4.10:** Identifies the topic, main idea(s), supporting details, and theme(s) in text across the content areas and from a variety of sources in appropriate-level texts.
- 6.1.4.11:** Identifies and describes the author's purpose and basic techniques the author uses to achieve that purpose.
- 6.1.4.15:** Distinguishes between fact and opinion and recognizes propaganda (e.g., advertising, media), bias, and stereotypes in various types of appropriate-level texts.
- 6.2.1.1:** Describes different aspects of major and minor characters (e.g., their physical traits, personality traits, feelings, actions, motives) and explains how those aspects influence characters' interactions with other characters and elements of the plot, including resolution of the major conflict.
- 6.2.1.2:** Identifies and describes the setting (e.g., environment, time of day or year, historical period, situation, place) and explains the importance of the setting to the story or literary text.
- 6.2.1.3:** Identifies major and minor events related to the conflict in a story (e.g., problem or conflict, climax, resolution) and explains how one event gives rise to another.

Table 5.13  
*Correlation Coefficients for 2007 Grade 7 Reading KAMM Indicators*

<b>N = 1,146</b>	<b>7.1.3.1</b>	<b>7.1.3.4</b>	<b>7.1.4.2</b>	<b>7.1.4.7</b>	<b>7.1.4.8</b>	<b>7.1.4.10</b>	<b>7.1.4.15</b>	<b>7.2.1.1</b>	<b>7.2.1.2</b>
<b>7.1.3.4</b>	0.49								
<b>7.1.4.2</b>	0.39	0.37							
<b>7.1.4.7</b>	0.37	0.42	0.41						
<b>7.1.4.8</b>	0.36	0.37	0.36	0.39					
<b>7.1.4.10</b>	0.50	0.52	0.44	0.43	0.43				
<b>7.1.4.15</b>	0.22	0.24	0.27	0.26	0.27	0.28			
<b>7.2.1.1</b>	0.41	0.48	0.36	0.39	0.32	0.50	0.20		
<b>7.2.1.2</b>	0.38	0.43	0.38	0.43	0.32	0.47	0.21	0.44	
<b>7.2.1.3</b>	0.32	0.36	0.35	0.33	0.31	0.42	0.19	0.34	0.36

*Note:* All correlations are significant at the 0.01 level (2-tailed).

- 7.1.3.1:** Determines meaning of words or phrases using context clues (e.g., definitions, restatements, examples, descriptions, comparison-contrast, clue words) from sentences or paragraphs.
- 7.1.3.4:** Identifies and determines the meaning of figurative language, including similes, metaphors, analogies, hyperbole, onomatopoeia, personification, and idioms.
- 7.1.4.2:** Understands the purpose of text features (e.g., title, graphs/charts and maps, table of contents, pictures/illustrations, boldface type, italics, glossary, index, headings, subheadings, topic and summary sentences, captions, sidebars, underlining, numbered or bulleted lists) and uses such features to locate information in and to gain meaning from appropriate-level texts.
- 7.1.4.7:** Compares and contrasts varying aspects (e.g., characters' traits and motives, themes, problem-solution, cause-effect relationships, ideas and concepts, procedures, viewpoints, authors' purposes) in one or more appropriate-level texts.
- 7.1.4.8:** Explains cause-effect relationships in appropriate-level narrative, expository, technical, and persuasive texts.
- 7.1.4.10:** Identifies the topic, main idea(s), supporting details, and theme(s) in text across the content areas and from a variety of sources in appropriate-level texts.
- 7.1.4.15:** Distinguishes between fact and opinion and recognizes propaganda (e.g., advertising, media), bias, and stereotypes in various types of appropriate-level texts.
- 7.2.1.1:** Describes different aspects of major and minor characters (e.g., their physical traits, personality traits, feelings, actions, motives) and explains how those aspects influence characters' interactions with other characters and elements of the plot, including resolution of the major conflict.
- 7.2.1.2:** Identifies and describes the setting (e.g., environment, time of day or year, historical period, situation, place) and explains the importance of the setting to the story or literary text.
- 7.2.1.3:** Identifies major and minor events related to the conflict in a story (e.g., problem or conflict, climax, resolution) and explains how one event gives rise to another.

Table 5.14  
*Correlation Coefficients for 2007 Grade 8 Reading KAMM Indicators*

<b>N = 748</b>	<b>8.1.3.1</b>	<b>8.1.3.4</b>	<b>8.1.4.2</b>	<b>8.1.4.7</b>	<b>8.1.4.8</b>	<b>8.1.4.10</b>	<b>8.1.4.14</b>	<b>8.1.4.15</b>	<b>8.2.1.1</b>	<b>8.2.1.2</b>
<b>8.1.3.4</b>	0.27									
<b>8.1.4.2</b>	0.30	0.48								
<b>8.1.4.7</b>	0.26	0.44	0.46							
<b>8.1.4.8</b>	0.28	0.52	0.56	0.48						
<b>8.1.4.10</b>	0.26	0.42	0.40	0.37	0.48					
<b>8.1.4.14</b>	0.26	0.36	0.40	0.32	0.40	0.35				
<b>8.1.4.15</b>	0.25	0.34	0.34	0.33	0.34	0.29	0.28			
<b>8.2.1.1</b>	0.19	0.43	0.40	0.34	0.50	0.39	0.36	0.23		
<b>8.2.1.2</b>	0.26	0.49	0.52	0.37	0.47	0.43	0.38	0.26	0.52	
<b>8.2.1.3</b>	0.24	0.34	0.26	0.32	0.35	0.27	0.31	0.20	0.25	0.32

Note: All correlations are significant at the 0.01 level (2-tailed).

- 8.1.3.1:** Determines meaning of words or phrases using context clues (e.g., definitions, restatements, examples, descriptions, comparison-contrast, clue words, cause- effect) from sentences or paragraphs.
- 8.1.3.4:** Identifies and determines the meaning of figurative language, including similes, metaphors, analogies, hyperbole, onomatopoeia, personification, idioms, imagery, and symbolism.
- 8.1.4.2:** Understands the purpose of text features (e.g., title, graphs/charts and maps, table of contents, pictures/illustrations, boldface type, italics, glossary, index, headings, subheadings, topic and summary sentences, captions, sidebars, underlining, numbered or bulleted lists) and uses such features to locate information in and to gain meaning from appropriate-level texts.
- 8.1.4.7:** Compares and contrasts varying aspects (e.g., characters' traits and motives, themes, problem- solution, cause-effect relationships, ideas and concepts, procedures, viewpoints, authors' purposes, persuasive techniques) in one or more appropriate-level texts.
- 8.1.4.8:** Explains cause-effect relationships in appropriate-level narrative, expository, technical, and persuasive texts.
- 8.1.4.10:** Identifies the topic, main idea(s), supporting details, and theme(s) in text across the content areas and from a variety of sources in appropriate-level texts.
- 8.1.4.14:** Identifies the author's position in a persuasive text and describes techniques the author uses to support that position (e.g., bandwagon approach, glittering generalities, testimonials, citing statistics, other techniques that appeal to reason or emotion).
- 8.1.4.15:** Distinguishes between fact and opinion and recognizes propaganda (e.g., advertising, media), bias, and stereotypes in various types of appropriate-level texts.
- 8.2.1.1:** Describes different aspects of major and minor characters (e.g., their physical traits, personality traits, feelings, actions, motives) and explains how those aspects influence characters' interactions with other characters and elements of the plot, including resolution of the major conflict.
- 8.2.1.2:** Identifies and describes the setting (e.g., environment, time of day or year, historical period, situation, place) and explains the importance of the setting to the story or literary text.
- 8.2.1.3:** Identifies major and minor events related to the conflict in a story (e.g., problem or conflict, climax, resolution) and explains how one event gives rise to another.

Table 5.16

**Correlation Coefficients for 2007 Grade 11 Reading KAMM Indicators**

<b>N = 744</b>	<b>11.1.3.1</b>	<b>11.1.3.4</b>	<b>11.1.4.2</b>	<b>11.1.4.7</b>	<b>11.1.4.8</b>	<b>11.1.4.10</b>	<b>11.1.4.14</b>	<b>11.1.4.15</b>	<b>11.2.1.1</b>	<b>11.2.1.2</b>
<b>11.1.3.4</b>	0.41									
<b>11.1.4.2</b>	0.50	0.40								
<b>11.1.4.7</b>	0.54	0.44	0.57							
<b>11.1.4.8</b>	0.46	0.40	0.57	0.61						
<b>11.1.4.10</b>	0.52	0.41	0.62	0.60	0.59					
<b>11.1.4.14</b>	0.40	0.32	0.47	0.47	0.46	0.48				
<b>11.1.4.15</b>	0.48	0.40	0.56	0.57	0.53	0.56	0.46			
<b>11.2.1.1</b>	0.33	0.32	0.38	0.40	0.38	0.39	0.27	0.33		
<b>11.2.1.2</b>	0.46	0.42	0.48	0.55	0.51	0.55	0.38	0.48	0.39	
<b>11.2.1.3</b>	0.35	0.34	0.40	0.45	0.44	0.43	0.35	0.39	0.39	0.44

Note: All correlations are significant at the 0.01 level (2-tailed).

- 11.1.3.1:** Determines meaning of words or phrases using context clues (e.g., definitions, restatements, examples, descriptions, comparison-contrast, clue words, cause-effect) from sentences or paragraphs.
- 11.1.3.4:** Identifies, interprets, and analyzes the use of figurative language, including similes, metaphors, analogies, hyperbole, onomatopoeia, personification, idioms, imagery
- 11.1.4.2:** Understands the purpose of text features (e.g., title, graphs/charts and maps, table of contents, pictures/illustrations, boldface type, italics, glossary, index, headings, subheadings, topic and summary sentences, captions, sidebars, underlining, numbered or bulleted lists, footnotes, annotations) and uses such features to locate information in and to gain meaning from appropriate-level texts.
- 11.1.4.7:** Compares and contrasts varying aspects (e.g., characters' traits and motives, themes, problem-solution, cause-effect relationships, ideas and concepts, procedures, viewpoints, authors' purposes, persuasive techniques, use of literary devices, thoroughness of supporting evidence) in one or more appropriate-level texts.
- 11.1.4.8:** Explains and analyzes cause-effect relationships in appropriate-level narrative, expository, technical, and persuasive texts.
- 11.1.4.10:** Identifies the topic, main idea(s), supporting details, and theme(s) in text across the content areas and from a variety of sources in appropriate-level texts.
- 11.1.4.14:** Identifies the author's position in a persuasive text and describes techniques the author uses to support that position (e.g., bandwagon approach, glittering generalities, testimonials, citing statistics, other techniques that appeal to reason or emotion).
- 11.1.4.15:** Distinguishes between fact and opinion and recognizes propaganda (e.g., advertising, media), bias, and stereotypes in various types of appropriate-level texts.
- 11.2.1.1:** Describes different aspects of major and minor characters (e.g., their physical traits, personality traits, feelings, actions, motives) and explains how those aspects influence characters' interactions with other characters and elements of the plot, including resolution of the major conflict.
- 11.2.1.2:** Identifies and describes the setting (e.g., environment, time of day or year, historical period, situation, place) and explains the importance of the setting to the story or literary text.
- 11.2.1.3:** Identifies major and minor events related to the conflict in a story (e.g., problem or conflict, climax, resolution) and explains how one event gives rise to another.

Table 5.17  
*Summary of Median Inter-Indicator Correlations per Grade*

<b>Content Area</b>	<b>Grade 3</b>	<b>Grade 4</b>	<b>Grade 5</b>	<b>Grade 6</b>	<b>Grade 7</b>	<b>Grade 8</b>	<b>High School</b>
<b>Mathematics</b>	0.32	0.29	0.27	0.33	0.23	0.28	0.21
<b>Reading</b>	0.46	0.40	0.43	0.38	0.37	0.34	0.44

## CRITERION-RELATED EVIDENCE FOR THE VALIDITY OF KANSAS KAMM ASSESSMENT SCORES

While the primary evidence for the validity of KAMM lies in the processes used to develop and design the system, it is also informative to collect evidence related to the degree to which a test correlates with one or more outcome criteria, or what is called criterion-related validity evidence. This type of validity evidence is needed to support inferences about an individual's current or future performance by demonstrating that test scores are systematically related to other indicators or criteria. The key is the degree of relationship between the assessment items or tasks and the outcome criteria. To help ensure a good relationship between the assessment and the criterion, the criterion should be relevant to the assessment, and it should also be reliable.

Criterion validity (also known as predictive validity) refers to "how adequately a test score can be used to infer an individual's most probable standing on some measure of interest – the measure of interest being the criterion" (Cohen & Swerdlik, 2002, p. 160). Test scores are used to predict a student's future performance on a criterion; criterion validity is used to assess how accurate the test predicts the criterion. A criterion is "defined as the standard against which a test or a test score is evaluated" (Cohen & Swerdlik, 2002, p. 160). An evaluation of KAMM scores' criterion validity includes assessing the relationships of groups of scores in mathematics and reading for grades 3 through 11. The size of the correlation coefficient between these group scores will indicate the strengths of the relationships among the measures.

An analysis documenting the relationship between math and reading KAMM assessment scores was performed to provide a source of criterion-related validity evidence. This evaluation was based on a sample of students who were administered the KAMM Mathematics and Reading Assessments in the spring of 2007. Pearson Product-Moment Correlations were calculated using the total test scores for mathematics and reading for grades 3-8 and high school.

In order to estimate the strength of relationship of the underlying construct, correlations were corrected for attenuation using the following formula:

$$r_{x_t y_t} = \frac{r_{xy}}{\sqrt{r_{xx} r_{yy}}}$$

where  $r_{x_t y_t}$  is the estimated correlation between the true scores of the measures x and y,  $r_{xy}$  is the observed correlation, and  $r_{xx}$  and  $r_{yy}$  are the reliabilities of x and y, respectively.

Results from the validity assessments are displayed in Table 5.16 which details the intercorrelations for grades 3-8 and high school, including the observed correlations and the correlations after being corrected for attenuation.

Intercorrelations for all grades and subjects were positive and greater than 0.44 (observed) and 0.53 (corrected). Grades 3 through 6 appear to have higher intercorrelations between

mathematics and reading than grades 7 through high school. All correlations indicate students who performed well on the mathematics section of the state assessment also performed well in the reading section.

Table 5.16  
*Inter-Correlation Coefficients (r) Between Mathematics & Reading 2007 KAMM Assessments*

<b>Grade</b>	<b>r (N)</b>	<b><math>\alpha</math> for Mathematics</b>	<b><math>\alpha</math> for Reading</b>	<b>Correlation After Correction for Attenuation</b>
<b>3</b>	0.61 (673)	0.87	0.86	0.70
<b>4</b>	0.62 (683)	0.83	0.86	0.73
<b>5</b>	0.61 (535)	0.83	0.89	0.71
<b>6</b>	0.62 (773)	0.86	0.87	0.72
<b>7</b>	0.46 (567)	0.80	0.86	0.55
<b>8</b>	0.55 (572)	0.82	0.87	0.65
<b>High School</b>	0.44 (291)	0.77	0.91	0.53



## REFERENCES

- AERA, APA, & NCME (1999). *Standards for educational and psychological testing*. Washington, D.C.
- Berk, R. A. (Ed.). (1982). *Handbook of methods for detecting test bias*. Baltimore, MD: The Johns Hopkins University Press.
- Bock, R. D., Gibbons, R., & Muraki, E. J. (1988). Full information item factor analysis. *Applied Psychological Measurement, 12*, 261 – 280.
- Hanson, B. A. (1991). *Method of moments estimates for the four-parameter beta compound binomial model and the calculation of classification consistency indexes*. ACT Research Report 91-5. Iowa City, IA: American College Testing.
- Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Applications to psychological measurement*. Homewood, IL: Dow Jones-Irwin.
- Jöreskog, K. G. & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software.
- Kingston, N. M., Kahl, S. R., Sweeney, K. P., & Bay, L. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards* (pp. 249-282). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Kolen, M. J. & Brennan, R. L. (1995). *Test equating: Methods and practices*. New York, NY: Springer-Verlag.
- Livingston, S. A. & Zieky, M. J. (1982). *Passing scores: A manual for setting standards of performance on educational and occupational tests*. Princeton, NJ: Educational Testing Service.
- Lord, F. M. (1965). A strong true-score theory, with applications. *Psychometrika, 30*, 239 – 270.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- McDonald, R. P. (1967). Nonlinear factor analysis, *Psychometric Monographs, 15*. The Psychometric Society.
- McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34*, 100 – 117.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*, 5-11.

- Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The Bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Setting performance standards* (pp. 249-282). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Muthén, L. K. & Muthén, B. O. (1998-2001). *Mplus User's Guide*. Los Angeles, CA: Muthén & Muthén.
- Poggio, J., Glasnapp, D. R., Yang, X., & Poggio, A. J. (2005). A comparative evaluation of score results from computerized and paper & pencil mathematics testing in a large scale state assessment program. *The Journal of Technology, Learning, and Assessment*, 3 (6). Available from <http://www.jtla.org>.
- Poggio, J., Glasnapp, D., Yang, X., Beauchamp, A., & Dunham, M. (2005). *Moving from paper and pencil to online testing: Findings from a state large scale assessment program*. A series of papers presented at the NCME annual meeting, Montreal, April.
- Steiger, J. H. (1990). Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Research*, 25, 173 -180.
- Stout, W. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika*, 52, 589 – 617.
- Tate, R. (2003). A comparison of selected empirical methods for assessing the structure of responses to test items. *Applied Psychological Measurement*, 27, 159 – 203.
- Tittle, C. K. (1982). Use of judgmental methods in item bias studies. In R.A. Berk (Ed.), *Handbook of methods for detecting item bias*. Baltimore, MD: The Johns Hopkins University Press.
- Wood, R. L., Wilson, D., Gibbons, R., Schilling, S., Muraki, E., & Bock, D. (2002). *TESTFACT 4.0: test scoring and item factor analysis*. [computer program] Chicago: Scientific Software Inc.