Kansas Alternate Assessment Alignment Study

Links for Academic Learning

Report to the Kansas State Department of Education

Julia Shaftel, Ph.D.

Brooke Nash, M.S. Ed.

May 1, 2011

Center for Educational Testing and Evaluation

University of Kansas

1122 West Campus Road

Lawrence, KS 66045

Table of Contents

Kansas Alternate Assessment and Kansas Extended Standards

The Kansas Alternate Assessment (KAA) was originally developed by the Kansas State Department of Education (KSDE) in response to the requirements of the Individuals with Disabilities Education Act of 1997 (IDEA)—and later the No Child Left Behind Act of 2001 and the Individuals with Disabilities Education Improvement Act of 2004—that all students with disabilities be included in statewide accountability assessment systems. The KAA system is individualized to each student's specific needs and curricular goals (Poggio, Yang, Irwin, Glasnapp, & Poggio, 2007). A student's score is based on independent ratings of evidence selected for folio files. These folio files target learning outcomes from the state's curriculum standards and are selected by the student's Individualized Education Plan (IEP) team as part of the student's instructional goals for the year (Poggio et al., 2007).

In accordance with IDEA, the Kansas Extended Standards were developed to be consistent with the general standards for the purpose of ensuring that the education of all students, including those with the most significant disabilities, is uniform with goals and standards for students without disabilities as established by the Kansas State Board of Education (KSBE). The extended standards serve as the basis of the KAA. The state of Kansas uses an extended approach to alignment in that there is a direct reference between the general curriculum standards and the KAA performance indicators. Extended standards and their associated benchmarks are worded identically to the general curricular standards for that subject area. Downward extensions of the general standards are explicit only at the finest grain, which in Kansas is the indicator. Each extended indicator is intended to be used for instruction and assessment of students with severe cognitive disabilities within its corresponding benchmark and standard. Therefore, the extended indicators are referenced to general achievement indicators under those same benchmarks and standards. The Kansas Extended Standards for Reading consist of 41 instructional indicators for reading, grouped into six benchmarks within two standards. The Kansas Extended Standards for Science comprise 43 instructional indicators in 17 benchmarks within seven standards. The Kansas Extended Standards for Math include 71 indicators in 13 benchmarks and four standards.

Within each benchmark, the performance indicators address the curricular needs of students with significant disabilities. The extended indicators are not grade specific. Every extended indicator is referenced to several general indicators at one or more grade levels, and every general indicator is linked to one or more extended indicators. With this two-way content-based referencing, IEP teams have a choice of extended curricular content indicators for each grade-level indicator, allowing for a range of difficulty levels and types of performance for students at each grade level.

Once it has been determined that a student is eligible to take the KAA, indicators are selected to be assessed based on appropriate instructional goals for the student and aligned with the student's IEP. In other words, the IEP team determines which indicators are appropriate for students taking the KAA (KSDE, 2010). Five indicators for each assessed content area must be selected at the student's grade level. Of these five indicators, at least one indicator must be selected from each of the extended standards areas. In Reading, there are two standards, Reading and Literature. At least one indicator from each of these areas must be selected for a student's assessment in Reading. The other three indicators may be selected from one standard or spread across two areas. In Mathematics, there are four standards: Number and Computation, Algebra, Geometry, and Data. At least one indicator from each of these four areas must be selected for a student's assessment in Mathematics. The remaining indicator may be selected from any of the four standards. Students who are assessed in Science at grade 4, grade 7, and in high school must have one indicator chosen from each of the first four standards: Science as Inquiry, Physical Science, Life Science, and Earth and Space Science. The remaining indicator may be selected from any of the seven standards.

Once indicators have been chosen and appropriate instruction on the indicators has been provided, three pieces of evidence are to be collected during the assessment window and submitted to a student's folio. Each piece of evidence must consist of five trials, either performance opportunities or paper-and-pencil tasks. This process will result in 15 separate pieces of evidence in each content area. Evidence can take the form of worksheets, DVDs, photographs, audiotapes, and data sheets that record observations of performance tasks.

After the evidence has been collected, each piece in the folio is to be independently rated by three local raters, with the student's current special education teacher as one of the raters (KSDE, 2010). It is recommended that the other two raters should be professionally licensed educators who do not work directly with the student, since this will help ensure a more objective review of the evidence. Raters are to be trained in the review, evaluation, and scoring of student data folios (Poggio et al., 2007).

For the 2009–2010 school year, a number of changes were made to improve the KAA process (KSDE, 2010). Administrators must approve KAA plans to confirm that: (a) no functional activities are being used for a student's assessment (by order of the U.S. Department of Education), (b) indicators are aligned to the activities chosen, (c) only appropriate assessable activities are used for the folio file, and (d) a student's opinion is not being used as an assessable activity. Furthermore, a deadline for selection of indicators to be assessed was set at December 31, 2009, ensuring that teachers would select indicators before the testing window opened and with time to instruct the student on that curricular content.

Links for Academic Learning

Links for Academic Learning (LAL) (Flowers, Wakeman, Browder, & Karvonen, 2007) is a comprehensive and flexible methodology for assessing alignment of a state's alternate assessment based on alternate achievement standards (AA-AAS) with its general curricular content standards and assessments. LAL data come from three sources: special educators, content experts, and current classroom teachers of students participating in a state's AA-AAS. The first two groups of reviewers evaluate the state's AA-AAS, including its extended standards, if the state has created them, and its alternate assessment items or tasks. For Kansas, the targets of those reviews were the Kansas Extended Standards in Reading, Mathematics, and Science, and the Kansas Alternate Assessment (KAA) itself, including procedures for the collection of evidence into portfolios, the process whereby portfolios are evaluated and student progress is rated, and the professional development and support materials available to Kansas teachers who use the KAA. The third source of LAL data involves surveying teachers about their training and instructional practices.

Before continuing to describe the results of the study, however, it is essential to acknowledge the structure and efficacy of the LAL protocol as it has been conceived and formulated. The LAL is a powerful and flexible methodology for evaluating a state's AA-AAS system, permitting modification to fit the diverse needs of individual states. Without the massive effort involved in the development of this process for assessing alignment of an AA-AAS, the task would have been formidable. As it was, the thoughtfully implemented questions, procedures, and evaluation criteria enabled the Kansas team to approach the task as a series of studies that were manageable within a fairly restrictive budget and to complete the project within one year.

Links for Academic Learning Criteria for Instruction and Assessment

LAL defines eight criteria for instruction and assessment that are intended to ascertain the link of a state's AA-AAS—including its extended standards (if they are used) and its performance tasks or test items—with grade-level curriculum. These criteria are defined as follows (Flowers et al., 2007):

Criterion 1: The content is academic and includes the major domains/strands of the content area as reflected in state and national standards (e.g., Reading, Math, Science).

Criterion 2: The content is referenced to the student's assigned grade level (based on chronological age).

Criterion 3: The focus of achievement maintains fidelity with the content of the original grade-level standards (content centrality) and when possible, the specified performance.

Criterion 4: The content differs from grade level in range, balance, and depth of knowledge, but matches high expectations set for students with significant cognitive disabilities.

Criterion 5: There is some differentiation in content across grade levels or grade bands.

Criterion 6: The expected achievement for students is for the students to show learning of grade-referenced academic content.

Criterion 7: The potential barriers to demonstrating what students know and can do are minimized in the assessment.

Criterion 8: The instructional program promotes learning in the general curriculum.

Each criterion of the LAL is evaluated by special educators, content area experts, or teachers of students participating in the AA-AAS. This procedure allows the AA-AAS to be reviewed from different perspectives and by individuals with diverse backgrounds and expertise. For the purpose of the Kansas alignment study, several modifications and additions were made to the LAL process as outlined in the 2007 manual, *Links for Academic Learning: An Alignment Protocol for Alternate Assessments Based on Alternate Achievement Standards* (Flowers et al.). We altered both the number and type of raters and the methodology for interpreting the reviews, but not the topics or content of the reviews.

The Kansas Alternate Assessment Alignment Study comprised three phases in order to address the eight criteria of the LAL and to incorporate a broader set of reviewers and perspectives. The first phase was a review of the KAA, the Kansas Extended Standards, and professional development materials by a national panel of special educators. Six university-based special educators external to Kansas were solicited to review both the KAA and the Kansas Extended Standards. A team of three experienced LAL reviewers rated the extended standards for all content areas; in addition, three independent reviewers rated the extended standards in their individual areas of expertise. The second phase consisted of teams of Kansas special education and content area teachers who collaboratively reviewed the Kansas Extended Standards and their intended linkage with general academic achievement standards. These teams were larger than required by the LAL, and disagreement and minority opinions among the reviewers were recorded for further analysis. The third phase was an online survey of Kansas special educators, each of whom was teaching a student participating in the KAA, to investigate professional development and instructional methods and tools.

Links for Academic Learning Student Population

The LAL was designed specifically to address the evaluation of an AA-AAS system, which is the state assessment geared toward students with significant cognitive disabilities. While each state has somewhat different criteria for its AA-AAS, all are limited to 1% of the student population that can be counted as proficient for the purposes of adequate yearly progress under the No Child Left Behind Act. Eligibility criteria for the KAA identify students with IEPs who are significantly delayed with respect to progress in the general grade-level curriculum, who require substantial adjustment to the curriculum, and who are unable to demonstrate their knowledge on general assessments even with accommodations. Participation in the KAA is not based on categorical labels or type of disability or on time spent in special education settings. Kansas provides a flowchart to assist in making decisions about which assessments are most appropriate for a particular student. Kansas developed a set of extended standards prior to 2000 in anticipation of participation in state assessments by students with significant disabilities as required by IDEA, and these have been available as instructional guides for teachers for over 10 years. Participation in the KAA is indicated when a student's instruction and IEP goals are based primarily on the extended standards, benchmarks, and indicators.

Purpose of the Kansas Alternate Assessment Alignment Study

The Kansas Alternate Assessment Alignment Study was undertaken at the suggestion of the Kansas Technical Advisory Committee. Members of the Technical Advisory Committee had raised questions concerning the alignment of the Kansas Alternate Assessment and the Kansas Extended Standards with the general curricular standards. After a review of the literature and consideration of options for an alignment study, LAL was suggested as an appropriate methodology and was approved by the Technical Advisory Committee in September 2009.

Special Education Reviewers

The first aspect of the LAL undertaking in Kansas was a review of the Kansas Extended Standards and the KAA by national expert special education reviewers. Three trained reviewers who had rated previous LAL projects for other states collaboratively rated these indicators. About 12% of the indicators were rated by more than one reviewer in order to establish consistency with the rating criteria and to reach sufficient levels of reviewer agreement. Three additional national special education experts performed solo ratings in their areas of expertise. The purpose of requesting additional independent reviews was to obtain the viewpoints of academic researchers familiar with different states' AA-AAS but not necessarily familiar with the LAL. Expertise of the special education reviewers is summarized in Table 1.

Table 1

Expertise of Special Education Reviewers

Re-viewer	Stan-dards Re-viewed	Current Role	Yrs. PK-12 Exp.	Yrs. Spec. Ed. Exp.	Degrees	Con-ducted Prof. Dev. for Tchrs	Taught Future Tchrs in Higher Ed.
1	Reading	Research Professor	30	30	M.S., Ed.D. Spec. Ed.	Yes	Yes
2	Math	Research Associate	9	18	M.S. Spec. Ed.; Ed.D. Leadership	Yes	Yes
3	Science	Retired Professor	5	20+	M.S. School Psych.; Ph.D. Spec. Ed.	Yes	Yes
4	All	Research Associate	10	10	M.S., Ph.D. Spec. Ed.	Yes	Yes

Reviewer	Standards Reviewed	Current Role	Yrs. PK-12 Exp.	Yrs. Spec. Ed. Exp.	Degrees	Conducted Prof. Dev. for Tchrs	Taught Future Tchrs in Higher Ed.
5	All	Research Assistant	16	16	M.Ed. Spec. Ed.	Yes	Yes
6	All	Research Assistant	10	10	M.Ed. Curr. & Inst.	Yes	Yes

Reviewer Materials

Both the team and independent special education reviewers were provided with a binder containing the following materials available from the KSDE website (www.ksde.org/):

- Description of the KAA
- KAA Eligibility Criteria and flowchart for participation decisions
- Information for Parents of Students with Disabilities
- Special Education Services information and contact names and phone numbers
- List of extended standards, benchmarks, and indicators in Reading, Mathematics, and Science
- Standards documents showing the intended linkages of the extended and general indicators
- Kansas Alternate Assessment Teacher’s Guide
- KAA Content Checklist for the Student Portfolio
- Professional development materials, including clarifying examples for the KAA

In addition, the binder included the Kansas Alternate Assessment Implementation Guide and examples of KAA portfolio cover sheets and evidence labels from the website of the University of Kansas Center for Educational Testing and Evaluation (CETE) that are made available to teachers and administrators conducting the KAA.

Special education reviewers were provided coding sheets and instructions for completing them. The coding sheets consisted of spreadsheets of the extended indicators with columns for coding the required constructs as well as documents from the appendices of the LAL manual. Coding worksheets were to be used to record each reviewer's assessment of the required communication level of each extended indicator, the highest referenced grade, and the age appropriateness of each indicator. Instructions included rubrics for coding the level of symbolic communication and age appropriateness as provided in the LAL manual. Reviewers were also asked to respond to specific questions about the barriers that might be encountered by students with various disabilities participating in the KAA, the degree of inference required to assess student performance on the KAA, features of teacher professional development materials, and other program quality indicators. Descriptions of the tasks and results of the study are presented in the order that they were completed by reviewers in the three phases of the study, not in the order of the LAL criteria, which often require different data sources for a single criterion.

Levels of Symbolic Communication

Special education reviewers were first asked to rate the level of symbolic communication required for each indicator as described in Symbolic Access Rating for Criterion 7 (Flowers et al., 2007, p. 38). While the LAL manual refers to ratings of AA-AAS items, ratings of the extended indicators themselves were completed for the KAA due to the individualized nature of actual KAA tasks developed by teachers. No bank of KAA tasks used by teachers exists, and one of the challenges for teachers in creating appropriate tasks is the potential for lack of alignment with the KAA. Therefore all reviews were at the level of the extended indicator. LAL definitions for levels of communication are shown in Table 2.

Table 2

Symbolic Communication Levels Used in Criterion 7 of LAL

Category	Description
Awareness/Presymbolic	Has no clear response and no objective in communication; communicates with gestures, purposeful moving to object, sounds;
Concrete Symbolic	Beginning to use pictures or other symbols to communicate within a limited vocabulary
Abstract Symbolic	Speaks or has vocabulary of signs, pictures to communicate. Recognizes some sight words, numbers, etc.

In contrast, the Kansas study distinguished four levels of communication for the special education reviewers, as shown in Table 3. The lowest two levels from the LAL were separated for the purposes of the special educator review.

Table 3

Symbolic Communication Levels Used in Criterion 7 of the KAA Alignment Study

Code	Level	Description
1	Awareness	Has no clear response and no objective in communication
2	Pre-symbolic	Communicates with gestures, purposeful moving to object, sounds
3	Concrete Symbolic	Beginning to use pictures or other symbols to communicate within a limited vocabulary
4	Abstract Symbolic	Speaks or has vocabulary of signs, pictures to communicate. Recognizes some sight words, numbers, etc.

The above definitions are inconsistent with the categories of communication for coding indicators as “not academic,” which is required in Criterion 1 (Flowers et al., 2007, p. 39) and shown in Table 4. Those categories combine all symbolic communication into a single category but include “purposeful moving toward objects,” which is intentional and might include some natural gestures, within the nonsymbolic category. The use of different definitions for different LAL tasks may have created confusion for reviewers. This is further suggested by the use of the term “Early Symbolic” instead of “Concrete Symbolic” in some places in the LAL manual (Flowers et al., 2007, pp. 39, 81).

Table 4

Symbolic and Nonsymbolic Communication Codes Used in Criterion 1 of LAL

Codes	Definitions
S	Symbolic: Item/task is answered through symbolic communication (pictures, symbols, signs, speech)
N	Nonsymbolic: Item/task is answered through nonsymbolic communication (gesture, purposeful moving toward object, sounds)

Another set of communication skill descriptions occurs in the third part of the LAL, the Curriculum Indicators Survey (CIS), where a three-part categorization of symbolic communication is used (Table 5). In these descriptions, the lowest level may include both intentional and nonintentional forms of nonsymbolic communication, including gestures. The higher two levels both involve clear use of symbols and may therefore be difficult to distinguish consistently. The differences between the numbers of communication categories and their definitions make comparisons across the components of the LAL difficult.

Table 5

Descriptions of Levels of Symbolic Communication From the Curriculum Indicators Survey

Level	Description
1	Has not yet acquired the skills to discriminate between pictures or other symbols (and does not use symbols to communicate). May or may not use objects to communicate. May or may not use idiosyncratic gestures, sounds/vocalizations, and movements/touch to communicate with others. A direct and immediate relationship between a routine activity and the student’s response may or may not be apparent. The student may have the capacity to sort very different objects, maybe by trial and error. Mouthing and manipulation of objects leads to knowledge of how objects are used. May combine objects (e.g., place one block on another).

Level	Description
2	May use some symbols to communicate (e.g., pictures, logos, objects). Beginning to acquire symbols as part of a communication system. May have limited emerging functional academic skills. Representations probably need to be related to the student's immediate environment and needs.
3	Communicates with symbols (e.g., pictures) or words (e.g., spoken words, assistive technology, ASL, home signs). May have emerging or basic functional academic skills. Emerging writing or graphic representation for the purpose of conveying meaning through writing, drawing, or computer keying.

Levels of symbolic communication have been recently addressed in relevant literature. Browder, Ahlgrim-Delzell, Courtade-Little, and Snell (2005) described students' levels of symbolic communication and comprehension abilities with respect to academic instruction. The authors introduced three stages of symbolic communication: presymbolic, early symbolic, and extended symbolic. In the chapter, symbol use as described is primarily expressive and is taught to students for the purposes of making choices, responding to questions, and expressing preferences. Student examples assume the existence of symbolic receptive communication in forms such as listening to books on tape, following spoken directions, and understanding questions. The authors noted that it is difficult to assess receptive skills when an individual cannot demonstrate expressive symbolic communication due to individual challenges and disabilities. Therefore, explicitly teaching the use of objects and graphics as symbols to enhance active participation is necessary.

Four levels of symbolic communication—awareness, presymbolic, concrete symbolic, and abstract symbolic—were evaluated in a subsequent study (Browder, Flowers, & Wakeman, 2008) that surveyed teachers about their students' symbolic communication and their types of responses to 10 academic tasks. For that study, the awareness level was tested separately as a way to distinguish students who may have limited intentionality in communication. Only a small percentage of students (<5%) was identified at the awareness level while the majority (55%) was rated at the abstract symbolic level. However, there was no assessment of the academic task response measure other than a reliability analysis. Since the four choices of student's level of response for the academic tasks were specifically formulated to reflect the four levels of symbolic communication, teachers' ratings of student response to academic tasks had a predictably high correspondence with their ratings of students' symbolic communication categories. Significant differences among all four communication levels were found for the academic tasks using cluster analysis, and solutions for two, three, and four clusters of students were considered. Analyses of the results suggested combining the awareness category with that of presymbolic communication because a three-cluster solution showed the highest agreement with teacher ratings of communication level. Due to the lack of external validity assessment and the fact that cluster analysis found reasonable solutions for two to four clusters of students, those results should be considered preliminary. For that reason, all four levels were tested separately with the special education review portion of the current study.

In some places, the LAL reveals misinterpretation of the concept of a symbol or symbolic communication. For example, this occurs in an example in Appendix K (Flowers et al., 2007, p. 85), which states, "students do not need symbolic communication skills to rote count." Numerals and their names are symbols, so any use of numerals must be symbolic communication, regardless of the sophistication of numerical reasoning the student demonstrates. A similar misunderstanding or misstatement occurs in Appendix G (Flowers et al., 2007, p. 76), in which evaluators are asked to describe the flexibility of an AA-AAS with respect to a student who is "nonverbal; responds using printed words," and other similar descriptions. A student may not use oral communication, but a student who uses language in any modality must possess abstract symbolic communication ability and hence verbal skills.

These misinterpretations may be a reflection of confusion in the special education literature about the meaning of the term *symbol* and the extent to which a symbol may embody *iconic* elements. Definitions of *symbol* and *icon* reflect their historical denotation from the field of semiotics (i.e., the study of signs), of which both symbols and icons are types. An icon conveys its meaning by virtue of its similarity to the object being referenced while a symbol expresses its significance through customary or conventional usage (Atkin, 2010). For example, Merriam-Webster (2011) provides the following definitions:

> Symbol: something that stands for or suggests something else by reason of relationship, association, convention, or accidental resemblance; *especially* : a visible sign of something invisible <the lion is a *symbol* of courage>; an arbitrary or conventional sign used in writing or printing relating to a particular field to represent operations, quantities, elements, relations, or qualities

> Icon: a usually pictorial representation **:** image; a sign (as a word or graphic symbol) whose form suggests its meaning

Oxford (2011) defines *symbol* and *icon* as follows:

> Symbol: a mark or character used as a conventional representation of an object, function, or process, e.g., the letter or letters standing for a chemical element or a character in musical notation.
>
> Icon: a person or thing regarded as a representative symbol of something; a sign whose form directly reflects the thing it signifies

Rowland and Schweigert (1990) distinguished *abstract symbols,* such as words and manual signs, from *concrete symbols,* which include pictures and gestures such as miming drinking from a cup or throwing a ball. Concrete symbols are iconic, meaning they have a perceptual relationship to the objects to which they refer. *Tangible symbols* are a subset of the concrete symbols and refer to either two- or three-dimensional permanent objects with clear perceptual relationships to objects or ideas. The use of tangible symbols reduces cognitive demands on communication because they are permanent, unlike oral or signed language; they are manipulable and require only simple motor responses such as touching or pointing rather than the fine motor skills of writing or speaking; they facilitate recognition rather than recall; and they are readily discriminable both visually and tactilely. This paper (Rowland & Schweigert, 1990) is perhaps a source for the distinction between the LAL categories of concrete and abstract symbol use.

Additionally, the LAL descriptions fail to distinguish receptive and expressive language. One example is a reading objective listed in the example report in Appendix L (Flowers et al., 2007, Appendix L, p. 12), which is labeled as both foundational and presymbolic. It states, "Given a grade level book and the direction 'Show me the title of the book', student will identify the title." In this example, the student must have symbolic receptive oral language skills just to understand the request.

In contrast to the various LAL definitions, the Learner Characteristics Inventory (LCI) (Towles-Reeves, Kearns, Kleinert, & Kleinert, 2009) uses a three-stage model to describe expressive language and four categories for receptive language, as shown in Table 6.

Table 6

Expressive and Receptive Language Levels From the Learner Characteristics Inventory

Category	Level	Description
Expressive language	1	Uses symbolic language to communicate: Student uses verbal or written words, signs Braille, or language-based augmentative systems to request, initiate, and respond to questions, describe things or events, and express refusal
	2	Uses intentional communication, but not at a symbolic language level: Student uses understandable communication through such modes as gestures, pictures, objects, textures, pointing, etc., to clearly express a variety of intentions
	3	Student communicates primarily through cries, facial expressions, change in muscle tone, etc., but no clear use of objects/textures, regularized gestures, pictures, signs, etc., to communicate
Receptive Language	1	Independently follows one- or two-step directions presented through words (e.g. words may be spoken, signed, printed, or any combination) and does not need additional cues
	2	Requires additional cues (e.g., gestures, pictures, objects, or demonstrations/models) to follow one- or two-step directions
	3	Alerts to sensory input from another person (auditory, visual, touch, movement) but requires actual physical assistance to follow simple directions
	4	Uncertain response to sensory stimuli (e.g., sound/voice, sight/gesture, touch, movement, smell)

These descriptions contrast expressive and receptive symbolic communication. Awareness (see Table 3), which describes a receptive state, is subsumed under the category “Alerts to sensory input . . .” at a presymbolic level. Intentional presymbolic expressive communication is distinguished from apparently unintentional communication, with deliberate gestures placed at the intentional and presymbolic level. These definitions also place gestures, pictures, and objects at the presymbolic level and reserve the term "symbolic language" for expressive communication using the abstract symbols of words, written text, signs, and Braille. Another key difference between the LCI definitions and the LAL categories is that the LAL categories state that there is "no objective in communication" at the awareness level. While the individual may not have intentional communication, surely there is an objective in cries and facial expressions that indicate likes or dislikes such that the communication partner may be able to interpret that communication unambiguously. Alternatively, the last LCI category in expressive language states that the "student communicates primarily through cries, facial expressions, change in muscle tone, etc.," which presupposes that communication is taking place, even if not at an intentional or symbolic level.

There is a fairly rich literature base on the developmental trajectory of symbol use—both for typically developing children and for individuals with intellectual disabilities—which may help to clarify the distinctions between symbolic communication categories. Tomasello, Striano, and Rochat (1999) found that symbol use by young children proceeded through stages, from associating gestures with objects at 18 months of age (for example, combing hair with fingers elicited identification of a hairbrush), to using miniature objects to stand for their normal-sized counterparts at 26 months, to allowing concrete objects to stand for something else entirely at 35 months (for example, using a cup as a hat or a box as a shoe). In a second study (Tomasello et al., 1999), the youngest children were unable to use objects as symbols for other objects without adult modeling and scaffolding, whereas symbol use became more flexible and creative with age. Facility with symbols can therefore be described developmentally with gestures used and understood first, miniature objects (a type of tangible symbol in Rowland's & Schweigert's [1990] terms) next, and concrete objects used as abstract symbols last. This study supports the placement of gesture use at the presymbolic level and as the second category. Because individuals with cognitive disabilities tend to develop language in the same sequence as typically developing children (O'Toole & Chiat, 2006), albeit more slowly, this sequence may be useful in understanding growth in symbol use by individuals with significant cognitive disabilities.

The relationship between symbolic language and symbolic play in children with Down syndrome was investigated by O'Toole and Chiat (2006) using a procedure similar to that used by Tomasello et al. (1999). Gestures were easiest to comprehend, followed by miniature objects, and then symbolic play, with the 6- to 7-year-old children showing similar performance to the 35-month-old children in the earlier study and consistent with their age-equivalent language test scores. This study confirmed the developmental sequence of symbolic play for children with intellectual disabilities and highlighted the relationship of symbolic play with verbal, though not nonverbal, cognitive skills.

McLaughlin and Cascella (2008) investigated the teaching of distal gestures to individuals with severe cognitive disabilities on the basis of prior research indicating that distal gestures, such as pointing, serve a more advanced symbolic function than proximal gestures, such as physical contact with a communication partner. They found that students with presymbolic communication skills were able to use some distal gestures during dynamic assessment, whereas students whose communication was classified as nonsymbolic were not able to do so. These distinctions are not consistent with the LAL categorization, which puts gesture use at the presymbolic level. If the awareness level is intended to represent nonsymbolic or even nonintentional communication, then some proximal gestures, particularly reaching, touching, or pushing away, might better be included in this category, while distal gestures such as pointing belong at the intentional and presymbolic level. However, the LAL manual categorizes awareness and presymbolic communication together, so this distinction is unnecessary when those categories are merged.

Sutton, Trudeau, Morford, Rios, and Poirier (2009) discussed problems in the use of graphic symbols for children with disabilities in terms of the perceptual mismatch between receptive and expressive communication. Children who require graphic symbols, unlike children who learn either oral or visual (signed) language systems for both receptive and expressive use, are exposed to oral speech as a receptive modality but graphic symbols as an expressive modality. The authors described the concept of *iconicity,* or resemblance to the symbol’s intended meaning, as a guiding factor in the creation of symbol sets. Symbols can be defined as *transparent, translucent,* or *opaque,* depending on the extent to which the meaning of the symbol can be discerned from its visual form. The authors studied typically developing children in order to investigate the nature of transposing from one symbol system to another, which can be tested directly with children who use both systems but can only be inferred from work with children who use assistive technology. They found that three- and four-year-old children were not able to select three transparent symbols (photographs) to correspond to spoken sentences in subject-verb-object format, often omitting the verb symbol, and they did not generalize the use of graphic symbols from one type of task to another. Outcomes indicated that young children do not approach oral language tasks in the same way as visual tasks and that the modality change presented an obstacle. The children may have had to learn the graphic symbols as second-order representations of the receptive verbal symbols of spoken words, which would constitute a more complex symbolic representation task that was too difficult for them.

These findings suggest that graphic symbols or icons may actually require greater symbolic capacity for children with receptive oral or visual language skills, rather than offering a simpler symbol system as suggested by terms such as *concrete symbols* and by the relative placement of that category in the LAL hierarchy of definitions shown in Table 2. Concrete symbols such as photographs denoting verbs may be more difficult to identify or infer meaning from than words, at least for individuals with receptive oral or visual language (Sutton et al., 2009). For an individual without oral or manual expressive language capability, the extent of receptive symbolic language should be a topic of investigation in terms of the learner's repertoire of nonsymbolic response skills prior to making the assumption that communication will be facilitated with concrete symbols or icons.

This conclusion is substantiated by a study involving printed symbols as adjuncts to written language, which determined that simple drawings used as symbols for language concepts along with text did not significantly increase comprehension for adults with intellectual disabilities over text alone (Poncelas & Murphy, 2007). Even though many of these symbols were iconic in that they illustrated an object or action directly, participants did not automatically attach meaning to them, and differences among individuals in both the symbol and no-symbol groups were a function of better language and reading comprehension skills. Only users who had prior knowledge of and familiarity with these particular symbols showed significantly better comprehension when using them.

In summary, these studies show that some gestures, particularly proximal gestures such as reaching or pushing (McLaughlin & Cascella, 2008), probably belong at a nonsymbolic level because they may indicate nonintentional communication by students with the most limited skills. Gestures that closely mimic actions involving a common object (Tomasello et al., 1999; O'Toole & Chiat, 2006) lie within an emerging intentional or presymbolic level of communication, along with distal gestures such as pointing to an object out of reach. There is clearly an empirical differentiation between these two categories that is not captured by the LAL when the two lowest communication levels are collapsed into one category.

Concrete and tangible symbols, which are icons, may constitute an alternative symbolic visual/tactile communication system learned by individuals without oral or signed language, or their use may develop in parallel with early symbolic language (O'Toole & Chiat, 2006; Tomasello et al., 1999). For both typically developing preschoolers and for young children with Down syndrome, use of tangible symbols such as miniature objects precedes abstract symbols, which substantiates their placement within concrete symbolic communication. However, pictures, whether photographs or iconic graphics, may require higher order symbolic representation for students who have receptive but not expressive symbolic language (Sutton et al., 2009). The fact that symbols are iconic does not mean that they are more easily apprehended than abstract symbols (Tomasello et al., 1999), and they may not offer any additional meaning when used in conjunction with text unless specifically learned and practiced (Poncelas & Murphy, 2007). These findings do not imply that pictures, iconic graphics, and tangible or concrete objects should not be taught and used with individuals who have significant disabilities and need alternate modes of communication (Sutton et al., 2009). However, they should be recognized as symbol systems that require the cognitive sophistication of symbolic communication. This conceptualization differs from the definitions used by the LCI (Towles-Reeves et al., 2009). The difference between concrete and abstract symbols appears to be a matter of degree rather than a cognitive distinction.

In line with confusion in the LAL descriptions and consistent with the empirical evidence cited above, there was some concern by academic reviewers that the communication categories in the LAL were not distinct from one another and may inaccurately describe students who appear superficially to be functioning at those levels. One reviewer provided comments about vocabulary problems in the LAL definitions:

> Presymbolic, by definition would indicate that the student is inconsistent in both his/her understanding and use of expressive output . . . The definition of presymbolic . . . is really much more like "emerging symbolic" communication or "illocutionary" communication. That is, the individual is clearly understandable and purposeful . . . but not using words or symbols. The definition for "awareness" is actually the definition of "presymbolic" or "perlocutionary" . . . "No objective in communication" is impossible – because all kids are communicating, it is the adults who are not interpreting the communication that they have. In addition, "awareness" is a receptive term, not an expressive one. So we have extensive confusion in the vocabulary used here.

> "Concrete symbolic" was originated in the literature of "deaf-blind" in which the objects are used to represent words and this is appropriate for that set of students. We do not see this term used for other populations in communication disorders literature. I know it occurs in education/special education literature, but it is confusing across disciplines . . . The beginning use of pictures and "other symbols" places the individual in an emerging symbolic level. If he [or she] can use pictures and is beginning to use a couple of actual signs, these are not "concrete" objects, but symbols! Indeed if this is limited, then we cannot say that he [or she] is fully symbolic, but at least this is emerging use of symbols. "Concrete symbolic" is not an understandable term across disciplines.
>
> Finally the definition for "abstract symbolic" also includes the use of pictures – pictures [are also] used in the definition of "concrete symbolic" . . . If we say that the individual combines pictures to compose short sentences or complex ideas, then sure, that is symbolic. Much more delineation needs to [be] included for the "abstract symbolic" definition.

Levels of Symbolic Communication Results

Table 7 shows the number and proportion of indicators rated at the various symbolic communication levels by the different reviewers. In Reading, trained LAL reviewers working as a team selected level 2, presymbolic, for 24 of the 41 indicators, with another 15 indicators rated as level 3, concrete symbolic; two were rated at level 4, abstract symbolic; and none were rated at the awareness level. In contrast, the academic reviewer rated the majority of the indicators, 26 out of 41, at level 4, symbolic communication, with 11 more indicators rated at level 3, concrete symbolic communication. Overall, only 15% of the Reading indicators were rated at the same symbolic level by both types of reviewers. In terms of the level of communication required to show comprehension of the Reading indicators, there was a clear mismatch between the reviewers, which further substantiates the issue of imprecision in these definitions.

In Math, the LAL review team rated 50 of the 71 indicators at the presymbolic level, 14 at concrete symbolic, and 7 at abstract symbolic. The academic reviewer rated 70 indicators as requiring at least presymbolic communication and one as requiring concrete symbolic communication for an overall consistency with the review team of 69%.

In Science, 38 of 42 indicators were rated at the presymbolic level by the LAL review team, with two at the awareness level and two at concrete symbolic. Fourteen indicators were rated as presymbolic by the academic reviewer, with 18 additional indicators at the concrete symbolic level and eight at abstract symbolic. (Due to a printing error, one Science indicator was not rated by the special education reviewers.) The percent of matching ratings was 26%.

Table 7.

Ratings of Levels of Symbolic Communication for Reading, Math, and Science by the Review Team and Independent Reviewers

Level of Communication	Reading Reviewers		Mathematics Reviewers		Science Reviewers	
	Team	Ind.	Team	Ind.	Team	Ind.
Awareness						
Number	0	0	0	0	2	2
Percent	0	0	0	0	5	5
Presymbolic						
Number	24	4	50	70	38	14
Percent	59	10	70	99	90	33
Concrete Symbolic						
Number	15	11	14	1	2	18
Percent	37	27	20	1	5	43
Abstract Symbolic						
Number	2	26	7	0	0	8
Percent	5	63	10	0	0	19

The awareness level was utilized only two times, and even then different indicators were rated at that level by different reviewers, suggesting that almost all of the extended indicators require at least a presymbolic level of communication skill. This outcome validates the original decision of the LAL authors to combine the two lowest categories of this hierarchy.

Similarly, abstract symbolic communication was rated infrequently as a prerequisite skill for the Kansas extended indicators. Because the independent reviewers were different for each subject area, no comparison can be made across subject areas with respect to their ratings. It does appear that Reading required the highest level of symbolic communication skill overall as rated by both types of reviewers, followed by Math and then Science. This is not surprising given the nature of the Reading indicators, which deal for the most part with written symbols and the meaning of text. While numerals and mathematical signs are symbols, many Mathematics objectives deal with concepts like quantity, correspondence, and patterns that may not require sophisticated symbol use. The independent reviewers rated the Kansas extended indicators as requiring higher levels of symbolic understanding than did the review team, with the exception of the independent mathematics reviewer. It is unknown to what extent the lack of clarity in the descriptions of the levels of symbolic communication skills contributed to inconsistent ratings. This would be a useful topic of further investigation with the LAL.

Referenced Grade

Referenced grade, or intended grade alignment, was the second question to be assessed by the special education reviewers as described in Criterion 2 (Flowers et al., 2007, pp. 22–23). Referenced grade refers to the grade levels at which extended standards are defined by the state to have a connection with general standards. Special education reviewers were asked to review the general and extended indicators and simply identify the highest grade-level link for each extended indicator. The highest referenced grade level specified in the state standards was requested because there are often several links of the extended and general indicators at different grade levels. Since the reviewers were next asked to assess the grade appropriateness of the extended indicators, the highest grade-level linkage was selected as the most likely to be inappropriate and therefore the most crucial target for rating. Reviewers were not asked to give their opinion of the most appropriate grade alignment; that task was to be completed by the content experts working in teams with Kansas special educators.

Grade 10 has been specified in the Kansas Extended Standards as the highest referenced grade for all 40 indicators in Reading. Extended Reading Indicator 2.2.3—"demonstrates understanding of the cultural differences in language"—was added as a new indicator for 2010 and was not linked to any general indicator, perhaps simply due to an oversight. In Math, references for general and extended indicators occur at all grades from 3 to 10. The largest number of grade references is at grade 10, with 67% of extended indicators linked to 10th-grade general indicators. Grade 8 has 11% of the intended links, with smaller numbers throughout the lower grades. In Science, 76% of indicators are referenced at high school, with 24% at grade 7 and none at grade 4.

Age Appropriateness of Indicators

Special education experts assessed the age appropriateness of each indicator as requested by Criterion 5 of the LAL (Flowers et al., 2007, p. 35) at the highest referenced grade level. The codes and definitions for this rating are shown in Table 8.

Table 8

Codes for Rating Age Appropriateness of Extended Indicators

Code	Description
1	Adapted from grade-level content (e.g., *Roll of Thunder, Hear My Cry*)
2	Not grade specific; neutral; themes are appropriate for all ages (e.g., pets)
3	Inappropriate for teens (e.g., circus)
4	Inappropriate even for elementary age (e.g., Barney)

For Reading, the group of LAL reviewers rated all but three indicators as 2, neutral in content, and therefore appropriate for any age. They explained that those three indicators were rated as 3, inappropriate for teens, because the content did not link well to the general benchmark and standard. If that content were to be relevant for teens, it would have to be listed under a different standard. The academic reviewer, in contrast, rated all indicators as 1, adapted from grade-level content. Reviewer results are shown in Table 9.

Table 9

Ratings of Age Appropriateness of Extended Indicators

	Reading		Math		Science	
Rating	Team	Ind.	Team	Ind.	Team	Ind.
1	0	41	0	5	0	9
2	38	0	71	40	42	33
3	3	0	0	26	0	0
4	0	0	0	0	0	0

In Math, all indicators were rated at level 2, neutral and not grade specific, by the team of three LAL reviewers. The independent reviewer rated 7% at level 1, adapted from grade-level content; 56% at level 2, neutral; and 36% at level 3, inappropriate for teens; for an overall agreement of 56%. Examples of the indicators rated inappropriate for teens included "counts by rote," "matches like numerals," "adds one more to a set," "recognizes and/or identifies shapes," and "demonstrates understanding of calendar use."

In Science, the LAL review team again rated all indicators as 2. The independent reviewer rated 21% of the indicators at level 1, adapted from grade-level content, and the other 79% at level 2, neutral, for a 78% agreement rate.

The consistent neutral ratings by the LAL review team are not surprising given that the Kansas extended indicators were deliberately intended not to be grade or age specific. This causes problems when seeking evidence of growth in skills over the grade span, as the LAL does in the content expert review portion (next section). The only extended indicators that the independent reviewer rated as inappropriate for teens were 26 of the Math indicators, all of which referenced content found in the general indicators at the elementary level. Other than those, all indicators were rated as either adapted from grade-level content or neutral and appropriate for all ages.

Degree of Inference About Student Learning

Special education reviewers completed four appendices from the LAL manual (Flowers et al., 2007). Appendix F requested their evaluations of the degree of inference that can be made about student learning from the KAA. All six special education reviewers responded to questions about the level of accuracy, the level of independence, evidence of new learning, generalization across people and settings, generalization across materials and activities, standards for proficiency, and program quality indicators. Reviewers used all KAA materials to assess these aspects of student learning, including the KAA Teacher's Guide and the KAA Implementation Guide containing the KAA evidence collection and scoring instructions.

The rubric for these ratings was somewhat unclear in its definitions. For example, "High Student Inference" had the subheading "Can clearly infer student showed learning." More commonly, a high-inference situation is one in which a high level of inference or interpretation is required because little objective information is available upon which to base a conclusion. In contrast, a low-inference situation is one in which a large amount of objective and valid data is available, and therefore little or no inference is required to reach a conclusion. Nevertheless, because this part of the review was structured as a rubric, each reviewer could select the descriptive category with which she agreed. Ratings on these elements are shown in Table 10.

Table 10

Ratings of Degree of Inference About Student Learning

Criterion	High Student Inference: Can clearly infer student showed learning	Low Student Inference: Student performance mixed with educator performance	No Student Inference: Can clearly infer student did not have to show any learning/teacher or program performance rated
Level of Accuracy	5	0	1
Level of Independence	2	4	0
New Learning	1	4	1
Generalization across People and Settings	2	0	4
Generalization across Materials and Activities	3	0	3
Standard Setting	1	2	2
Program Quality Indicators	1	1	2

The first category of student performance, *level of accuracy*, was rated as high by most reviewers based on the detail required in the evidence labels and the 5-point KAA skill performance rubric. The second category, *level of independence*, was rated low by most reviewers. Rationales for that rating included a lack of clarity in the Implementation Guide about the use of prompting. *New learning* was also rated low by most reviewers because there is little grade-level differentiation in the Kansas Extended Standards, and guidelines for when KAA assessment data can be collected were vague. Since there is no baseline test or data collection, it may not be evident to KAA scorers what was learned by the student that year.

Two categories address generalization, the first with respect to people and settings, and the second dealing with conceptual generalization across materials and activities, which is considered to be a more important aspect of generalization according to the LAL. While some reviewers rated both categories as high, a majority reported having no available information about generalization of student learning. The rationale for these ratings referred to the lack of required information about people, settings, and materials on the evidence labels. While there is a place to record notes about collected evidence, information on generalization is not required, and therefore generalization cannot be inferred from collected performance data. One reviewer who rated the generalization elements as high noted that the Teacher's Guide requires each piece of evidence to be unique in some way and that the source of uniqueness must be recorded on the evidence label. This reviewer also commented that professional development materials on the KAA specifically instruct teachers to collect evidence of student performance across people, settings, and materials.

Standard setting in the LAL refers to the clarity of the proficiency standards for student performance on the AA-AAS, rather than the process for setting and describing those standards. Ratings were inconsistent because of confusion about the 5-point scoring rubric. On the one hand, the scoring rubric specifies the percent of trials or probes that must be completed for each level of proficiency, and some reviewers rated this element high because of the rubric. However, students can earn points for a low percentage of correct responses, and some reviewers felt that even a chance level of responding could earn too high a score with this rubric.

Finally, program quality indicators is an element of student learning that addresses the possible confounding of teacher ability and program quality in inferences drawn about student performance. One reviewer rated this category as high because no program quality indicators were described as part of the KAA process, implying that only student performance influenced evidence ratings and portfolio scores. Two other raters rated this category as allowing no student inference for the same reason, and two reviewers did not rate this category.

Minimizing Barriers for Students

Appendix G of the LAL deals with minimizing barriers for students with different characteristics, such as students who are blind or have low vision, student who are deaf or hard of hearing, and students who use various communication modalities and systems. In each case, reviewers were asked to identify whether there are no provisions for students with that characteristic, whether flexibility is built into the tasks, or whether accommodations or modifications are available. This particular element of the LAL posed a challenge for reviewers. In some cases, wording was such that the reviewers did not understand the response option. For example, one response option for each student characteristic was "No provision for students with these characteristics." One reviewer did not know whether that meant the assessment made no provision for individual student characteristics or whether no provision for these students was allowed on the assessment. Another issue concerned descriptions of student characteristics. Students who are blind or have low vision are described as "visual impairment/legally blind." Students who are deaf or hard of hearing are described as "hearing impaired." Both of these labels could be construed as pejorative by these students and their advocates. These descriptions need to be updated by the LAL authors to reflect current consensus about the appropriate terminology used to describe these students.

A considerable issue concerns communication, as was noted in the previous discussion of symbolic communication skills. For example, one student description read, "nonverbal; responds using printed words," with other options for picture, manual sign, and eye gaze responses. None of these response methods suggest that a student is nonverbal, and in fact any of these responses would suggest that a student does indeed have symbolic language abilities. The student may not use oral communication, but that does not imply that the student is nonverbal, only that an alternate communication system is required. A better description would simply have focused on the communication modality that a particular student may use, whether that is oral language, visual or sign language, pictures, or a communication board or other assistive technology. As mentioned earlier, a distinction between receptive and expressive language may be valuable here. There is a specific question for reviewers about the ability to measure performance for a student who does not yet have clear intentional communication even at the nonsymbolic level, which is a critical issue to address.

Of the six reviewers, only two provided responses to each question; therefore those responses and the comments of all the reviewers will be summarized here. Almost all reviewers noted that there was no explicit discussion of accommodations, modifications, or other supports in the KAA administration materials other than the recommendation that any accommodations or assistive technology used during assessment be the same as those used during instruction. Without better guidance, reviewers were unable to answer questions about barriers for students with various characteristics. The two reviewers who did respond to these questions chose the response that flexibility is built into the tasks and therefore they are accessible to all students. Two reviewers responded that the KAA had no way to capture responses from students who have no clear and intentional communication, while one reviewer thought that such a student could be assessed. Four reviewers responded that accommodations, modifications, and supports were not sufficiently defined to achieve standardized administration, while one reviewer thought that standardized administration was possible.

Alignment of Professional Development Resources

Appendix H of the LAL calls for reviewers to evaluate professional development materials and answer a series of 15 questions about teacher training. Reviewers generally agreed that teachers in Kansas are trained to use the extended standards, to use the student's grade level to determine what to teach, and to address the academic priorities set by the state for this population. Five out of six reviewers did not find evidence that teachers were trained to review grade-level content standards. The dissenting reviewer noted that the teacher must review the general standards in order to identify an appropriate linked extended standard as a target for assessment. Five out of six reviewers responded that teachers are not trained to align content for instruction with state standards, probably because intended alignment is provided by the state through the grade-referenced general indicators. Five out of six reviewers noted that there is no evidence that teachers are trained to promote active student learning by targeting independent responses. The other questions were answered in the negative by all reviewers. Topics included delivery of instruction at various depths of knowledge, guidance to plan for increasing expectations across grade levels or bands, decreasing prompting and increasing independent responding, generalization, promoting student mastery of skills, and minimizing barriers for students with sensory or physical impairments or different levels of symbolic communication skill. One reviewer noted that the training materials were focused on assessment administration and scoring, not on broader content addressing best practices.

Program Quality Indicators

The final element of the LAL reviewed by the national special education reviewers was Appendix I, the Program Quality Indicators Checklist. These ratings are provided in Table 11.

Table 11

Program Quality Indicators

Do the alternate assessment and professional development promote:	Yes	No
1. Opportunities for instruction in general education classrooms for students with significant cognitive disabilities?	0	5
2. Opportunities for instruction with typical peers for students with significant cognitive disabilities?	0	5
3. Opportunities for students with significant cognitive disabilities to make choices, problem solve, self-advocate, self-evaluate?	1	4
4. The provision of assistive technology for students who need it?	4	2
5. The access and use of typical classroom resources within instruction (e.g., science kits, grade-level books, textbooks)?	0	5
6. Literacy being promoted across the content areas for students with significant cognitive disabilities (e.g., the pairing of text with picture symbols and objects)?	0	5

Do the alternate assessment and professional development promote:	Yes	No
7. The meaningful linking of academic skills in functional contexts?	0	5

One reviewer did not rate most of these questions, commenting that there was no evidence for a response. The other reviewers generally responded that the KAA did not promote these quality indicators because there was no evidence in favor of them. The exceptions to these ratings were for question 4; reviewers noted that the KAA Implementation Guide specifically states that assistive technology should be used for the KAA as it is in teaching and that the evidence labels require the identification of accommodations, which could include assistive technology. One reviewer responded affirmatively to question 3 because clarifying examples of the extended indicators available to teachers provided samples of these skills.

Summary

The special education review portion of the Kansas Alternate Assessment Alignment Study increased the number of special educators from three to six, with three trained LAL reviewers working as a team to achieve consensus and interrater reliability according to the protocol described in the LAL. Three additional independent academic reviewers, mostly unfamiliar with the LAL, were solicited for their viewpoints on the KAA in their subject areas of expertise. This allowed us to obtain alternate viewpoints and compare and contrast reviewer ratings. All reviewers had master's degrees in educational fields, while three had doctorates in special education and two were special education doctoral students. All had at least 10 years' experience in special education, had conducted professional development for teachers, and had taught pre-service teachers at the postsecondary level.

The Kansas extended indicators formed the unit of analysis for this alignment study. In Kansas, teachers develop tasks to assess each extended indicator that has been selected as an appropriate target of assessment for each student. There is no bank of tasks or specific guidance on how to create a task, and for this reason there is substantial variation in the types and qualities of tasks used as evidence of learning. Furthermore, the extended indicators are not grade specific and are referenced to different general indicators at several grade levels. Conversely, several general indicators may be referenced to a single extended indicator, creating a network of linked indicators across grade and difficulty levels. One of the objectives of this study was to identify an improved alignment structure, which can only be accomplished at the indicator level.

The special education reviewers evaluated three characteristics of each extended indicator: level of symbolic communication required to demonstrate achievement, highest referenced grade, and age appropriateness. Following their review of the extended indicators, the special education reviewers evaluated KAA administration and professional development materials to provide feedback on the degree of inference that can be made about student learning, the presence of barriers for students with specific characteristics, evidence of teacher training in best practices, and evidence of program quality indicators.

Alternate viewpoints representing raters with different backgrounds and prior experience with the LAL were notable in several areas of the review. For the first task, which involved the level of communicative competence required to demonstrate achievement on the extended standards, independent reviewers for Reading and Science concluded that facility with abstract symbolic communication was required to a greater extent than the review team found. The third independent reviewer rated the extended indicators in Mathematics very similarly to the review team, that is, somewhat lower in terms of required symbolic communication. These differences may have been partially due to inconsistencies in the descriptions of different levels or types of symbolic communication, as well as misstatement or misinterpretation in the LAL manual about symbol use and what kind of communication is symbolic.

Reviewers were asked to assess the age appropriateness of the extended indicators at the highest referenced grade. Again there were meaningful differences between reviewers. The trained LAL review team rated virtually all extended indicators, except for three that they believed did not fit well under their reading standard, as neutral in content and suitable for any age, perhaps because the indicators had been worded intentionally so that they could be used at any grade. The independent reviewers, in contrast, rated many of the extended indicators as adapted from grade-level content, with a few marked as inappropriate for teens due to their elementary-level content.

Confusion appeared again as special education reviewers were asked to evaluate the degree of inference that can be made about student learning on the KAA. Except for the level of accuracy measured by the KAA, reviewers showed disagreement about these criteria, with the independent reviewers providing generally more positive ratings than the review team, who rated every other criterion as permitting little or no inference about student learning.

Most reviewers were unable to complete the Minimizing Barriers for Students Checklist, either because they could not find relevant evidence in the KAA materials or because they disagreed with the meanings of the definitions for student characteristics. Two of the independent raters completed the checklist and selected “flexibility built into tasks” as the way in which the KAA can be adapted because of its complete individualization for students with various characteristics and disabilities.

In contrast, the independent reviewers tended to rate the KAA lower on alignment of professional development resources. The review team rated two of the 15 criteria positively: teachers are trained to use the student’s grade or grade band to determine what to teach, and teachers are trained in the academic priorities set by the state for this population. The independent reviewers agreed with the review team that teachers are trained to use the extended standards, but otherwise their positive ratings were sporadic.

The last element evaluated was the extent to which the alternate assessment and professional development materials promote best practices. Only four reviewers, including two independent reviewers and two members of the review team, reported that assistive technology was available to students on the basis of evidence found in the evidence labels and the Implementation Guide. Otherwise, the only program quality indicator endorsed was that students with significant cognitive disabilities have opportunities to make choices and advocate for themselves, and this was endorsed by an independent reviewer on the basis of clarifying examples.

On the whole, the independent reviewers tended to rate elements of the KAA and the extended indicators more positively than did the review team working collaboratively. The reasons for these differences are not immediately apparent. All reviewers had significant special education experience. The LAL review team had the benefit of previous experience with the LAL in alignment studies conducted for several other states whose AA-AAS tests vary widely. The independent reviewers had the benefit of fresh eyes on both the KAA and the LAL materials. This led to some criticisms of the LAL, particularly in levels of communication, but it also perhaps led to a more thorough analysis of the materials presented to them, because they identified sources for some of their responses that were missed by the review team. Overall, the disagreements provide as much cause for reflection as do the agreements.

Content Expert Reviewers

A total of 24 educators participated in the Content Expert Review component of the Kansas Alternate Assessment Alignment Study, all of whom were working as educators in Kansas public schools. The content experts included English language arts (ELA), math, and science teachers as well as special educators working in classrooms with students eligible for the KAA. Each content team consisted of four content area experts and two special educators representing elementary and secondary grade levels. One team was formed for Science and one for Reading while two teams were assigned to review the much larger number of extended indicators in Mathematics. Both Mathematics teams represented all age and grade levels; Mathematics indicators were split into halves for the review process.

Of the 24 educators, 18 were female. Three of the educators indicated that ELA was their main content area, nine reported math (one of whom was also a special education teacher), four reported science, seven indicated special education, and one did not indicate a main subject area. Many teachers were qualified in more than one content area, and some special educators were also certified in content areas. When asked about their current professional role or job title, 18 of the participants reported that they were classroom teachers, three indicated that they were instructional coaches, two were library media specialists, and one was a district special education facilitator. Teachers also provided information about the number of years of teaching experience that they had in total, as well as specifically for the subject area that they currently teach. This information is displayed in Tables 12 and 13, respectively. While all 24 educators indicated that they had at least a bachelor's degree, 18 had a master's degree and five had another type of graduate degree. Other types of attained education that teachers listed included English as a Second Language (ESL) endorsements and Adaptive Special Education endorsements. Furthermore, 19 of the educators who participated in this project indicated that they had conducted professional development for teachers in their respective content areas, 18 had a leadership role in curriculum planning in their school or district, five were certified by the National Board for Professional Teaching Standards, and one had taught teachers in a higher education setting.

Table 12

Total Years of Teaching Experience

Years	%
5–10	29.2
11–16	12.6
17–22	25.1
23–28	12.5
29–34	8.4
35–40	12.6

Table 13

Years of Teaching in Content Area

Years	%
1–5	4.2
6–10	20.8
11–16	16.7
17–22	33.3
23–28	16.7
29–34	4.2

The Kansas Extended Standards were evaluated on several technical aspects related to their alignment with general indicators and national standards, as well as the depth of knowledge required by each indicator. Each group of content reviewers collaboratively rated the alignment of each extended indicator link with the various general indicators. The review process resulted in differing total numbers of ratings for each extended indicator depending on the number of times that extended indicator was referenced to general indicators as defined in the extended standards. Furthermore, the extended indicators could be interpreted differently based on their various referenced general indicators, resulting in some of the links being evaluated as having higher content or performance fidelity or better alignment than others. Even though it was a collaborative process, reviewers were not required to form a consensus and the recording of a minority viewpoint was allowed. To account for the variability in the number of links associated with each extended indicator, as well as the potential for disagreement among group members, ratings were evaluated as a percentage of the total number of ratings for individual extended indicators.

Reading. For Reading, six reviewers, including four reading/ELA educators and two special education experts, rated the relationship of the Kansas extended indicators to their referenced general indicators. Forty-one extended indicators in Reading are linked to between four and 10 general indicators each. For example, Extended Reading Indicator E.R.1.1.2, which states that the student assigns meaning to auditory symbols, is referenced to seven different general indicators across 3rd through 8th grades and high school. E.R.1.1.2 is linked to general indicator R.4.1.4.10 at 4th grade, which states that the student identifies the topic, main idea(s), and supporting details in appropriate level texts. However, it is also referenced to R.5.1.4.10 at 5th grade, which further adds that the student identifies theme(s) in appropriate level texts.

Mathematics. In Math, 12 reviewers consisting of nine math educators, one of whom also had special education certification, and three additional special education experts were divided into two groups of six. These were subsequently divided into four groups of three for some of the alignment activities due to the large number of Mathematics indicators and general indicator references. Thus, while most indicators were evaluated by groups of three, some were reviewed by groups of six. The individual general Mathematics indicators, unlike those in Science and Reading, are broken down into sub-indicators that define smaller portions of the overall content. This increases the number of potential references for the extended indicators. Therefore, each of the 71 Mathematics extended indicators is linked with two to 42 general indicators. For example, Extended Math Indicator E.M.1.4.3, which states that the student adds or subtracts to create a new set, is referenced to 21 sub-indicators for nine general indicators across 3rd through 8th grades and high school. Due to the large number of indicator links that needed to be reviewed in Math, each group of raters was assigned to review a portion of the indicators.

Science. In Science, six reviewers, including four science educators and two special education experts, rated 43 extended indicators linked with one to 12 general indicators each. For example, Extended Science Indicator E.S.3.4.1 states that the student adapts to environmental changes such as temperature, weather, light, etc., and this is linked to seven different general indicators across 7th grade and high school.

Academic Content

As noted by Flowers et al. (2007), it should not be assumed that the content of the extended standards (and thus the alternate assessment) is academic in nature. In accordance with Criterion 1 of the LAL (Flowers et al., 2007, p. 18), content needs to be evaluated prior to alignment to ensure that it is academic and not foundational. Extended standards that are rated as *not academic* should not be included in the alignment study. Thus, reviewers were asked to determine if the content associated with each extended indicator was considered academic. To make these decisions, reviewers were asked whether the item could be defined by a national standard for that content area. National standards for reading were defined by the National Council of Teachers of English (NCTE, Table 14). National standards for math were defined by the National Council of Teachers of Mathematics (NCTM,

Table 15). National standards for science were defined by the National Science Education Standards (NSES, Table 16). As recommended by Flowers et al. (2007), if the extended standard can be labeled according to a strand within the national standards, the reviewer codes it as academic and indicates the primary national standard link. A secondary link may be indicated if a secondary relationship is identified. Ratings for primary and secondary national standard links are displayed in Appendix A.

Table 14

Codes for National English Language Arts Standards

Code	National Standard	Description
E1	Reading	Decoding text including deciphering symbols (letter, pictures, Braille)
E2	Writing	Generating information-make a useful mark, composing to a scribe, or creating a printed product (symbols that represent text (e.g., picture symbols) to share with another person-like a book made of pictures made by student
E3	Speaking	Generating nonwritten communication
E4	Listening	More than response to sound; intentional response within context
E5	Viewing or Visually Representing	Purposeful focus on (or creation of) nontextual information and a response to what is seen
E6	Research	Obtaining new information

Note. Adapted from NCTE content standards.

Table 15

Codes for National Mathematics Standards

Code	National Standard	Description
M1	Number and Operations	Understand numbers, ways of representing numbers, relationships among numbers, and number systems; understand meanings of operations and how they relate to one another; compute fluently and make reasonable estimates
M2	Algebra	Understand patterns, relations, and functions; represent and analyze mathematical situations and structures using algebraic symbols; use mathematical models to represent and understand quantitative relationships; analyze change in various contexts
M3	Geometry	Analyze characteristics and properties of two- and three-dimensional geometric shapes and develop mathematical arguments about geometrical relationships; specify locations and describe spatial relationships using coordinate geometry and other representational systems; apply transformations and use symmetry to analyze mathematical situations; use visualization, spatial reasoning, and geometric modeling to solve problems
M4	Measurement	Understand measurable attributes of objects and the units, systems, and processes of measurement; apply appropriate techniques, tools, and formulas to determine measurements
M5	Data Analysis and Probability	Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them; select and use appropriate statistical methods to analyze data; develop and evaluate inferences and predictions that are based on data; understand and apply basic concepts of probability
M6	Problem Solving	Build new mathematical knowledge through problem solving; solve problems that arise in mathematics and in other contexts; apply and adapt a variety of

Code	National Standard	Description
		appropriate strategies to solve problems; monitor and reflect on the process of mathematical problem solving
M7	Reasoning and Proof	Recognize reasoning and proof as fundamental aspects of mathematics; make and investigate mathematical conjectures; develop and evaluate mathematical arguments and proofs; select and use various types of reasoning and methods of proof
M8	Communi-cation	Organize and consolidate mathematical thinking through communication; communicate mathematical thinking coherently and clearly to peers, teachers, and others; analyze and evaluate the mathematical thinking and strategies of others; use the language of mathematics to express mathematical ideas precisely
M9	Connections	Recognize and use connections among mathematical ideas; understand how mathematical ideas interconnect and build on one another to produce a coherent whole; recognize and apply mathematics in context outside of mathematics
M10	Represent-ation	Create and use representations to organize, record, and communicate mathematical ideas; select, apply, and translate among mathematical representations to solve problems; use representations to model and interpret physical, social, and mathematical phenomena

Note. Adapted from NCTM content standards.

Table 16

Codes for National Science Standards

Code	National Standard	Description
S1	Science as Inquiry	Abilities necessary to do scientific inquiry; understanding about scientific inquiry
S2	Physical Science	Properties of objects and materials; structure, properties, and changes of properties in matter; position and motion of objects; motion and forces; light, heat, electricity, magnetism; transfer of energy; chemical reactions; conservation of energy and increase in disorder; interactions of energy and matter
S3	Life Science	Characteristics and life cycles of organisms; organisms and environments; structure and function in living systems; reproduction and heredity; molecular basis of heredity; biological evolution; interdependence of organisms; regulation and behavior; populations and ecosystems; diversity and adaptations of organisms; matter, energy, and organization in living systems
S4	Earth & Space Science	Properties of earth materials; structure of the earth system; earth's history; earth in the solar system; energy in the earth system; geochemical cycles; origin and evolution of the earth system; origin and evolution of the universe
S5	Science & Technology	Understanding about science and technology; natural vs. man-made objects; abilities of technological design
S6	Science in Personal & Social Perspectives	Personal and community health; characteristics and changes in populations; population growth; natural resources; environmental quality; changes

Code	National Standard	Description
		in environments; science and technology in society; natural- and human-induced hazards; risks and benefits; science and technology in local, national, and global challenges
S7	History & Nature of Science	Science as a human endeavor; nature of science and scientific knowledge; history of science; historical perspectives

Note. Adapted from NSES content standards.

Reading. In Reading, all 41 extended indicators were deemed to be academic. The percentages of primary and secondary ratings for each national Reading content strand are displayed in Figure 1 . Categories E2, *Writing*, and E6, *Research*, were not chosen by any of the reviewers, probably because the Kansas Reading Standards do not address these national ELA standards. The majority (approximately 88%) of the Reading extended indicators were primarily defined by the *Reading* strand within the national standards. Most extended indicators (almost 90%) were not given a secondary rating.

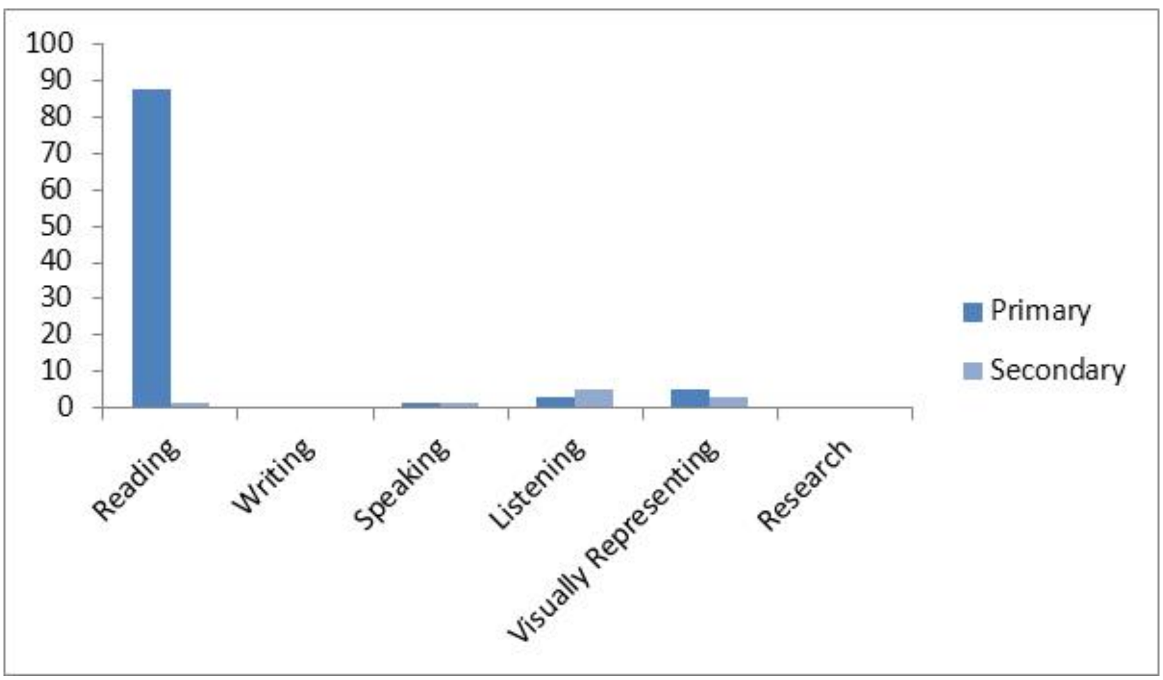

Figure 1. Percentages of Ratings for Primary & Secondary National Reading Standards Links With Extended Indicators.

Math. In Math, all but one extended indicator was deemed to be academic; the review team unanimously concluded that E.M.1.1.10 was a foundational skill. This extended indicator requires the student to identify symbols for dollar and cents notation and was not deemed to be academic because it was not found in any national standard and was therefore omitted from the following analyses. The percentages of primary and secondary links for the other 70 extended indicators with each national content strand are displayed in Figure 2. Approximately 42% of the Math extended indicators were primarily defined by the *Number and Operations* strand within the national standards, whereas the same percentage of secondary ratings was for the *Connections* strand.

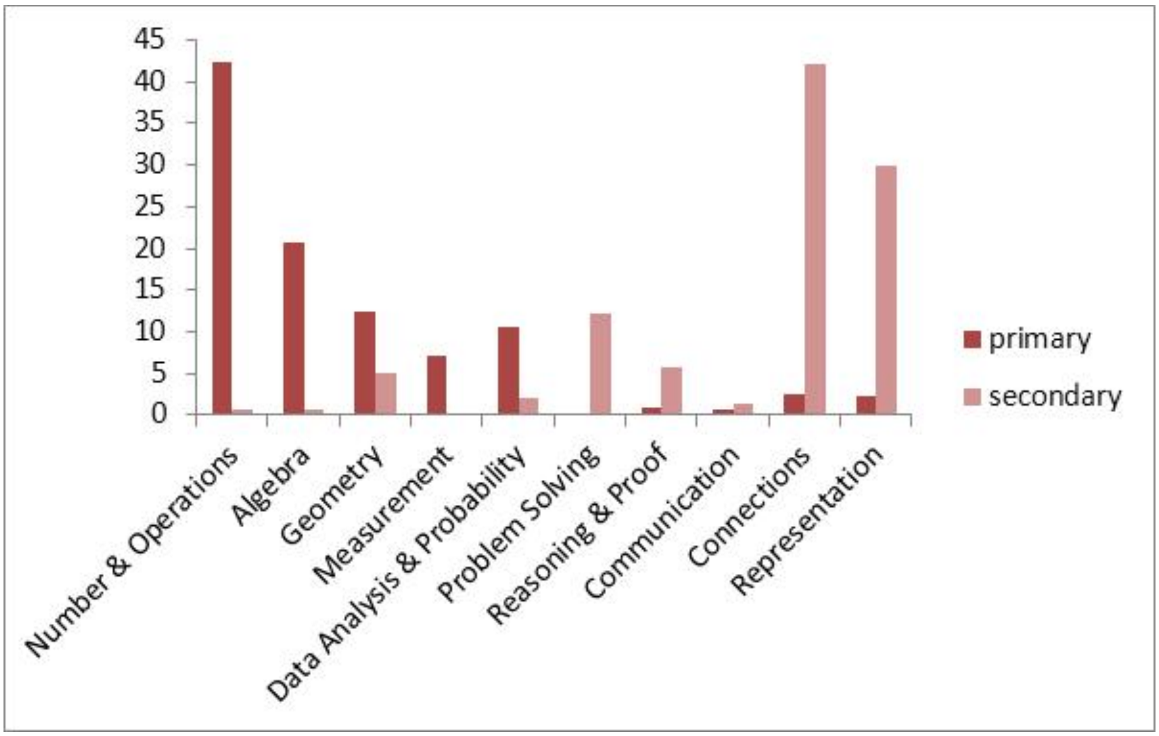

Figure 2. Percentages of Ratings for Primary & Secondary National Math Standards Links With Extended Indicators.

Science. In Science, all 43 extended indicators were deemed to be academic based on their links with the national standards. Percentages of primary and secondary ratings for each national content strand are displayed in Figure 3. Most of the extended indicators were defined by the *Science as Inquiry*, *Physical Science*, or *Life Science* strands. Only three indicators were given secondary ratings, all of which were for *Science in Personal & Social Perspectives*.

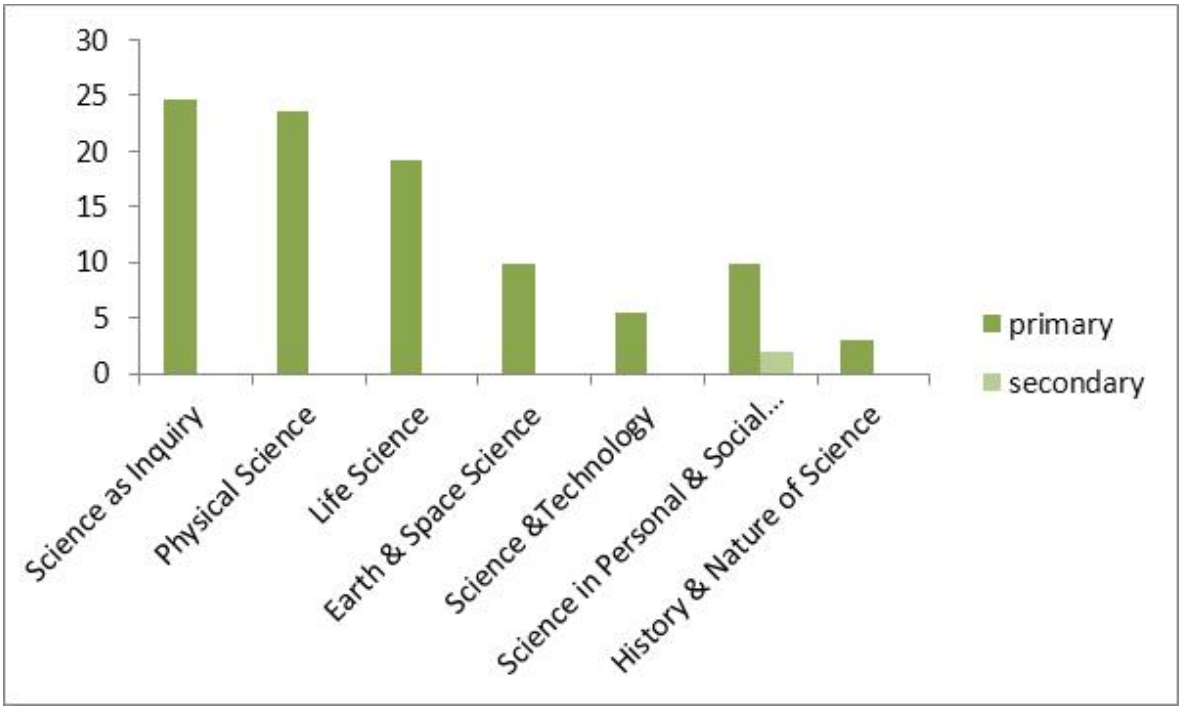

Figure 3. Percentages of Ratings for Primary & Secondary National Science Standards Links With Extended Indicators.

Grade-Level Content

After determining which extended indicators were academic for the purpose of the remaining alignment activities, the reviewers then evaluated the association of the extended indicators with grade-level content as described in Criterion 3 (Flowers et al., 2007, pp. 24–26). Based on the LAL protocol, a state's extended standards should be compared to the general curricular standards for both content and performance fidelity. Content fidelity is rated as a near link, a far link, or no link based on the consistency of the content of the referenced grade-level indicator to the extended indicator. Flowers et al. (2007, p. 12) provide the following example:

An extended standard of *Identify weather conditions* may have no link to the grade-level content standard *Analyze and identify types of clouds*. An extended standard of *Identify clouds* may be considered a far link because even though it is dealing with clouds, it still does not address the total domain of the original standard, which references types of clouds. A near link for an extended standard would be *Identify cumulus and not cumulus clouds.*

When no link was found between the extended indicators and the referenced grade-level indicators, further coding was provided to indicate the reason for the problem. Coding definitions for the content fidelity scale are provided in Table 17.

Table 17

Codes for Content Fidelity Scale

Code	Content Fidelity	Description
0	No link	The item does not measure the standard. Further coding for reason of no link:
	A	Mismatch—an error in identifying the correct standards
	B	Overstretch—the item has lost the intention and meaning of the standard
	C	Backmap—fitting a functional activity to an academic standard
	D	Nonspecific—standard is too broad to adequately align to the item
1	Far link	The item measures some of the original content standard
2	Near link	The standard is specific and the item clearly measures the content

Because each extended indicator can have a defined and intended reference to several different general curricular indicators within a grade and with similar general indicators at different grade levels, each extended indicator may have received near-link endorsements at one grade level as well as far-link or no-link endorsements for references to different general indicators or at other grade levels. For this reason, results are reported in terms of the percentage of total endorsements for individual extended indicators. Ratings for grade-level content fidelity, as well as the number of links between extended indicators and general curricular standards, are displayed in Appendix B.

Reading. One new Reading extended indicator—ER.2.2.3, the student demonstrates understanding of the cultural differences in language—had not yet been referenced to any general indicators, thus it was not included in any of the following alignment activities. For the remaining 40 Reading extended indicators and their referenced general indicators, a near link was chosen 85.9% of the time, a far link was chosen 10.9% of the time, and no link was chosen 3.2% of the time. All reviewers rated 24 extended indicators as having a near link to all of their associated general indicators (i.e., 100% of the ratings for the extended indicator were for a near link). Two extended indicators had no ratings for a near link. Only one extended indicator (ER.1.2.3) was rated by all teachers as having no link to all seven of the general indicators with which it was referenced. All reviewers agreed that there was a mismatch between this indicator and its linked general indicators. Three additional extended indicators had small percentages of no-link ratings because of nonspecificity or overstretch.

Math. For the 70 Math extended indicators, reviewers rated the fidelity of each content link between extended indicators and grade-level standards. Of these ratings, a near link was chosen 26.1% of the time, a far link was chosen 35.2% of the time, and no link was chosen 38.6% of the time. Only two extended indicators, EM.1.1.2 and EM.1.3.4, were rated by all reviewers as having a near link to all of their associated general indicators (i.e., 100% of the ratings for the extended indicator were in favor of a near link). Fourteen extended indicators had no ratings for a near link. One extended indicator (EM.3.2.3) was rated by all reviewers as having a far link to all of the general indicators. EM.2.4.2 was rated by all teachers as having no link to any of the four referenced general indicators because of a mismatch. Ten extended indicators were never rated as having no link to any of their associated general indicators. For the indicators that received some no-link ratings, the most common reason was mismatch and the second most common reason was overstretch. Neither backmapping nor lack of specificity was selected as a reason for no link between extended and general indicators.

Science. For the 43 Kansas Science extended indicators, a near link was chosen 31.0% of the time, a far link was chosen 50.0% of the time, and no link was chosen 19.0% of the time. Three extended indicators (ES.1.1.7, ES.6.1.2, and ES.6.3.1) were rated by all reviewers as having a near link to all of their associated general indicators (i.e., 100% of the ratings for the extended indicator were in favor of a near link). Ten extended indicators had no ratings for a near link. ES.2.1.1 was rated by all reviewers as having a far link to all of the general indicators. There were 22 extended indicators that were never rated as having no link to any of their associated general indicators. For the extended indicators that received some no-link ratings, the most common reason was mismatch and the second most common reason was overstretch. Lack of specificity was reported for two of the extended indicators. Backmapping was not indicated as a reason for any of the links between extended and general standards. Percentages of ratings for each category of content fidelity (i.e., near, far, none) across Reading, Math, and Science are displayed in Figure 4.

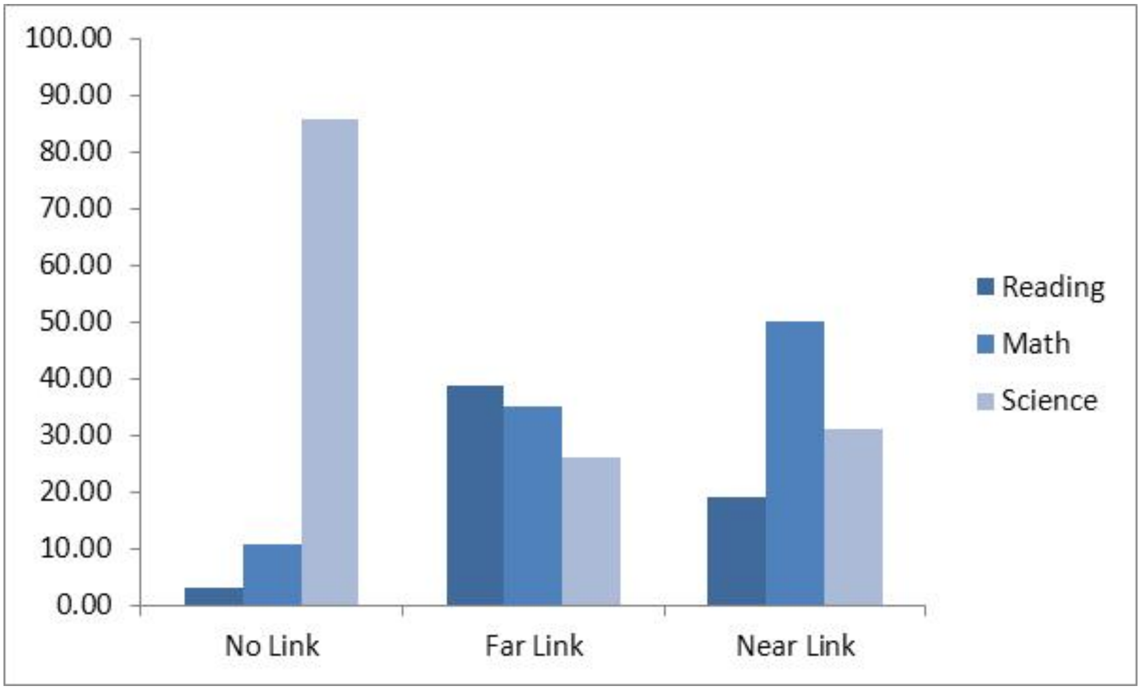

Figure 4. Percentages of Ratings for Content Fidelity by Content Area.

Grade-Level Performance

Criterion 3 of the LAL (Flowers et al., 2007, pp. 24–26) also describes performance fidelity between extended indicators and grade-level standards as the degree of consistency between the types of performance (e.g., select, identify, compare, analyze, and evaluate) required for the extended standards and the types of performance found in the grade-level content standards. The authors further explain that while content fidelity should remain high (i.e., either a near or far link), alternate achievement standards allow for an alternate level of performance (i.e., below grade-level performance) of the standards; thus, there is more flexibility in the expected levels of agreement. Furthermore, performance fidelity may be lower than content fidelity due to the difficulty of developing tasks and items on which students can show achievement (Flowers et al., 2007). Performance fidelity is rated on a three-point rating scale depending on the degree of consistency between extended and general performance requirements. When the performance of the extended indicator was identical to the performance of the general indicator, an "all" rating was made. A "some" rating was given when the performance of the extended indicator partially matched that of the general indicator. Finally, when the performance of the extended indicator was different from the performance of the general indicator, consistency was rated as "none." Ratings for grade-level performance fidelity are displayed in Appendix C.

Reading. An identical match occurred 75.1% of the time, a partial match 21.8% of the time, and no match 3.1% of the time. Eighteen extended indicators were rated by all reviewers as having an identical performance match to all of their corresponding general indicators. Conversely, six extended indicators did not have any ratings for an identical performance match to any of the general indicators. Three extended indicators were rated by all reviewers as having at least a partial performance match to all of their corresponding general indicators. Only one extended indicator (ER.1.2.3) was rated by all reviewers as having no match on performance with any of the grade-level indicators.

Math. Of the total number of ratings, there was an identical match 18.6% of the time, a partial match 29.4% of the time, and no match 51.9% of the time. Two extended indicators were rated by all reviewers as having an identical performance match to all of their corresponding general indicators. Conversely, 24 extended indicators did not have any ratings for an identical performance match to any of the general indicators. Two extended indicators were rated by all reviewers as having at least a partial performance match to all of their corresponding general indicators. Three extended indicators (EM.2.2.1, EM.2.2.2, and EM.2.4.2) were rated by all reviewers as having no match on performance with any of the grade-level indicators.

Science. Of the total number of ratings, there was an identical match 23% of the time, a partial match 54.1% of the time, and no match 45.9% of the time. Two extended indicators were rated by all reviewers as having an identical performance match to all of their corresponding general indicators. Conversely, there were 18 extended indicators that did not have any ratings for an identical performance match to any of the general indicators. Four extended indicators were rated by all reviewers as having at least a partial performance match to all of their corresponding general indicators. Two extended indicators (ES.4.3.2 and ES.5.1.1) were rated by all reviewers as having no match on performance with any of the grade-level indicators. Finally, there were 12 indicators that were never rated as having no match with grade-level performance. Percentages of ratings for each category of performance fidelity (i.e., none, some, all) across Reading, Math, and Science are displayed in Figure 5.

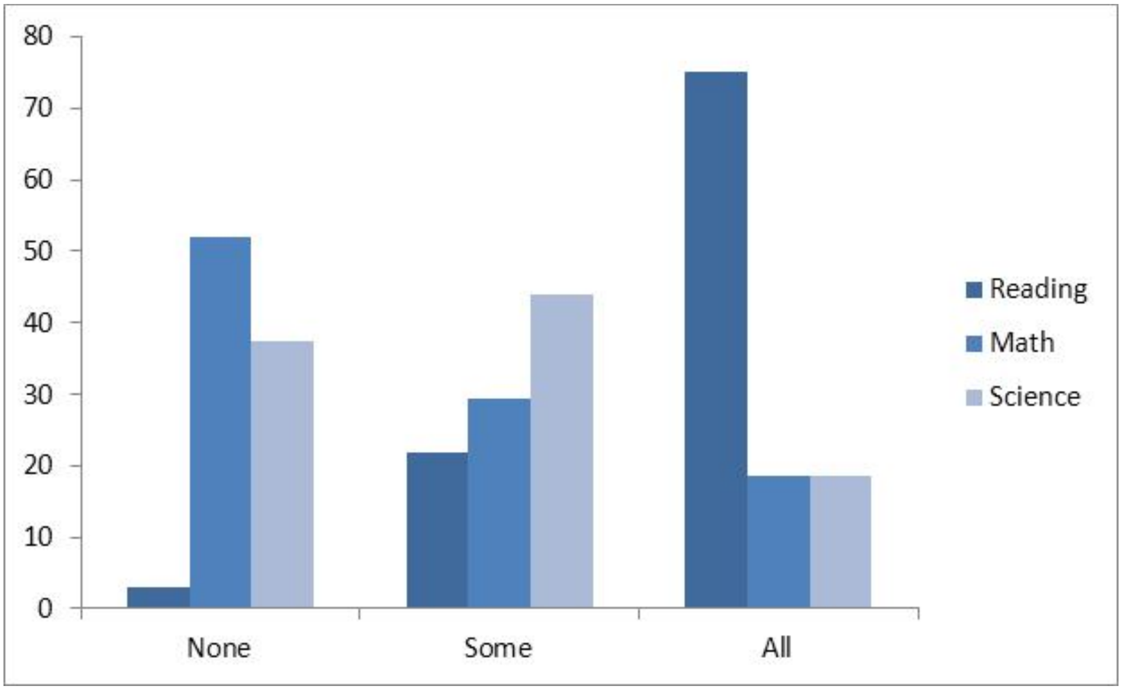

Figure 5. Percentages of Ratings for Performance Fidelity by Content Area.

Categorical Concurrence

Four key alignment constructs of categorical concurrence, depth of knowledge, range of knowledge, and balance of representation deal with four related aspects of alignment between a state's AA-AAS, extended standards, and general standards. According to Criterion 4 (Flowers et al., 2007, p. 26), categorical concurrence is determined through an inspection of the consistency of the content between general curricular standards, the extended standards, and the AA-AAS. The purpose of investigating categorical concurrence is to evaluate whether students assessed by a state's AA-AAS are reliably assessed through a sufficient number of test items or observations on the same content strands as students taking a general assessment. However, that content is likely to be of lower difficulty than the general assessment and it may not cover the same range. Depth of knowledge is an evaluation of the cognitive complexity of test items and/or extended indicators used for assessment on an AA-AAS. All extended indicators that were identified as having academic content were reviewed by content expert review teams for depth of knowledge. Range of knowledge, in contrast, looks at whether students with significant disabilities are assessed on a similar breadth of content as students who participate in a general assessment. The LAL proposes a goal of at least 50% coverage for the objectives or benchmarks within the general curricular standards. Finally, balance of representation reflects the distribution of AA-AAS items or extended indicators with respect to the general content standards and objectives. Balance of representation is evaluated on the number of "hits" (i.e., AA-AAS items or extended indicators) that are deemed to be aligned with the general content standards on the basis of either a near or far rating for content fidelity. These four topics will be described in the next sections.

Categorical concurrence measures the number of assessment items matching the content of the general curricular standards with the purpose of determining that reliable assessment can occur for students with significant disabilities over that content. In Kansas, the KAA assessment items are selected on the basis of extended indicators under each general curricular standard, a practice that conforms exactly to the LAL guidelines. For Reading and Math, with two and four standards respectively, extended indicators for assessment must be chosen from each standard, with extra indicators chosen from any standard at the IEP team's discretion. In Science, which has seven standards, extended indicators are selected from each of the first four standards with an additional indicator chosen from any one of the standards at the IEP team's discretion. Therefore, the KAA is consistent with the general curricular standards in Math and Reading and with more than half of the general curricular standards in Science.

Three unique tasks are required to measure each extended indicator, which is less than the six items recommended by the LAL (Flowers et al., 2007, p. 27). However, each of these tasks or activities must contain at least five trials, which themselves could be test items or performance events. This means that there are, at minimum, 15 individual items or examples of performance to inform the scoring of a student's achievement on each extended indicator. This is consistent with the LAL, which requires each individual with significant disabilities to demonstrate reliable performance on each standard.

In Mathematics and Reading, where there are more required KAA test items than there are standards, KAA students can unequivocally be said to have met the definition of categorical concurrence. Since only Mathematics and Reading are assessed at grades 3, 5, 6, and 8, KAA test takers will have complete assessment coverage of all six general curricular standards in those grades. In Science, which is assessed at grades 4 and 7 and again in high school, KAA students are assessed on four or five out of seven standards, for a concurrence of 57% to 71%. Categorical concurrence percentages are shown in Table 18.

Table 18

Categorical Concurrence for KAA in Reading, Mathematics, and Science

Content Area	Number of Standards	Number of KAA Items	Categorical Concurrence
Reading	2	5	100%
Mathematics	4	5	100%
Science	7	5	57–71%

Depth of Knowledge

Depth of knowledge addresses the cognitive complexity of the extended indicators or assessment items, which is expected to differ on an AA-AAS from the cognitive demands of general assessment items (Flowers et al., 2007). According to Criterion 4 of the LAL (Flowers et al., 2007, pp. 26–34), the extended standards should be skewed to lower levels than the general state standards, and this lower cognitive demand defines the difference between grade-level achievement and alternate achievement. As recommended by Flowers et al. (2007), a modified version of Bloom's learning taxonomy was utilized in the present study; specifically, "the scale was extended downward to incorporate a level that captures the response processes for students with the most significant cognitive disabilities" (p. 13). The final, six-level taxonomy is displayed in Table 19.

Table 19

Codes for Depth of Knowledge Scale

Code	Depth of Knowledge	Description
1	Attention	touch, look, vocalize, respond, attend
2	Memorize/recall	list, describe facts, identify, state, define, label, recognize, record, match, recall, relate
3	Performance	perform, demonstrate, follow, count, locate, read
4	Comprehension	explain, conclude, group/categorize, restate, review, translate, describe concepts, paraphrase, infer, summarize, illustrate
5	Application	compute, organize, collect, apply, classify, construct, solve, use, order, develop, generate, interact with text, implement
6	Analysis, Synthesis, Evaluation	pattern, analyze, compare, contrast, compose, predict, extend, plan, judge, evaluate, interpret, cause/effect, investigate, examine, distinguish, differentiate, generate
0	Can't score/too vague	

Note. Adapted from Flowers, et al. (2007, p. 56).

Reading. The six Reading content reviewers coded each of the 41 extended indicators with regard to their depth of knowledge according to the scale above. Of the these ratings, none endorsed the *attention* category; 20.7% endorsed *memorize/recall;* 17.1% of the ratings were for *performance*; 10.2% were for *comprehension*; 14.2% endorsed the *application* category; 37.8% endorsed *analysis, synthesis, and evaluation*; and all extended indicators were scored. Thus, most extended indicators were rated at the *analysis, synthesis, and evaluation* level and none were rated at the lowest level, *attention*. Depth-of-knowledge ratings for extended indicators by content area are displayed in Appendix D.

Math. Groups of six content reviewers coded each of the 70 extended indicators that were evaluated as having academic content with regard to their depth of knowledge according to the scale in Table 19. Of the these ratings, none endorsed the *attention* category; 15.7% endorsed *memorize/recall;* 18.8% of the ratings were for *performance*; 12.9% were for *comprehension*; 31.2% endorsed the *application* category; 5.7% endorsed *analysis, synthesis, and evaluation*; and all extended indicators were scored (i.e., none were rated as too vague to score). Thus, most extended indicators were rated at the *application* level and none were rated at the lowest level, *attention*.

Science. The group of six content reviewers collaboratively coded 25 of the 43 extended indicators with regard to their depth of knowledge. Due to time constraints, the other 18 indicators were evaluated by groups of three raters. Of these ratings, none endorsed the *attention* category; 21.3% endorsed *memorize/recall;* 12.4% of the ratings were for *performance*; 15.3% were for *comprehension*; 20.3% endorsed the *application* category; 23.3% endorsed *analysis, synthesis, and evaluation*; 5.4% indicated that the extended indicator could not be scored; and 2% of the data were missing. The ratings for each depth-of-knowledge category across the three content areas are displayed in Figure 6.

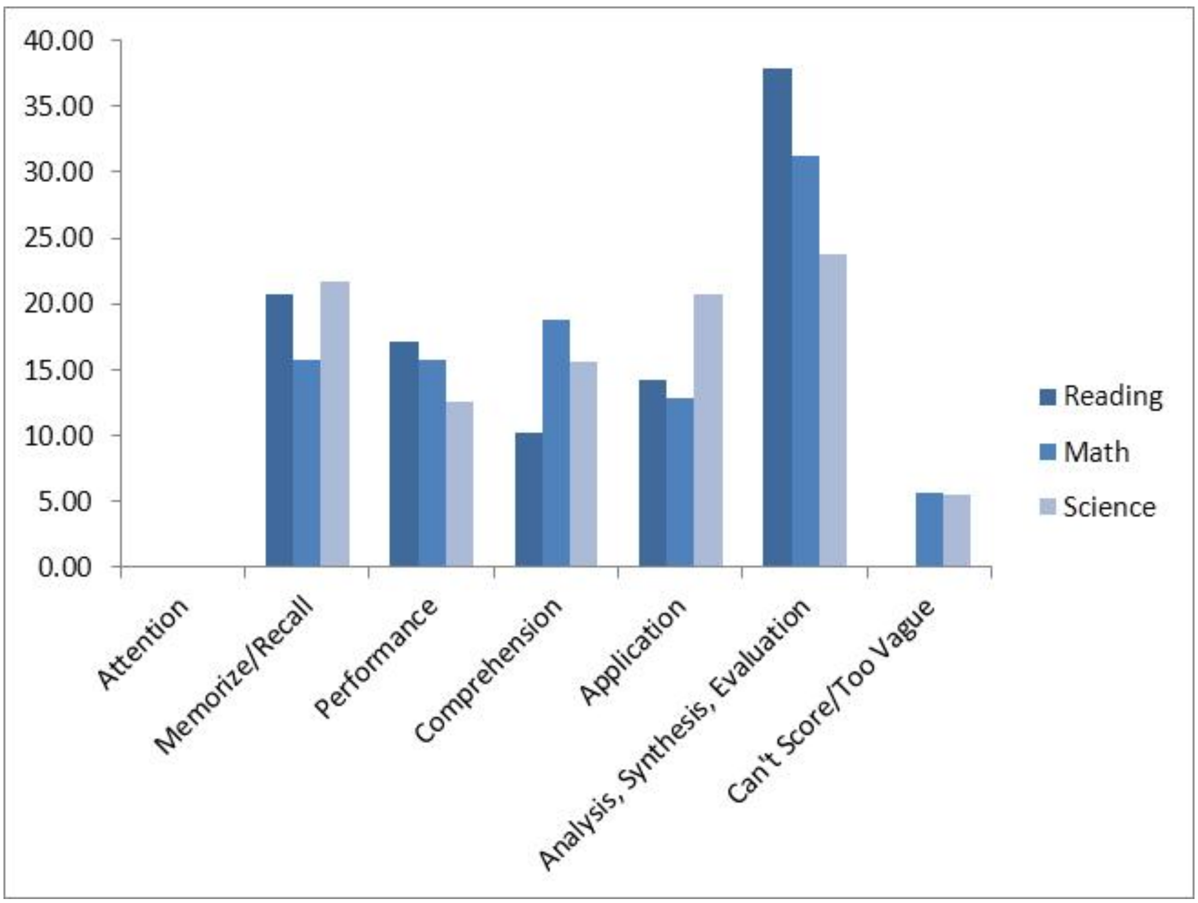

Figure 6. Percentages of Ratings for Depth of Knowledge by Content Area.

LAL also recommends that the general standards be evaluated on depth of knowledge in order to contrast these with the corresponding referenced extended indicators. In Kansas, the general curricular standards have previously been evaluated for depth of knowledge. However, taxonomies specific to each content area were used, so direct comparisons are not feasible. In Science, however, KSDE has evaluated the extended indicators for depth of knowledge using Bloom's taxonomy. KSDE makes information on depth of knowledge for all content areas available on its website (www.ksde.org).

Range-of-Knowledge Correspondence

Range-of-knowledge correspondence evaluates the extent to which AA-AAS students are being assessed on the breadth of knowledge covered by the extended indicators. For this alignment construct, the number of assessment items, tasks, or performances is irrelevant. The number of assessed indicators per standard is the crucial variable. The LAL manual (Flowers et al., 2007, p. 29) warns that assessment items may be aligned to only a few objectives within the content standard. Alignment is said to be achieved with at least 50% coverage and weakly achieved if there is 40–50% coverage. Given that content standards may have many objectives and an AA-AAS is likely to have fewer items than a general assessment, meeting these goals may be difficult.

This aspect of alignment is particularly relevant to Kansas. For example, Extended Reading Standard 1 has 28 objectives, or extended indicators, and anywhere from one to four of those indicators would be included on the KAA for content coverage of 4% to 14%. On Extended Reading Standard 2, with 12 extended indicators, content coverage would be correspondingly higher, from 8% to 33%. With one extended indicator assessed in Standard 1 and four in Standard 2, content coverage for the entire Reading assessment would be the average of the coverage for those two standards, or 18.5%. A single value for content coverage cannot be assigned to any one extended standard because teachers are able to select the standards to be assessed on the KAA after minimum requirements have been met. Therefore, content coverage can best be expressed as a range of values based on the minimum and maximum number of extended indicators per standard that can be assessed by the KAA (Table 20).

Table 20

Percent of Content Coverage for Extended Standards in Reading, Mathematics, and Science

Content Area	St. 1	St. 2	St. 3	St. 4	Sts. 5–7	Overall
Reading						
Number of ext. indicators	28	12	NA	NA	NA	40
Range of Knowledge*	4–14%	8–33%				11–18%
Mathematics						
Number of ext. indicators	31	15	16	9	NA	71
Range of Knowledge**	3–6%	7–13%	6–13%	11–22%		8–10%
Science						
Number of ext. indicators	9	8	9	7	10	43
Range of Knowledge**	11–22%	13–25%	11–22%	14–29%	0–10%***	12–13%

Note. *based on 1–4 extended indicators per standard; **based on 1–2 extended indicators per standard; ***a maximum of one extended indicator may be chosen from these three standards.

This table illustrates the differences in assessment coverage for the various content areas at the indicator level. None of the content areas reaches even the weak alignment of 40% coverage. Because the range-of-knowledge value is the percent of standards with at least 50% of the objectives having one or more hits, it is apparent that the range of knowledge for Kansas is 0% in each content area.

Balance of Representation

Balance of representation refers to the distribution of KAA items within the general curricular standards and objectives. For example, states may choose to focus their AA-AAS items or tasks on particular objectives to the exclusion of others. In Kansas, the standard–benchmark–indicator hierarchy is used for both general and extended standards. At the finest grain, or indicator level, the general indicators are referenced to the extended indicators such that each general indicator has from one to 12 linked extended indicators. In this way, the extended indicators fully complement the general indicators. No general indicator fails to reference one or more complementary extended indicators.

Balance of representation is evaluated in terms of KAA tasks, which are created by the teacher in order to assess a student's performance on a single extended indicator. Obviously, the number of assessment tasks that can be requested of the student must be limited, and these limits will affect the content coverage that can be attained. The KAA requires five extended indicators to be selected for assessment within each subject area. Math and Reading are assessed at grades 3–8 and once in high school. Science is assessed at grade 4, grade 7, and high school. For the two standards of Reading and four standards of Mathematics, each standard must be represented by the selection of at least one extended indicator, while the remaining extended indicators may be selected from any standards. In Science, the first four of the seven standards must be represented by at least one extended indicator, and the fifth extended indicator can be chosen from any standard. In each case, only one extended indicator may be chosen to correspond to any general indicator, so each assessed extended indicator addresses a different general content indicator. However, any one standard in Reading or Mathematics may be represented by only a single extended indicator. In Science, four standards will be represented by at least one extended indicator, but the three remaining standards may not be assessed at all. This relationship was first captured by range of knowledge and can be further analyzed by assessing the balance of representation.

Balance of representation is evaluated using the following formula from the LAL manual (Flowers et al., 2007, p. 30):

$$1-(\sum_{i=1}^{k}\left|\frac{1}{O}-\frac{I_k}{H}\right|)/2$$

In this formula, O represents the total number of objectives judged to be aligned for the standard, I_k is the number of items corresponding to objective k, and H is the total number of items aligned within the content standard. The recommended value for adequate balance of representation is .7 with .6 indicating a weakly acceptable balance. In Kansas, this formula must be computed over the range of possible assignment of indicators to standards, because the number of assessed indicators within a standard varies by student, resulting in different values for I_k/H. Using Kansas data for Reading Standard 1 as an example, O is 27 out of 28 objectives (extended indicators) aligned with general content based on a near or far link for content fidelity, I_k ranges from 1 to 4 for the indicators to be assessed, and *H* is 5. For Reading Standard 2, I_k again ranges from 1 to 4 for the remaining indicators to be assessed. The balance of representation index or range of indices, in the case of Science, is shown in Table 21. The index for Reading just misses the weak alignment goal of .60, Mathematics is weakly aligned, and Science meets the adequate alignment goal of .70. The rationale for the differences is that Reading has a small number of standards with many objectives per standard while Science has a larger number of standards with fewer objectives.

Table 21

Range of Balance of Representation Indices for Reading, Mathematics, and Science

Content Area	Number of Standards	Number of Assessed Indicators	Balance of Representation Index
Reading	2	5	.56
Mathematics	4	5	.64
Science	7	5	.75–.80

To summarize the four key constructs that comprise Criterion 4 of the LAL, the KAA extended standards and assessment items show excellent categorical concurrence because of the comprehensive linking of the extended standards with the general standards and the requirement that assessed indicators be distributed among the standards for maximum coverage. In terms of depth of knowledge, the extended indicators were consistently rated as suggestive of and amenable to higher levels of cognitive complexity. The extended indicators have a high rate of alignment on the basis of content fidelity as measured by the number of endorsements as near and far links to the general curricular indicators. These endorsements are the first step in determining the final two attributes in Criterion 4. Where the KAA falls short is in the content coverage of the extended standards as assessed by the range of knowledge and balance of representation constructs. The range-of-knowledge evaluation reveals that only a small proportion of each content area is actually assessed by the KAA items and tasks—always less than the recommended 50%. The balance of representation indices for the three content areas shows that Reading fails to meet recommended LAL values while Mathematics is weakly aligned and Science is adequately aligned.

As a follow-up to content and performance fidelity between the Kansas Extended Standards and the grade-level standards described in Criterion 3 of the LAL (Flowers et al., 2007, pp. 24–26), reviewers were asked to determine which general indicator best matched a given extended indicator. Since Kansas allows for extended indicators to be linked to more than one general indicator, content reviewers were instructed to determine the best link between the extended and general indicators. In other words, of the various general indicators associated with one extended indicator, which link is optimal? Reviewers were instructed to choose a primary match and a secondary match. However, reviewers were allowed to choose more than one primary and secondary match; thus, the number of possible ratings for each extended indicator varied across indicators. Furthermore, in Math, each extended indicator could comprise several sub-indicators and reviewers rated each sub-indicator match with the corresponding general indicator. These sub-indicator ratings were summed for each indicator, which also resulted in differing numbers of possible ratings for each extended indicator depending on the number of associated sub-indicators. Ratings for primary and secondary matches by content area are displayed in Appendix E.

Reading. For Reading, six content reviewers rated each of the four to 10 links between the 40 extended indicators that are referenced to general standards with regard to a primary match and a secondary match. Most extended indicators were rated as having one primary and one secondary match; however, an additional choice for a primary and/or secondary match was made for several indicators. As an example, for extended indicator E.R.1.4.3, all six ratings for primary match were for general standard R.3.1.4.2 and all six secondary match ratings were for R.4.1.4.2. On the other hand, of the six primary match ratings for extended indicator E.R.2.1.7, all of them were for R.7.2.1.3; of the 12 secondary match ratings, half were for R.8.2.1.3 and the other half were for R.HS.2.1.3. Thus, for E.R.1.4.3, the strongest link to grade-level standards was at the 3rd grade level while the second strongest link was at the 4th grade level. Conversely, for E.R.2.1.7, the strongest link to grade-level standards was at the 7th grade level while the next two strongest links were at the 8th and high school grade levels.

In general, most primary matches tended to be at the lower grade levels, with no primary matches at the high school level. While secondary matches tended to follow the same pattern as primary matches, this pattern started at 4th grade and continued on through high school. Figure 7 displays the number of primary and secondary matches between Reading extended indicators and grade-level standards for each grade.

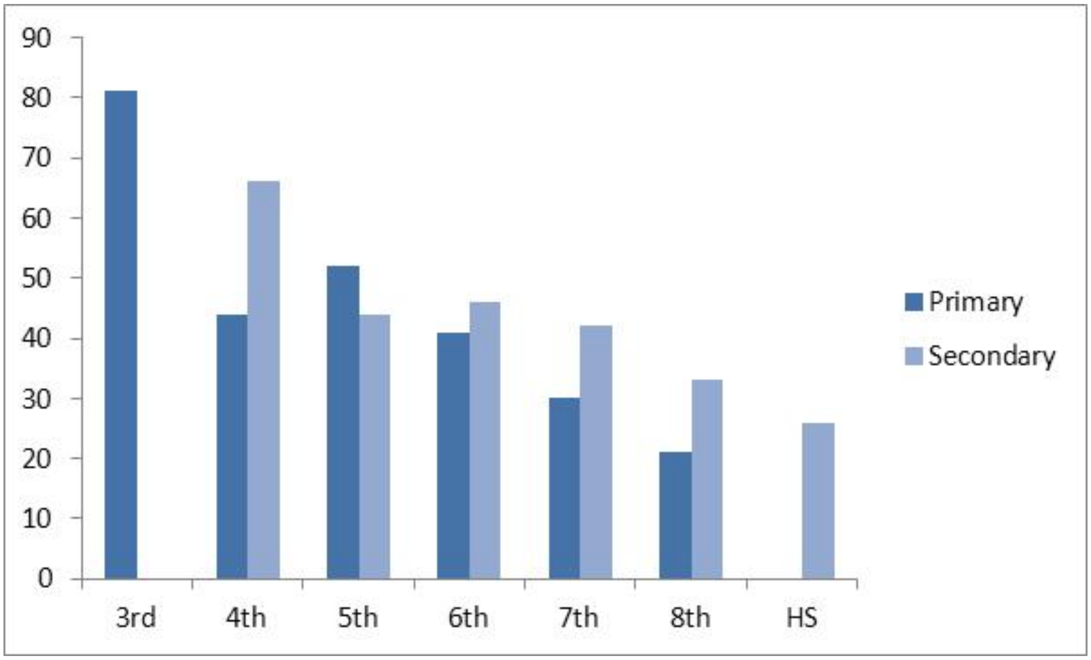

Figure 7. Frequency of Reading Grade-Level Standards Rated as a Primary or Secondary Match With Extended Indicator by Grade.

Math. In Math, 12 reviewers rated anywhere from two to 42 links between the 70 academic extended indicators and their corresponding grade-level standards. While several extended indicators had one primary and one secondary match, most were rated as having several primary and secondary matches with grade-level standards. For example, extended indicator E.M.1.4.1 was rated as having two primary matches and five secondary matches with grade-level standards. Of the 12 primary ratings, three were for general standard M.3.1.1.K3 and nine were for M.4.1.4.K6a, b, and d. It should be noted that the latter general standard comprises three sub-indicators that reviewers rated separately. Of the 18 secondary ratings, 12 were divided evenly (three each) across M.3.1.1.K4, M.3.1.4.A1a, M.4.1.4.A1a, and M.8.1.4.K2a while six of the ratings were for M.5.1.4.A1a, and e. Thus, the strongest links to grade-level standards were made at the 3rd and 4th grade levels while the next strongest links were made at the 3rd, 4th, 5th, and 8th grade levels. There were also a few extended indicators that were rated as having only a secondary match and no primary match with grade-level standards. As an example, for extended indicator E.M.2.4.1, there were three ratings for a secondary match with M.8.2.4.A2 but no ratings for a primary match.

In general, most primary matches tended to be at the lower grade levels, particularly at 3rd grade, with a decrease in the number of matches as grade level increased. Secondary matches to grade-level standards followed the same general pattern, although there was a slight increase in the number of matches at the 6th grade level than at 4th or 5th. Figure 8 displays the number of primary and secondary matches between Mathematics extended indicators and grade-level standards for each grade.

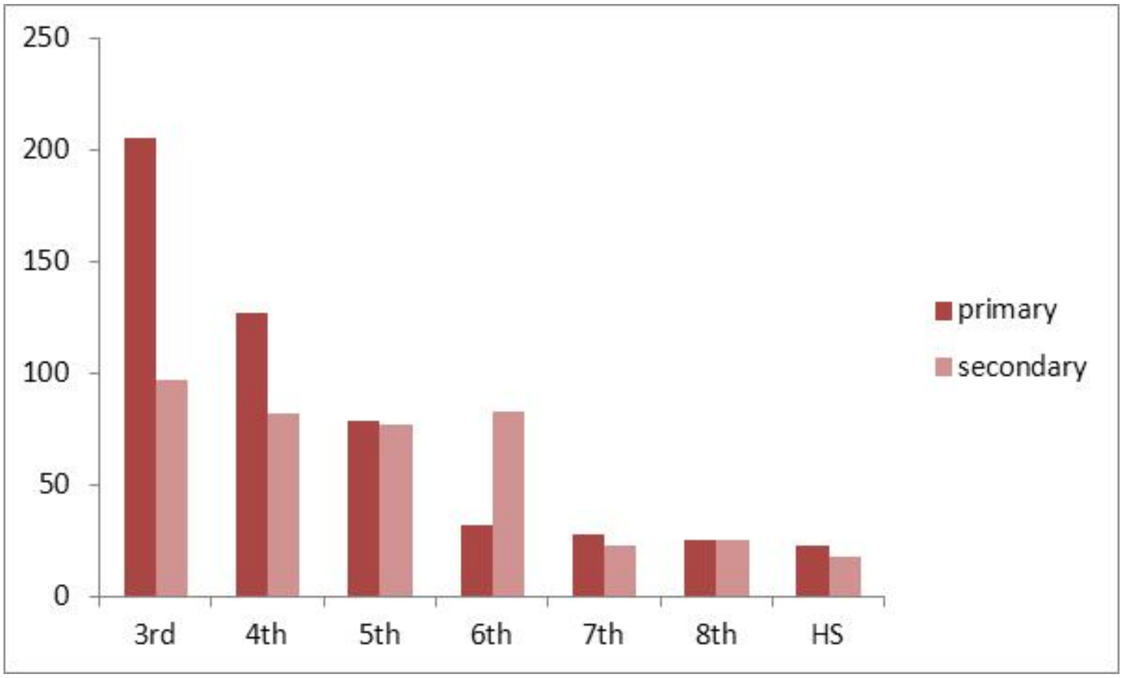

Figure 8. Frequency of Math Grade-Level Standards Rated as a Primary or Secondary Match With Extended Indicator by Grade.

Science. In Science, six content reviewers rated each of the one to 12 links between 43 extended indicators and their corresponding general standards with regard to a primary match and a secondary match. It should be noted that Science is only tested at 4th, 7th, and high school grade levels. Most extended indicators were rated as having at least one primary and one secondary match, with several indicators rated as having more than one choice for a primary and/or secondary match. However, some extended indicators were rated as having only a primary match and no secondary match. For example, extended indicator E.S.3.2.1 had 12 primary rating endorsements, six of which were for a link with general standard S.4.3.2.1 and the other six for S.HS.3.7.3; and there were six secondary ratings, all of which were for the link with S.7.3.2.1. Thus, the strongest two links were at the 4th and high school grade levels while the second strongest link was at the 7th grade level. On the other hand, E.S.2.1.5 had six endorsements for a primary link at the 4th grade level with S.4.2.1.2 and no secondary ratings.

Again, most primary matches tended to be at the 4th grade level, with the fewest number of matches at the high school level. However, most secondary matches occurred at 7th grade, with fewer matches at the 4th and high school grade levels. Figure 9 displays the number of primary and secondary matches between Science extended indicators and grade-level indicators for each grade.

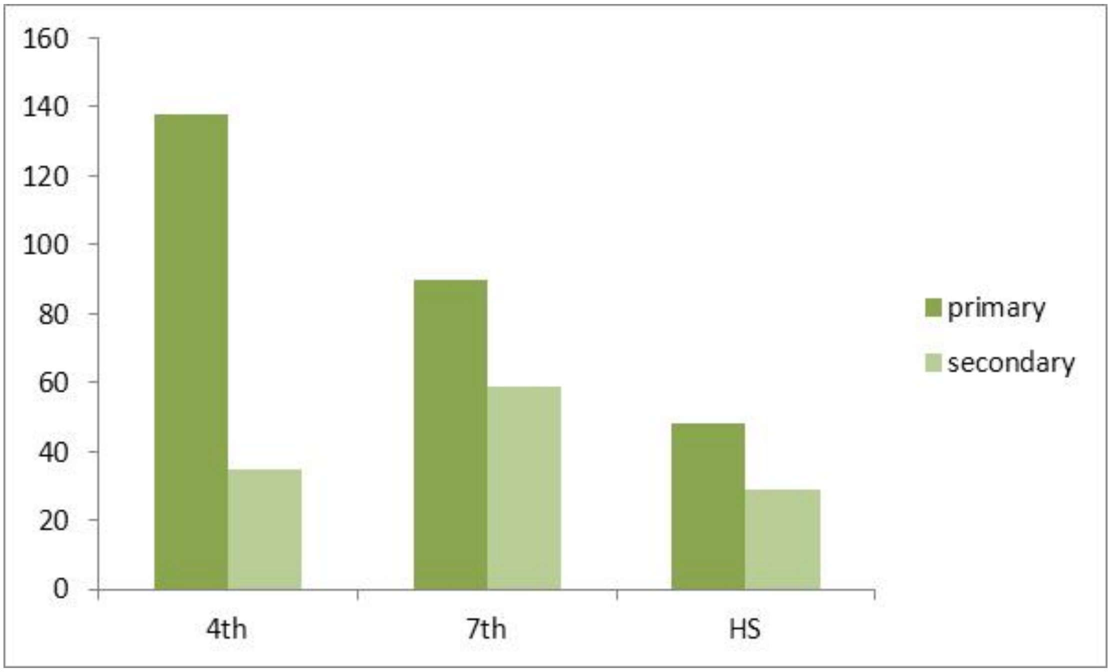

Figure 9. Frequency of Science Grade-Level Standards Rated as a Primary or Secondary Match With Extended Indicator by Grade.

Differentiation of Content Across Grades/Grade Bands

After conducting the content alignment portion of the review, content expert reviewers were asked to complete Appendix E of the LAL (Flowers et al., 2007, p. 72) to indicate the extent to which extended indicators differentiated between grade levels or grade bands. Reviewers were asked to form holistic conclusions about the percentage of extended indicators that represented broader or deeper content in higher grades, the presence of new skills at higher grades (as opposed to simply repeating identical skills), and the percentage of the indicators at lower grades that represented prerequisite skills for higher grades.

This particular rating form was confusing for content reviewers. With five categories to rate, four referred to changes in indicators at higher grade levels while one referred to features of extended indicators at lower grade levels, which is conceptually inconsistent. The form itself was structured as a matrix with types of vertical relationships in the leftmost column and subsequent columns representing categories marked 75% (clear), 50% (partial), 25% (limited), and 0% (none). Therefore, if reviewers believed that 20% of the extended indicators were to be rated as fitting one of the vertical relationships, they had to enter a different percent than those offered and weren't sure of which category to use. The format also suggested that the total of the percentages entered into each category should be 100%, yet the higher-grade-level change categories were not mutually exclusive with the lower-grade-level category of prerequisite skills. In other words, lower-grade-level indicators might be prerequisite skills for higher-grade-level indicators that measured new or deeper content. A more straightforward approach would have been to skip the matrix and simply ask reviewers to rate the percentage of indicators that defined skills in each of the five categories. Reviewers would then have had more flexibility to assign percentages instead of feeling constrained by the existing column categories, which were unnecessary.

Reviewers' results are shown in Table 22. It is evident that Science and Reading reviewers found a larger percentage of identical indicators at different grade levels, with a small percentage of prerequisite skills at lower grade levels and new indicators at higher grade levels. In Mathematics, reviewers found that the largest percentage of extended indicators represented broader application of content at higher grade levels, with the next-largest percentage measuring prerequisite skills at lower grade levels.

Table 22

Percentage of Indicators Showing Differentiation in Content by Grade Levels or Grade Bands

Vertical Relationships	Math	Science	Reading
Broader	40	0	0
Deeper	23	0	0
Prerequisite	33	7	22
New	16	10	10
Identical	21	83	60
Total	100	100	100

Summary

The content expert review portion of the Kansas Alternate Assessment Alignment Study was conducted to evaluate the alignment of the Kansas Extended Standards with the Kansas general standards across three content areas: Reading, Mathematics, and Science. Alignment was evaluated with respect to links with national standards in each subject area, fidelity with grade-level content and performance, specific alignment of extended indicators in terms of the best matches for those indicators among the general indicators, differentiation of content by grade/grade band, and four additional constructs related to alignment. These four constructs included categorical concurrence between the content of the general and extended standards, cognitive depth of knowledge associated with each extended indicator, range-of-knowledge correspondence between assessed indicators and the general content standards, and balance of representation of the assessed indicators with respect to the extended indicators.

All 24 content reviewers had at least a bachelor's degree, 18 had a master's degree, and five had another type of graduate degree. Furthermore, 19 of the reviewers had conducted professional development for teachers in their respective content area, 18 had played a leadership role in curriculum planning in their school or district, five were certified by the National Board for Professional Teaching Standards, and one had taught teachers in a higher education setting.

Content expert reviewers were first asked to determine if the content of the extended standards was academic or foundational. Content was deemed to be academic if it could be defined by a strand within the national standards for the given content area. Across the three content areas, only one extended indicator (in Math) was unanimously determined to be foundational rather than academic. All other indicators were linked to national standards and retained for the subsequent alignment activities.

In Kansas, the standard–benchmark–indicator hierarchy is the framework for both general and extended standards. The general indicators are referenced to the extended indicators such that each general indicator has anywhere from one to 12 linked extended indicators; conversely, each extended indicator has from two to 42 links with general indicators, both within and across grade levels. In this way, the extended indicators fully complement the general indicators. Once extended indicators were determined to be academic, these multiple links with corresponding general indicators were evaluated with respect to both content fidelity and performance fidelity. In other words, the degree of consistency between the extended indicators and their referenced grade-level indicators was rated with respect to content as well as the type of performance required.

With respect to content fidelity, Reading extended indicators were rated as having more near links to their general indicators than Math or Science indicators. Reading also had the smallest number of extended indicators that were rated as having no link with their corresponding general indicator. Both Math and Science extended indicators had more no-link or far-link ratings with general indicators than near-link ratings.

In Reading, there were only four extended indicators that had no-link ratings. Reasons for these ratings included several mismatches (i.e., an error occurred in identifying the correct standards), one overstretch (i.e., the intention and meaning of the standard was lost), and one nonspecific reason (i.e., standard was too broad for adequate alignment). In Math and Science, the most common reason indicated for no-link ratings was a mismatch between the extended indicator and grade-level standards; the second most common reason was overstretch.

In terms of performance fidelity, Reading extended indicators were again rated more often as having an identical performance with their corresponding general indicators than Math and Science. Similarly, Reading also had the fewest ratings for extended indicators that were not identical to general indicator performance. Math extended indicators were rated most often as having no performance fidelity with general indicators.

While these alignment activities were relatively straightforward and easy for the content reviewers to complete, summarizing the data in a meaningful way became more of a challenge. LAL protocol requires reviewers to come to a group consensus on the different variables analyzed; however, for the Kansas alignment study, reviewers were allowed to disagree with one another. In order to capture dissenting ratings across the groups of reviewers, results were summarized across the multiple references for each extended indicator as well as across the reviewers on the content review team. Percentages were calculated to describe relative proportions of each category. However, the total number of ratings across categories differed based on the number of references and the number of reviewers, making direct comparisons of proportions potentially misleading. For instance, while more extended indicators were evaluated in Math, in many cases, there were also fewer raters evaluating those indicators than for Reading or Science. In any case, the descriptive data provided in this report does support the intent of the study, which was to provide information regarding the quality of alignment between Kansas extended indicators and grade-level standards.

Following the evaluation of content and performance fidelity, Kansas educators were asked to assess the depth of knowledge of each extended indicator. Depth of knowledge is a measure of cognitive complexity and is expected to be lower in an AA-AAS than in a general assessment. For all three content areas, the *analysis, synthesis, and evaluation* category was endorsed most frequently. The *attention* category was not endorsed for any extended indicator within any of the three content areas, and only a small percentage of ratings indicated that the extended indicator was too vague to score. The other four depth-of-knowledge categories were rated variously across Reading, Math, and Science extended indicators. Thus, content reviewers endorsed generally high levels of cognitive complexity for the extended indicators, which implies that indicators selected for assessment do not limit the scope of the cognitive activities that can be requested of students with significant disabilities.

Reviewers next determined which link between the extended indicator and referenced grade-level indicators was the best (as indicated by a primary rating) and which was the second best (as indicated by a secondary rating). Across all three content areas, there was a general pattern of primary matches being identified the most at lower grades and the least at higher grades. This was also generally true for secondary matches, with the exception of Science indicators, which had more secondary ratings at the 7th grade level than at 4th grade or high school. Summarizing these data again presented some challenges as multiple primary and secondary ratings were allowed for each extended indicator. In other words, one extended indicator could have more than one primary match with grade-level standards. This flexibility became especially apparent in Math, where reviewers rated several sub-indicators as corresponding to one extended indicator. Most often, reviewers rated all sub-indicators as primary or secondary matches, but this was not always the case (i.e., some reviewers would rate one sub-indicator as a primary or secondary match and not rate the other sub-indicators). For summary purposes, data for all content areas were collapsed across reviewers and further collapsed across sub-indicators in Math.

Reviewers then evaluated differentiation of content across grades or grade bands by forming holistic conclusions about the percentage of extended indicators that represented broader or deeper content in higher grades, the presence of new skills at higher grades (as opposed to simply repeating identical skills), and the percentage of the indicators at lower grades that represented prerequisite skills for higher grades. This particular rating form, in matrix format, was confusing and unnecessarily complex. Nonetheless, reviewers concluded that Science and Reading indicators were largely identical at higher grades, with a small percentage of prerequisite skills at lower grade levels and new skills at higher grades. In Mathematics, reviewers found that the largest percentage of extended indicators represented broader application of content at higher grade levels, with the next-largest percentage measuring prerequisite skills at lower grade levels.

After the content review teams had concluded their work, three additional alignment constructs were computed based on the relationship of extended indicators to general indicators and the number of assessed indicators within each content area standard: categorical concurrence, range of knowledge, and balance of representation. Categorical concurrence relates the general and extended standards to determine the content coverage of the extended indicators. For the Kansas Extended Standards, with their comprehensive linking and cross-linking with the general standards, that relationship was complete. Range-of-knowledge measures the extent to which the KAA covers the content in the general standards, and this was less than the recommended 50% coverage for all assessed content areas. Balance of representation looks at the balance of coverage of the extended standards by the assessed indicators, and again this coverage was weak in LAL terms. Reading fell short of recommended balance, Math was weakly balanced, while Science met LAL guidelines for acceptable balance.

Curriculum Indicators Survey

Criterion 8 in the LAL (Flowers et al., 2007, pp. 40–42) recommends gathering evidence that the general curriculum can be accessed in the instructional program of those students who take an alternate assessment. As part of the evidence gathered a sample of Kansas special education teachers who currently taught a student participating in the KAA completed the online Curriculum Indicators Survey (CIS). The CIS is intended to assess the enacted curriculum for students with significant cognitive disabilities who participate in alternate assessments as well as information with regard to instructional resources and professional development (Karvonen, Wakeman, Flowers, & Browder, 2007).

The first part of the survey generally assesses teacher background information and evidence of best practices in instructional programming. Specifically, information is collected about teachers' credentials, professional development, classroom characteristics, instructional resources, use of particular types of classroom assessment, and instructional influences (Karvonen, Wakeman, Flowers, & Browder, 2007). Teachers are instructed to answer these survey items with all of their students in mind, which could be based on students from a self-contained classroom or from a case load that consists of students in multiple settings (Karvonen, Wakeman, Flowers, & Browder, 2007).

Conversely, the second part of the survey is completed with a particular target student in mind. The contents of Part II consist of separate sections for each content area (mathematics, ELA and science) and each of these sections contains a list of topics or strands (e.g., geometry) and specific concepts within each topic. As described by Karvonen, Wakeman, Flowers, & Browder (2007), teachers rate the intensity of coverage of each concept within topics that they teach to target students and also indicate the highest performance expectation (cognitive level) for that student during the year. Within each section of Part II, teachers are also asked to indicate the grade level from which teachers adapted materials, activities, and contexts, the intensity of use of a variety of instructional strategies (e.g., individualized instruction, independent practice), and the level of expectation for student participation in these activities (Karvonen, Wakeman, Flowers, & Browder, 2007).

Participants. A random sample of 400 educators who currently had CETE KAA accounts were sent invitations to participate in the survey. Of those 400 invitations, 84 sent a response indicating that they were not eligible to participate because they did not currently teach a KAA-eligible student, and an additional 30 were returned due to invalid email accounts. Of the remaining 286 surveys that were sent, 84 of them were completed, for a response rate of 29%.

Of the 84 teachers who responded to the survey, 76 (90%) were female. In terms of grade level(s) taught, 32 (38%) indicated that most of their students were in kindergarten through 2nd grade, 39 (46%) mostly taught 3rd through 5th graders, 32 (38%) mostly taught 6th through 8th graders, and 21 (25%) mostly taught 9th through 12th graders. Teachers also reported on the number of students in their classroom or caseload: 13 (15%) indicated that they had between three and five students, 20 (24%) had between six and eight students, 15 (18%) had between nine and 11 students, 20 (24%) had between 12 and 15 students, and 16 teachers (19%) reported that they had more than 15 students in their classroom or caseload.

The amount of education and type of teaching certifications varied across the teachers surveyed. While all teachers had at least a bachelor's degree, almost 70% of the teachers also had a master's degree, and one person had a specialist (6-year) degree. In terms of teacher licensure, 14% of the teachers had a concentration on English language/reading, 6% had a concentration on math, and approximately 1% concentrated on science. However, 92% of teachers had a special education certification, 77% were certified in elementary education, 43% in middle school education, 29% in secondary, 2% were certified by the National Board, and 21% of teachers indicated that they had another type of certification not listed (e.g., Early Childhood Education and ESL). The breakdown of teacher education, license concentration, and certification can be found in Table 23 .

Table 23

Number of Teachers in Categories of Education, License Concentration & Certification

Credentials	N
Degree	
Bachelor's degree only	25
Masters	58
Specialist (6-year)	1
License Concentration	
ELA/Reading	12
Math	5
Science	1
Certification	
Special Education	77
Elementary	65
Middle	36
Secondary	24
National Board	2
Other	18

The amount and type of teaching experience also varied across the survey respondents. Of the 84 teachers, nine reported that they had zero to three years of teaching experience, 19 had between four and 10 years of experience, 31 had between 11 and 20 years of experience, 18 had between 21 and 30 years of experience, and 7 had 31 years or more of teaching experience. The amount of teaching experience by subject area (i.e., reading, math, science, and special education) can be found in Table 24 .

Table 24

Number of Teachers Endorsing Each Category of Teaching Experience by Subject

Years of Teaching Experience	ELA/ Reading	Math	Science	Special Education	Total
0–3	12	13	22	18	9
4–10	20	21	29	24	19
11–20	30	31	19	25	31
21–30	16	13	10	13	18
31 or more	6	6	4	4	7
Total	84	84	84	84	84

Knowledge of Kansas Extended Standards. Teachers were asked how many of the Kansas Extended Indicators are worded clearly so that the instructional purpose is easy to understand. Responses used a five-point scale, with 1 = *almost none* (n = 1), 2 = *about one quarter* (n = 7), 3 = *about one half* (n = 13), 4 = *about three quarters* (n = 22), and 5 = *almost all* (n = 40).

Professional Development. Teachers indicated how much time in the last 12 months they had spent engaged in professional development activities geared towards instructional strategies and content standards specific to each content area (e.g., instructional strategies in ELA and related content standards). Professional development activities could include workshops, inservices, college courses, and summer institutes. There appears to be a consistent pattern for the most participation in professional development activities in the ELA/reading content area and the least participation in the science content area. Teachers reported that they most commonly spent one to five hours in ELA- and reading-related professional development activities. Most teachers indicated that they spent no time in science-related professional development activities. The percentages of teachers' levels of engagement in professional development activities across the three content areas are displayed in Table 25 .

Table 25

Percentage of Teachers Participating in Professional Development Activities by Amount of Time

	ELA/Reading		Math		Science	
Time Engaged in Activity	Inst. Strat.	Cont. Stand.	Inst. Strat.	Cont. Stand.	Inst. Strat.	Cont. Stand.
None	11.9	23.8	21.4	29.8	69.0	66.7
1 to 5 hours	39.3	47.6	54.8	52.4	20.2	22.6
6 to 10 hours	28.6	14.3	13.1	7.1	7.1	6.0
11 to 15 hours	6.0	4.8	4.8	6.0	1.2	2.4
> 15 hours	14.3	9.5	6.0	4.8	2.4	2.4

Instructional Resources. Teachers were asked to indicate the types of instructional resources that they used throughout the school year to teach reading, math, and science to their students. These resources were categorized into types of materials, settings, and persons. Tables 26–28 describe the percentages of teachers who used various types of resources (materials, settings, and persons) to teach the three content areas. With respect to the types of materials used during instruction, teachers reported that they used materials or lessons from websites the most for all three content areas. In math, teacher-made manipulatives were also reported as being used as much as website resources. Across all content areas, teachers indicated that they used assistive technologies (e.g., Cheap Talk, Big Mac, DynaVox, text reader, talking calculator) the least. Patterns across the types of settings teachers used during instruction were less apparent. Teachers indicated that they used other settings in the community the least for all three content areas. For ELA/reading, teachers reported using other settings in the school the most. Teachers reported using natural setting materials (e.g., coins) the most in math and other settings in the school the most in science. However, both natural setting materials and inclusive class settings were used almost as much as other settings in science. With regard to types of persons used to assist with instruction, teachers reported that they used nondisabled peers the least. In reading teachers primarily used another staff member such as a speech therapist, for math they used other special education teachers, and in science teachers reported using teachers from other disciplines to assist with instruction.

Table 26

Percentage of Teachers Who Used Each Type of Material by Content Area

Type of Material	Reading	Math	Science
Commercially made materials adapted from general education	76.2	78.6	39.3
Commercially made manipulatives adapted from general education	63.1	78.6	38.1
Age-appropriate, commercially made print or text materials designed for this type of student	69.0	63.1	34.5
Age-appropriate, commercially made manipulatives designed for this type of student	57.1	67.9	31.0
Other commercially made print or text materials designed for this type of student	65.5	59.5	35.7
Other commercially made age-appropriate manipulatives designed for this type of student	50.0	59.5	27.4
Teacher-made books, workbooks, materials	82.1	81.0	58.3
Teacher-made manipulatives	79.8	84.5	45.2
Materials or lessons from websites	88.1	84.5	60.7
Computer	83.3	71.4	50.0

Type of Material	Reading	Math	Science
Assistive technologies (e.g., CheapTalk, Big Mac, DynaVox, text reader, talking calculator, etc.)	44.0	36.9	23.8

Table 27

Percentage of Teachers Who Used Each Type of Setting by Content Area

Type of Setting	Reading	Math	Science
Real life or natural setting materials (e.g., coins, community signs, telephones)	69.0	79.8	45.2
Inclusive class setting	53.6	51.2	45.2
Other settings in my school	73.8	67.9	47.6
Other settings in the community	36.9	42.9	20.2

Table 28

Percentage of Teachers Who Used Each Type of Person by Content Area

Type of Persons	Reading	Math	Science
Nondisabled peers	31.0	27.4	15.5
Teachers from other disciplines (e.g., academic or special subject areas)	35.7	34.5	31.0
Another staff member at the school (e.g., speech/occupational/physical therapist)	66.7	39.3	28.6
Other special education teachers	48.8	45.2	26.2

Classroom Assessment. Teachers were asked about how often they used various types of assessment in the three different content areas, including objective questions, on-demand performance, and teacher observation. Figures 10–12 display the percentages of teachers who used each type of assessment by level of frequency for each content area. Teachers reported using teacher observation the most frequently and objective questions the least frequently to assess their students in reading and math. This pattern also existed in science; however, many teachers also reported using performance on demand quite frequently to assess their students in science.

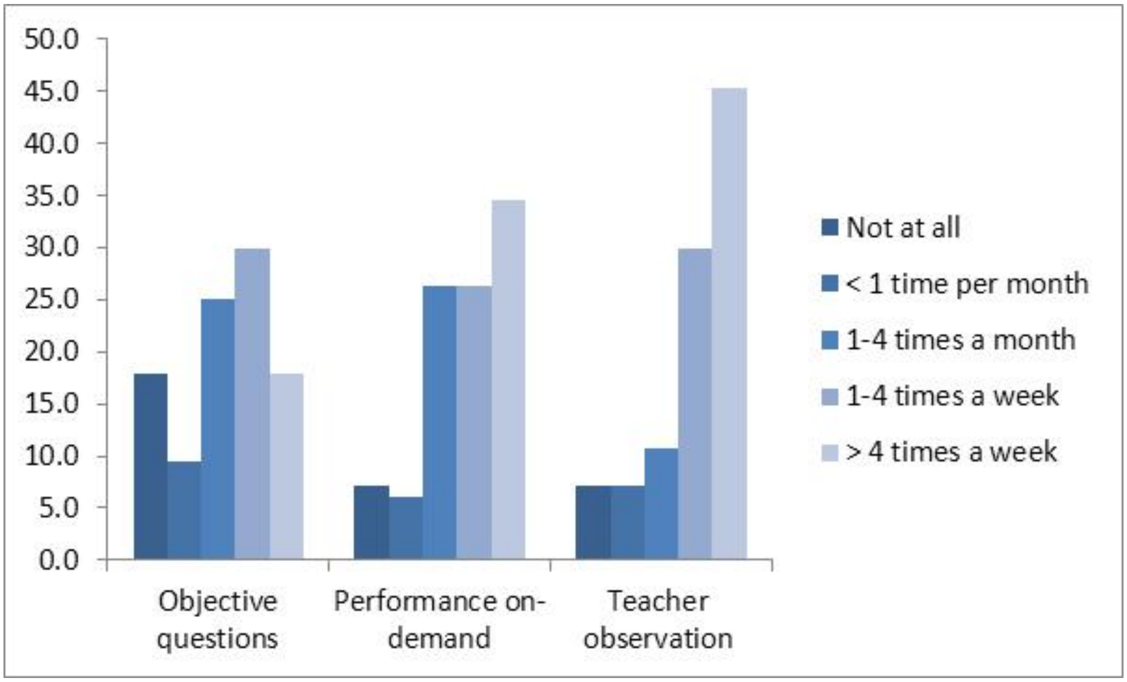

Figure 10. Percentage of Assessment Use Frequency in Reading.

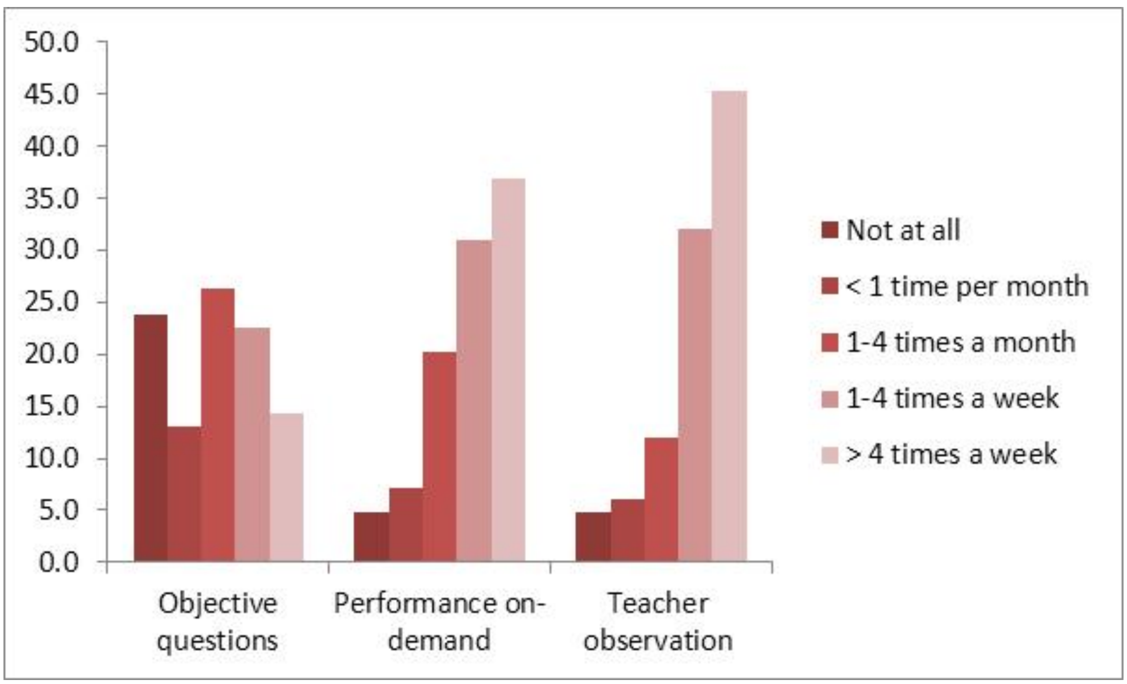

Figure 11. Percentage of Assessment Use Frequency in Math.

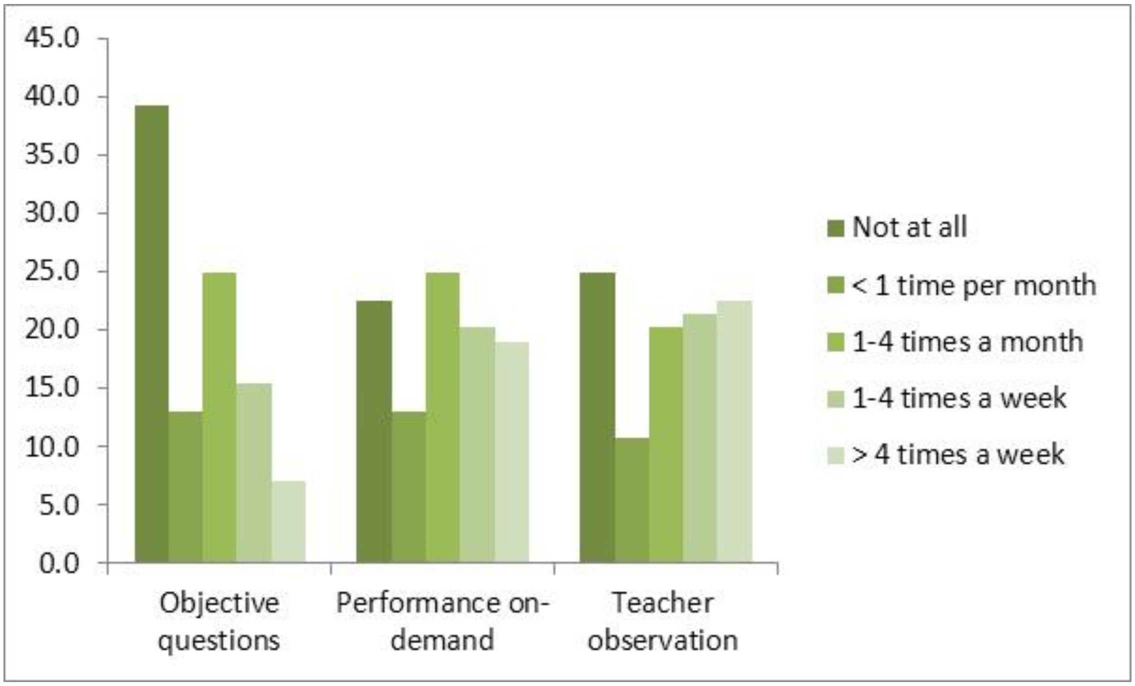

Figure 12. Percentage of Assessment Use Frequency in Science.

Instructional Influences. Teachers were also asked about the degree to which various types of factors influence their instruction in each of the three content areas. As is shown in Figures 13–15, which display the degree of and type of instructional influence for each content area, students' needs as documented by their IEPs were, by far, the most influential factor on teachers' instruction in all three content areas. National content standards appeared to be the least influential factor on instruction in all content areas.

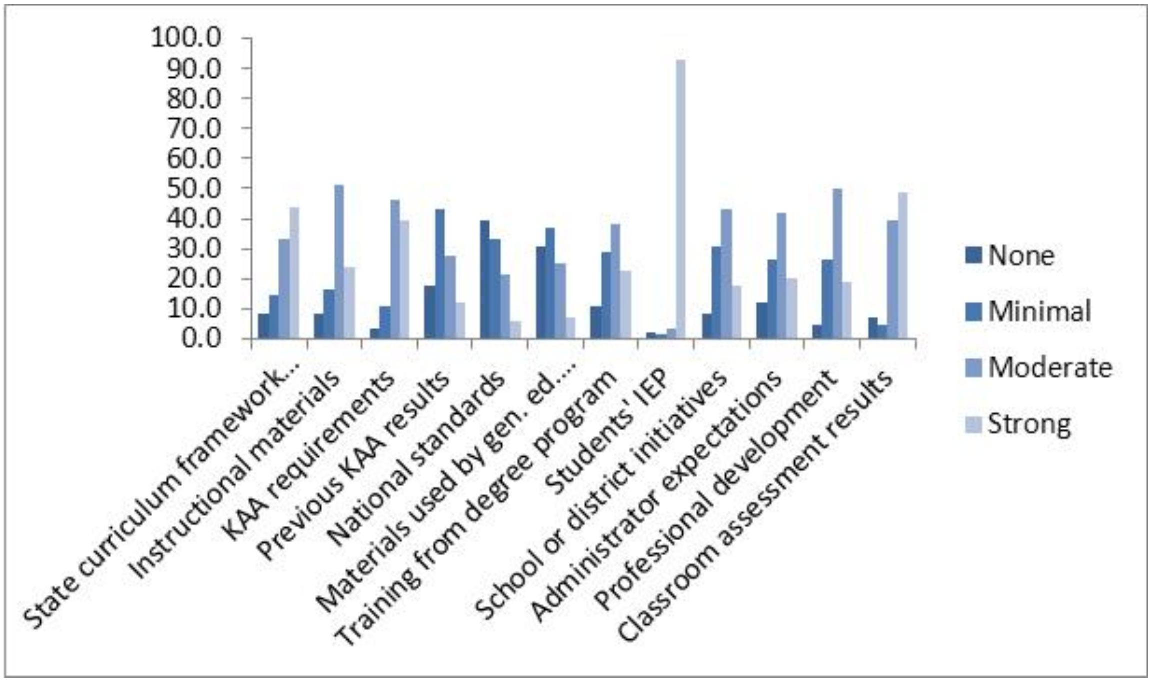

Figure 13. Percentage of Teachers Who Endorsed Each Type of Instructional Influence in Reading.

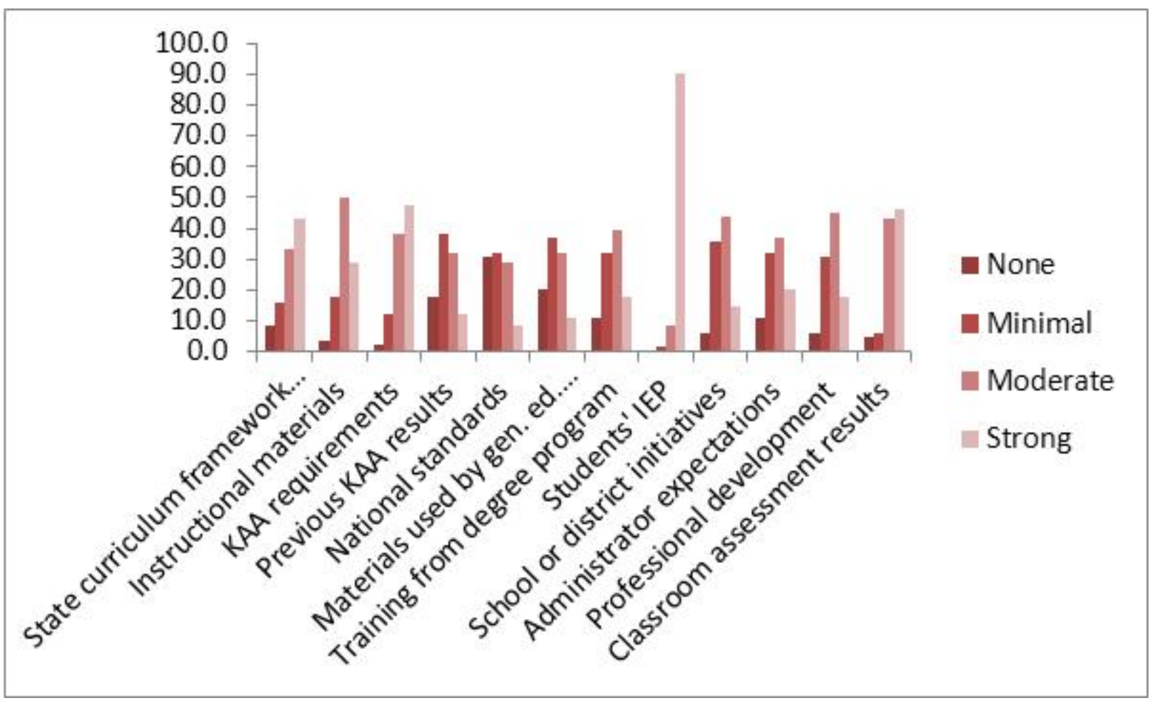

Figure 14. Percentage of Teachers Who Endorsed Each Type of Instructional Influence in Math.

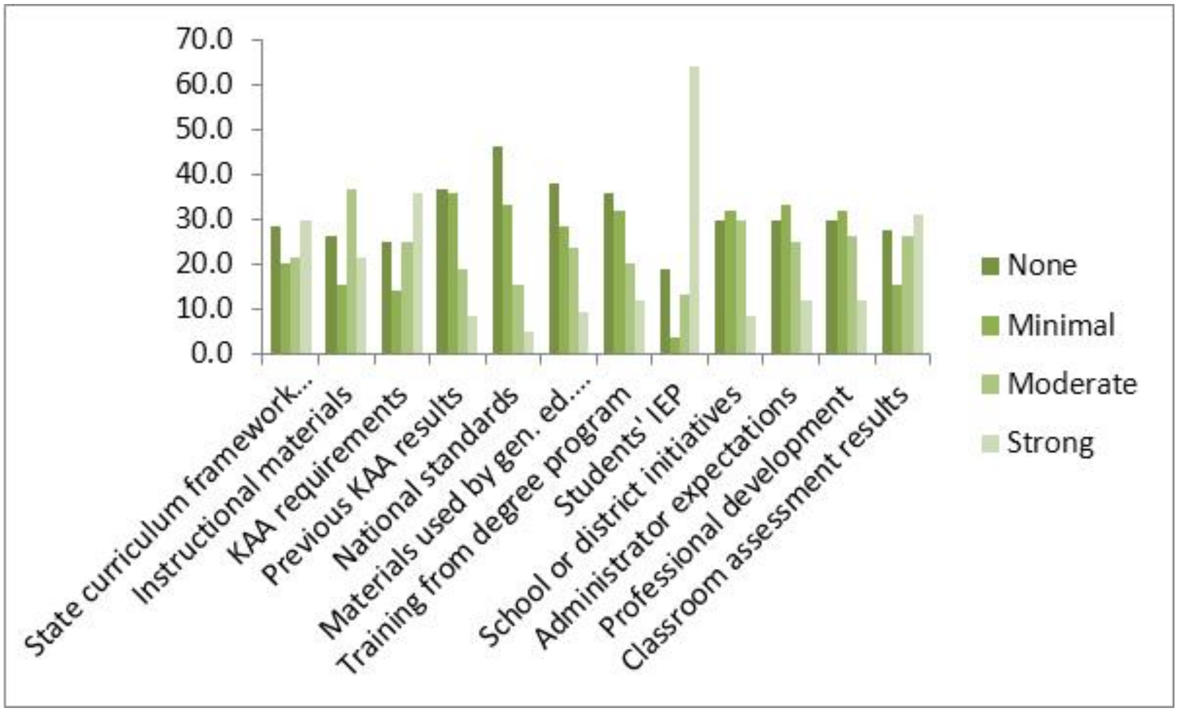

Figure 15. Percentage of Teachers Who Endorsed Each Type of Instructional Influence in Science.

Part II: Target Students

Target Student Characteristics. For the second part of the survey, teachers were asked to identify and describe a target student as a reference for the subsequent survey questions. Of the 84 teacher respondents, 49% indicated that their target students were in the 3rd through 5th grade, 32% were in 6th through 8th grade, 19% were in 9th through 12th grade, and 2% of the teachers reported that their target students were in a grade not listed. Target student grade level and age can be found in Tables 29 and 30. Teachers also specified one or more disability labels (as defined by IDEA) that applied to their target students; these can be found in Table 31.

Target students were further described by their current levels of communication, which were partitioned into three levels according to the definitions provided in Table 5. A level 1 student does not yet have the skills to discriminate between pictures or other symbols and does not use symbols to communicate; a level 2 student may use some symbols to communicate (e.g., pictures, logos, objects) and is beginning to acquire symbols as part of a communication system; and a level 3 student communicates with symbols (e.g., pictures) or words (e.g., spoken words, assistive technology, ASL, home signs). Of the 84 teachers surveyed, 6% described their target students' communication abilities as being at level 1, 20% indicated level 2, and 74% reported that their student was at level 3.

Table 29

Target Student Grade Level

Grade Level	N
3–5	39
6–8	27
9–11	16
Other	2
Total	84

Table 30

Target Student Age

Age Range	N
8–10	29
11–13	29
14–16	19
17–19	7
Total	84

Table 31

Target Student Disability Categories

IDEA Disability Label	N	%
Mental Retardation	49	58.3
Speech/Language Impairment	28	33.3
Autism	24	28.6
Multiple Disabilities	15	17.9
Other Health Impaired	11	13.1
Learning Disability	10	11.9
Orthopedic Impairment	5	6.0
Emotional Disturbance	5	6.0
Traumatic Brain Injury	1	1.2
Visual Impairment	1	1.2
Hearing Impaired/Blind	0	0.0
Deaf–Blind	0	0.0

Levels of symbolic communication for KAA participants on the CIS can be compared with the corresponding classifications on the Learner Characteristics Inventory (LCI), which was implemented in Kansas in 2007 by the National Alternate Assessment Center (Kearns & Towles-Reeves, 2007). Table 32 shows the comparison between the LCI descriptions of students in spring 2007 (Kearns & Towles-Reeves) and the CIS data collected in spring 2010. Also included are LCI data from three additional states for the purposes of comparison across states (Towles-Reeves et al., 2009).

It is apparent that the population of students taking the KAA is consistent from 2007 to 2010 in terms of classification into the categories of the CIS and the expressive portion of the LCI, which are reasonably consistent in their definitions (see Tables 5 and 6 for complete categorical descriptions). However, it is equally apparent that the categories of expressive and receptive communication on the LCI define students differently. Inclusion or classification into a particular expressive language category does not imply classification into a corresponding receptive language category. Of course, classification depends almost entirely on the clarity and mutual exclusivity of the category definitions, so comparison of two definitional systems requires close affiliation between the corresponding categories. Both the LCI and the CIS refer to concrete symbols or icons such as pictures and objects as intentional but not fully symbolic communication. This is in contrast to the LAL definitions for Criterion 1 (Table 4) and to other literature that associates iconic graphics with true symbol use rather than emerging or presymbolic communication (Poncelas & Murphy, 2007; Sutton et al., 2009; Tomasello et al., 1999).

Table 32

Comparison of Percent of Students Classified in LCI and CIS Communication Categories

Levels of Symbolic Communication	Kansas	State 1	State 2	State 3
Learner Characteristics Inventory (2007)				
Expressive				
Symbolic	74	71	63	74
Intentional	18	17	26	17
nonintentional	8	8	11	8
Receptive				
Symbolic	48	46	34	56
Cued	43	41	54	33
Alert	8	10	10	7
Uncertain	1	2	2	3
Curriculum Indicators Survey (2010)				
3 Symbolic	74			
2 Beginning symbolic	20			
1 Nonsymbolic	6			

Intensity of Coverage. After identifying a target student, teachers were asked to indicate the amount of instructional coverage they provided the student on each content-specific item or activity since the beginning of the school year. Instruction could include direct instruction through independent practice delivered by any staff member, peer, or volunteer. As shown in Table 33, intensity of coverage was measured by the number of lessons in a school year used to teach the concept. In Reading, teachers were asked about 27 reading concepts within four different topics or strands. In Math, there were 16 concepts within five different math topics, and in Science there were six topics that consisted of 21 concepts. As can be seen in Figure 16, across all three content areas teachers most frequently rated the various content-specific concepts as having no coverage (i.e., not an expectation for the particular topic during the current school year). In Reading, 73.5% of teachers reported that they do not cover *evaluating writing and presentations* at all during the school year. Approximately 70% of teachers indicated that they do not teach *formulas of measurement* in math, and 91.4% do not cover *atomic theory* in science. On the other hand, almost 60% of teachers indicated that they taught *beginning reading* at an intensive level and approximately the same percentage of teachers indicated that they taught *number sense* at an intensive level. In Science, the highest percentage of teachers who reported teaching a concept at this level was 25% for the *structure and energy in the earth's system* concept. Percentages of teacher ratings for the intensity of coverage for the target student by each content-specific concept are displayed in Appendix F.

Table 33

Codes for Assessing the Intensity of Coverage of Topic Concepts

Intensity of Coverage	Definition
None	not an expectation for this topic this school year
Slight	1 - 10 lessons over the course of the school year
Moderate	11 - 20 lessons over the course of the school year
Sustained	21 or more lessons over the course of the school year
Intensive	daily/nearly daily instruction throughout the school year
None yet	no coverage yet but planned for later this school year

Adapted from Flowers et al. (2007)

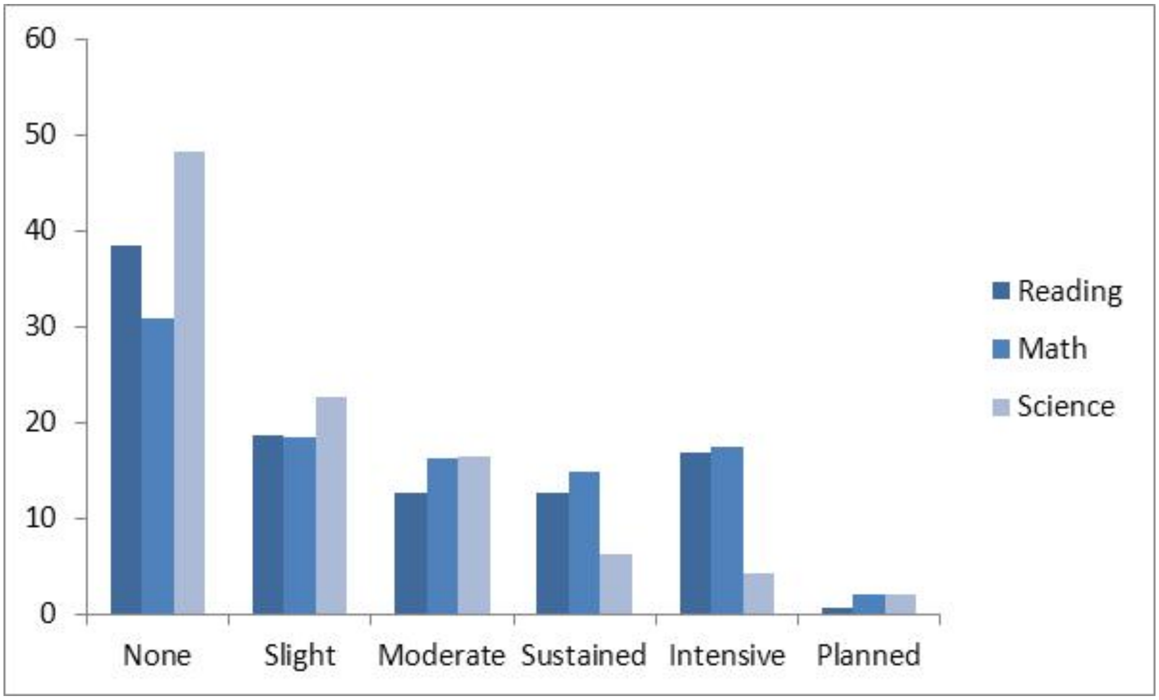

Figure 16. Percentage of Ratings for Intensity of Coverage for Content-Specific Activities by Content Area.

Highest Performance Expectation. Teachers were also instructed to indicate the highest expectation for performance they had for their target students on each of the concepts within the content strands. The cognitive levels of performance (depth of knowledge) are the same as those used in the content expert review section of this report and can be found in Table 19. Teachers most frequently indicated that their highest level of performance expectation for their target students on reading and science concepts was at the most basic level, *attention*. In math, the *performance* level of cognition was reported slightly more often than the *attention* level. The percentages of ratings for the highest performance expectation on the various content-specific concepts by content area are displayed in Figure 17. Percentages of teacher ratings for the highest performance expectation for target students by each content-specific concept are displayed in Appendix G.

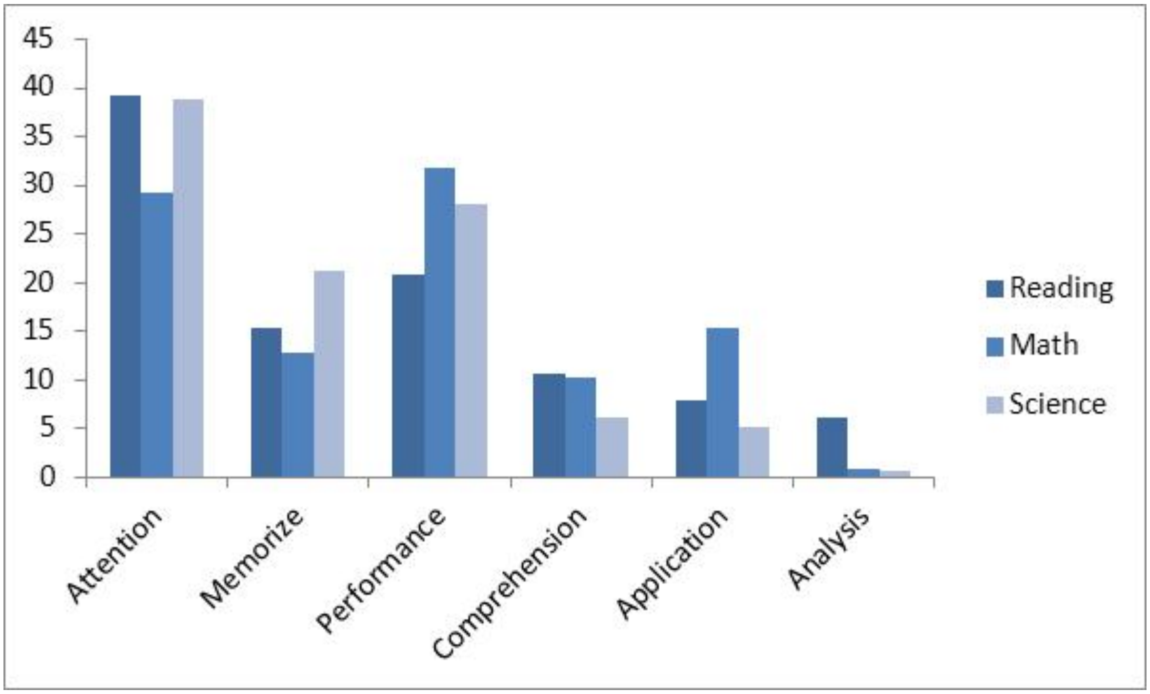

Figure 17. Percentages of Ratings for Highest Performance Expectation Across Content-Specific Concepts by Content Area.

Grade-Level Materials. After evaluating the various concepts within each content strand with regard to the intensity of coverage and cognitive expectation, teachers were instructed to indicate the grade level from which they adapted materials, activities, and contexts for instruction with target students. In general, teachers mostly reported adapting materials from the lower grade levels (pre-Kindergarten through 2nd grade) in order to teach their target students in all three content areas. Tables 34–36 display the percentages of teachers who adapted materials from the various grade-level bands by each content topic.

Table 34

Percentage of Teachers Who Adapted Reading Materials, Activities, or Contexts

Topic	Not taught	pK–2	3–5	6–8	9–12	Voca-tional	Un-graded
Language	14.3	57.1	22.6	4.8	0.0	10.7	15.5
Reading and Literature	8.3	59.5	17.9	7.1	0.0	9.5	10.7
Composition	32.1	47.6	9.5	2.4	0.0	7.1	10.7
Media	75.0	3.6	6.0	3.6	0.0	1.2	14.3

Table 35

Percentage of Teachers Who Adapted Math Materials, Activities, or Contexts

Topic	Not taught	pK–2	3–5	6–8	9–12	Voca-tional	Un-graded
Number Sense Operations	7.1	60.7	21.4	4.8	1.2	13.1	17.9
Patterns, Relations, Algebra	14.3	64.3	10.7	4.8	1.2	9.5	13.1
Geometry	28.6	52.4	10.7	4.8	0.0	7.1	13.1
Measurement	16.7	56.0	15.5	3.6	1.2	10.7	11.9
Data Analysis, Statistics/Probability	34.5	39.3	11.9	2.4	0.0	9.5	11.9

Table 36

Percentage of Teachers Who Adapted Science Materials, Activities, or Contexts

Topic	Not taught	pK–2	3–5	6–8	9–12	Voca-tional	Un-graded
Earth & Space Science	42.9	38.1	4.8	4.8	1.2	6.0	11.9
Life Science Biology	36.9	42.9	6.0	6.0	1.2	4.8	13.1
Physical Science, Chemistry, & Physics	50.0	29.8	4.8	4.8	0.0	4.8	13.1
Technology/Engineering	61.9	15.5	4.8	4.8	1.2	4.8	10.7
History & Nature of Science	78.6	4.8	2.4	2.4	0.0	1.2	9.5
Science as Inquiry	67.9	15.5	4.8	3.6	0.0	0.0	10.7

Instructional Time. The last portion of the CIS asked teachers about the intensity of use of a variety of instructional strategies (e.g., individualized instruction, independent practice) as well as the level of expectation for student participation in these activities. First, teachers rated the amount of time spent with target students on each content-specific instructional activity during the past week. Teachers indicated "little" if one hour or less was spent on the activity during the past week; "some" for two to four hours; "moderate" for five to seven hours; and "considerable" if eight or more hours were spent on the activity during the past week.

With respect to the most extreme categories (i.e., "none" and "considerable"), the most commonly reported reading and science instructional activity that received no instructional time was *engagement in speech or presentation* (reading: 64.3%; science: 79.8%), while the most commonly reported activity in both of these content areas that received a considerable amount of instructional time was *individualized instruction* (reading: 36.9%; science: 8.3%). For math, *using work centers* (54.8%) was reported the most as an activity that received no instructional time while *individualized instruction* (29.8%) was rated the most as an activity that received a considerable amount of instructional time. Tables 37–39 describe the percentages of teachers who endorsed each instructional activity based on the amount of instructional time that was spent on that activity with target students during the previous week. Activity descriptions are condensed from the original survey descriptions.

Table 37

Percentages of Teacher Ratings for Instructional Time in Reading Activities

	Amount of Instruction Time (hrs)				
Instructional Activity	None	≤1	2 to 4	5 to 7	≥8
Receive individualized instruction	1.2	4.8	31.0	25.0	36.9
Receive instruction in small group	6.0	9.5	29.8	32.1	21.4
Collect, summarize or analyze information	40.5	29.8	13.1	13.1	1.2
Work independently	16.7	27.4	35.7	14.3	4.8
Receive instruction with prompts	7.1	8.3	28.6	27.4	26.2
Learn demonstrate skills	7.1	8.3	22.6	38.1	21.4
Engage in writing process	26.2	28.6	29.8	11.9	1.2
Learn to use resources	41.7	28.6	14.3	7.1	2.4
Use hands on materials or manipulatives	3.6	14.3	27.4	29.8	23.8
Use computers or other assistive technology	10.7	22.6	31.0	27.4	7.1
Perform assessment skills for data collection	15.5	27.4	28.6	16.7	9.5
Take a test	36.9	45.2	10.7	2.4	2.4
Practice skills in different setting	15.5	33.3	32.1	14.3	3.6
Practice skills with a variety of similar materials	11.9	28.6	35.7	17.9	4.8
Engage in read-aloud activities	17.9	21.4	32.1	21.4	6.0
View multi-media presentations	38.1	39.3	16.7	1.2	0.0
Engage in speech or presentation	64.3	20.2	7.1	6.0	0.0

Instructional Activity	Amount of Instruction Time (hrs)				
	None	≤1	2 to 4	5 to 7	≥8
Use work center	52.4	20.2	14.3	8.3	2.4

Table 38

Percentages of Teacher Ratings for Instructional Time in Math Activities

Instructional Activity	Amount of Instruction Time (hrs)				
	None	≤1	2 to 4	5 to 7	≥8
Receive individualized instruction	2.4	11.9	23.8	31.0	29.8
Receive instruction in small group	14.3	11.9	20.2	35.7	16.7
Collect, summarize or analyze information	28.6	31.0	28.6	6.0	2.4
Work independently	14.3	28.6	32.1	20.2	0.0
Receive instruction with prompts	6.0	9.5	35.7	29.8	16.7
Learn demonstrate skills	7.1	11.9	36.9	23.8	15.5
Rote Count	17.9	32.1	21.4	21.4	4.8
Complete symbolic math problems	28.6	21.4	32.1	8.3	6.0
Learn to use math resources	36.9	34.5	16.7	2.4	6.0
Use hands-on math materials	6.0	10.7	39.3	22.6	15.5
Use computers, calculators, or other tech.	13.1	23.8	27.4	22.6	9.5
Perform math assessment skills	28.6	38.1	11.9	11.9	6.0
Take a test	46.4	38.1	6.0	4.8	2.4
Practice skills with different setting	28.6	31.0	21.4	9.5	3.6
Practice skills with a variety of materials	9.5	31.0	34.5	14.3	8.3
Apply mathematical concepts to real world	19.0	34.5	25.0	11.9	4.8
Use work center	54.8	13.1	13.1	10.7	2.4

Table 39

Percentages of Teacher Ratings for Instructional Time in Science Activities

Instructional Activity	Amount of Instruction Time (hrs)				
	None	≤1	2 to 4	5 to 7	≥8
Receive individualized instruction	34.5	21.4	23.8	7.1	8.3
Receive instruction in small group	28.6	29.8	22.6	9.5	4.8
Collect, summarize or analyze information	47.6	34.5	10.7	1.2	1.2
Work independently	48.8	27.4	11.9	3.6	1.2
Receive instruction with prompts	28.6	23.8	28.6	7.1	4.8
Learn demonstrate skills	36.9	22.6	27.4	4.8	3.6
Engage in inquiry process	64.3	21.4	3.6	3.6	2.4
Learn to use science resources	63.1	21.4	6.0	1.2	2.4
Use hands-on science materials	27.4	31.0	22.6	7.1	6.0
Use computers, calculators, or other tech.	41.7	31.0	10.7	6.0	4.8
Perform science assessment skills	58.3	27.4	4.8	2.4	2.4
Take a test	69.0	20.2	3.6	1.2	0.0
Practice skills with different setting	52.4	25.0	11.9	3.6	1.2
Practice skills with a variety of materials	46.4	29.8	14.3	1.2	2.4
Engage in science read-aloud activities	46.4	34.5	7.1	3.6	1.2
View multi-media presentations	53.6	26.2	9.5	6.0	0.0
Engage in speech or presentation	79.8	11.9	3.6	0.0	0.0
Use work center	63.1	13.1	8.3	3.6	1.2

Level of Participation. Finally, for each content-specific instructional activity that target students spent some amount of time on (i.e., at least one hour), teachers also rated the level of the student's participation in the activity. Teachers rated the participation level as *none*, *passive*, *active with supports*, or *independent active*. Across all three content areas, the level of student participation was characterized as *active with supports* most frequently for the majority of instructional activities. Tables 40–42 display rates of teacher endorsement for each instructional activity based on the target student's level of participation in that activity. Activity descriptions are condensed from the original survey descriptions.

Table 40

Percentages of Teacher Ratings for Participation in Reading Instructional Activities

	Level of Student Participation			
Instructional Activity	None	Passive	Active With Supports	Independent Active
Receive individualized instruction	0.0	10.7	72.6	8.3
Receive instruction in small group	3.6	15.5	61.9	7.1
Collect, summarize or analyze information	22.6	19.0	23.8	2.4
Work independently	10.7	17.9	28.6	27.4
Receive instruction with prompts	3.6	13.1	65.5	6.0
Learn demonstrate skills	2.4	15.5	61.9	7.1
Engage in writing process	10.7	17.9	45.2	4.8
Learn to use resources	20.2	21.4	22.6	3.6
Use hands on materials or manipulatives	0.0	13.1	58.3	17.9
Use computers or other assistive technology	3.6	11.9	45.2	25.0
Perform assessment skills for data collection	8.3	14.3	50.0	13.1
Take a test	15.5	11.9	40.5	9.5
Practice skills in different setting	7.1	19.0	50.0	7.1
Practice skills with a variety of materials	3.6	17.9	59.5	4.8
Engage in read-aloud activities	8.3	17.9	54.8	4.8
View multi-media presentations	16.7	34.5	16.7	6.0
Engage in speech or presentation	29.8	10.7	19.0	2.4
Use work center	21.4	9.5	19.0	13.1

Table 41

Percentages of Teacher Ratings for Participation in Math Instructional Activities

	Level of Student Participation			
Instructional Activity	None	Passive	Active With Supports	Independent Active
Receive individualized instruction	0.0	13.1	69.0	8.3
Receive instruction in small group	6.0	13.1	60.7	4.8
Collect, summarize or analyze information	11.9	19.0	41.7	1.2
Work independently	8.3	15.5	39.3	23.8
Receive instruction with prompts	2.4	13.1	69.0	2.4
Learn demonstrate skills	2.4	13.1	66.7	3.6
Rote Count	4.8	10.7	42.9	20.2
Complete symbolic math problems	10.7	9.5	45.2	8.3
Learn to use math resources	16.7	19.0	31.0	2.4
Use hands-on math materials	0.0	10.7	61.9	13.1
Use computers, calculators, or other tech.	2.4	15.5	42.9	20.2
Perform math assessment skills	14.3	10.7	42.9	9.5
Take a test	14.3	9.5	32.1	8.3
Practice skills with different setting	11.9	19.0	39.3	2.4
Practice skills with a variety of materials	2.4	16.7	59.5	4.8
Apply mathematical concepts to real world	6.0	16.7	52.4	4.8
Use work center	21.4	9.5	21.4	7.1

Table 42

Percentages of Teacher Ratings for Participation in Science Instructional Activities

	Level of Student Participation			
Instructional Activity	None	Passive	Active With Supports	Independent Active
Receive individualized instruction	6.0	17.9	33.3	3.6
Receive instruction in small group	4.8	20.2	38.1	3.6
Collect, summarize or analyze information	14.3	14.3	27.4	2.4
Work independently	11.9	19.0	17.9	8.3
Receive instruction with prompts	3.6	20.2	39.3	4.8
Learn demonstrate skills	7.1	16.7	31.0	3.6
Engage in inquiry process	14.3	11.9	15.5	2.4
Learn to use science resources	14.3	13.1	13.1	2.4
Use hands-on science materials	2.4	15.5	44.0	1.2
Use computers, calculators, or other tech.	11.9	16.7	26.2	4.8
Perform science assessment skills	16.7	6.0	23.8	3.6
Take a test	22.6	4.8	16.7	2.4
Practice skills with different setting	15.5	11.9	20.2	2.4
Practice skills with a variety of materials	9.5	13.1	27.4	3.6
Engage in science read-aloud activities	15.5	17.9	20.2	2.4
View multi-media presentations	13.1	17.9	19.0	2.4
Engage in speech or presentation	21.4	8.3	8.3	0.0
Use work center	21.4	7.1	13.1	2.4

Highest Performance Expectation Compared to Depth of Knowledge

As recommended by the LAL (Flowers et al., 2007), comparisons were made between the CIS cognitive level ratings and similar ratings provided in other elements of the system—specifically, the depth-of-knowledge ratings from the content expert review section. The ELA and Math content strands used in the CIS were originally derived from the Alternate Assessment Collaborative and frameworks developed by Massachusetts in 2001 (Karvonen, Wakeman, Flowers, & Browder, 2007). Teachers were instructed to rate these content strands in Reading, Math, and Science with respect to the highest level of cognitive performance they would expect from their identified target students. Conversely, the content expert reviewers in the second section of this report were instructed to rate each Kansas extended indicator with respect to the depth of knowledge that is required by that indicator. While there is not a straightforward correspondence between Kansas extended indicators and the content strands from the CIS, comparisons between general content areas were feasible. In Reading, the CIS content strands contained two additional topics (composition and media) that are not addressed in any of the Kansas general or extended standards; for this reason, these ratings from the CIS were removed from this analysis. All other ratings were summed across raters and across content areas in order to compare the percentage of ratings given for each cognitive level between content review experts and CIS respondents.

As shown in the following radar figure, the content review depth-of-knowledge ratings were rated far more frequently at higher levels of cognition (i.e., application and analysis) than the CIS target student performance expectations, which were rated most frequently at the lowest levels of cognition (i.e., attention and memorize). In other words, content reviewers were more likely to rate various extended indicators at higher levels of cognition than were teachers who rated similar content with a target student in mind.

Conclusions from this data should only be made with caution due to the lack of direct alignment between the types of content strands from each section as well as differences between the tasks (e.g., rating depth of knowledge of a particular indicator versus rating the highest performance expectation for a student). However, it may be of interest to note that the pattern described above is consistent across all three content areas. Further exploration is needed to determine the function of these patterns.

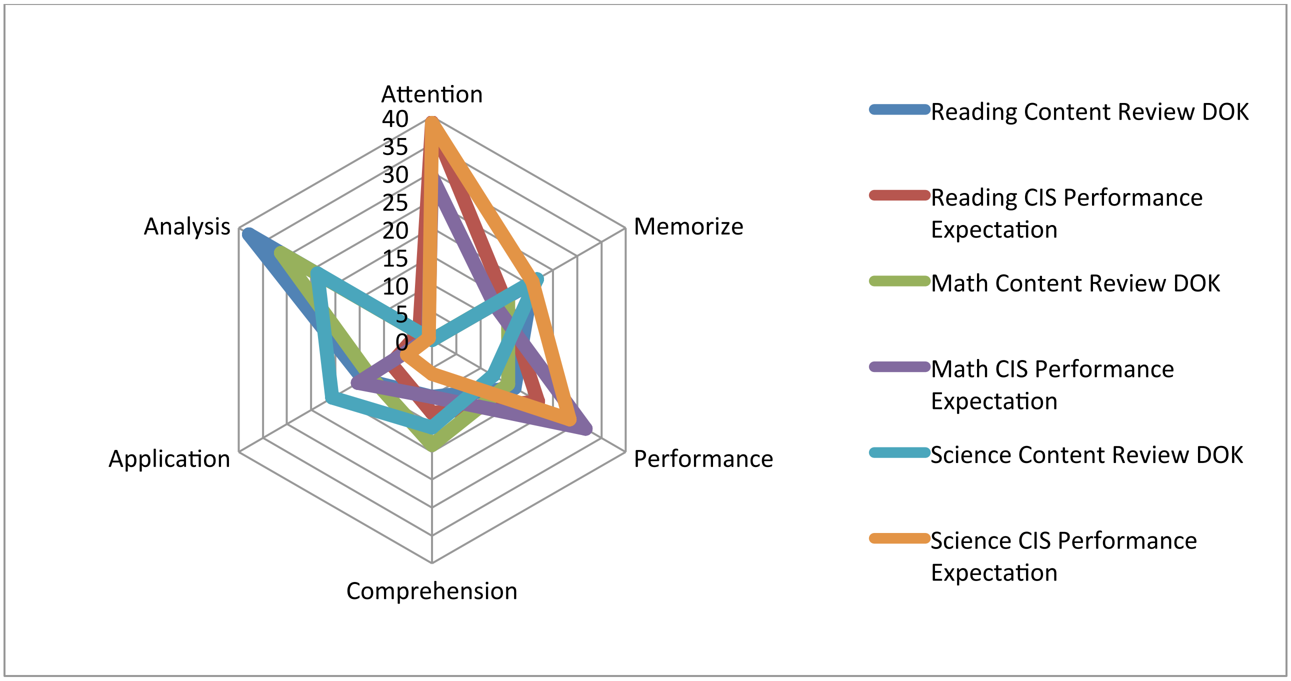

Figure 18. Percentages of Ratings for Depth of Knowledge (DOK) by Type of Rater for Each Content Area.

Summary

The CIS was intended to evaluate Criterion 8 of the LAL manual (Flowers et al., 2007, pp. 40–42), which gives consideration to the alignment of instruction with content standards. Flowers et al. (2007) explain that this is particularly important given that teachers are now required to teach this population state academic curriculum. Teachers in the state of Kansas who currently taught at least one student who participated in the KAA were surveyed about their background information and instructional practices for their classrooms, both as a whole and for individual target students. All 84 teachers who responded to the survey had at least three students in their classroom or workload, and 16 teachers had more than 15 students in their classroom or workload. All teachers had at least a bachelor's degree and almost 70% of those teachers had a master's degree; one person had a specialist degree. Of the 84 teachers, 77 indicated that they had a special education certification and 49 reported that they had between four and 20 years of teaching experience in special education.

Teachers were first asked to indicate the types of instructional resources that they used throughout the school year to teach reading, math, and science to their students. For all three content areas, teachers reported that they used materials or lessons from websites the most and assistive technologies the least. Teachers further indicated that they used settings within the community the least to teach reading, math, and science. Other settings within the school were used the most to teach reading and science, and natural setting materials were used the most to teach math. Finally, teachers reported that they used another staff member (such as a speech therapist) the most to assist with teaching reading, other special education teachers to help teach math, and teachers from other disciplines to help teach science. Across all three content areas, teachers reported using all of the various instructional resources (materials, settings, and persons) the least in science.

Teachers were also asked about their assessment practices in the three different content areas. Teachers reported using observation the most frequently and objective questions the least to assess their students in all three content areas. Performance on demand was also indicated as an assessment that was used frequently to assess students in all three content areas.

Several instructional influences (e.g., state curriculum framework or content, national content standards) were assessed by teacher respondents with regard to the degree of impact each had on their instructional program. Across all three content areas, students' needs as documented by their IEPs were the most influential factor on teachers' instruction while national content standards were the least influential factor on instruction.

The second part of the survey required teachers to identify a target student who the subsequent items would reference. Target student characteristics were found to be consistent with previous classification of KAA students with respect to levels of communication. Almost 60% of target students were classified as intellectually disabled and 74% were categorized at the symbolic level of communication.

Once teachers identified and described their target students, they were asked to indicate the amount of instructional coverage that the student received on specific concepts within each content strand as well as the highest performance expectation (cognitive level) they had for that student on each of those concepts. Across all three content areas, teachers most frequently rated content-specific concepts as not being covered during the current school year. However, in each of Reading and Math, there was one content-specific concept that was rated by the majority of teachers as intensely covered in the instructional program. In Science, only one quarter of teachers rated any one concept as being taught at the intensive level.

Furthermore, teachers most frequently indicated that their highest level of performance expectation for their target students on reading and science concepts was at the most basic level, *attention*. In Math, the *performance* level of cognition was reported slightly more often than the *attention* level. *Analysis* was reported the least as the highest performance expectation for target students across all three content areas.

Finally, teachers were asked about the intensity of use of a variety of instructional strategies (e.g., individualized instruction, independent practice) and the level of student participation expected in these activities. The most commonly reported reading and science instructional activity that received no instructional time was *engagement in speech or presentation,* while *using work centers* was reported the most as an activity that received no instructional time in math. Across all three content areas, *individualized instruction* was reported most frequently as receiving a considerable amount of instructional time. With respect to student participation in these activities, the level of expectation was characterized as *active with supports* most frequently for the majority of the instructional activities across all three content areas.

Recommendations

Specific recommendations have been formulated based on the reviewers' findings and the conclusions of the LAL study. Many strengths of the KAA are reiterated here, as well as corresponding suggestions for improvement. This section is organized by large categories rather than by the LAL criteria or the findings of any particular group of reviewers. There are several aspects of the KAA that could be improved through critical evaluation of the components that were rated lower by reviewers who participated in the Kansas Alternate Assessment Alignment Study. Following are suggestions for consideration by KSDE and its special education advisory panel.

Symbolic Communication. This topic is mentioned separately because of the amount of confusion evident in the LAL and in research into the communication skills of individuals with disabilities. KSDE has made a decision to avoid labeling students by level of symbolic communication attainment, which is commendable given the inconsistency of terminology in the literature as well as KSDE concerns about pigeonholing students into inflexible categories of communicative competence with consequent lowered expectations. Furthermore, federal peer review requirements included removing indicators targeting only communication because they were considered functional rather than academic skills. As a result, therefore, the KAA does not currently include increased communication skill development as an explicit objective. The following suggestions are made with the goal of incorporating expectations for increasing symbol use into extended indicators:

1. Targeting student growth in symbolic communication by including increased communication goals within academic indicators could be considered. Target skills may include the use of gestures, objects, and symbols, in addition to spoken, signed, and written words, to demonstrate both expressive and receptive language competence in the demonstration of academic content. These goals could be incorporated into extended indicators in any content area, even as suggestions or reminders for teachers to work on building symbolic language skills concurrently with academic knowledge.

2. Professional development activities could focus on sequences of communication skills to guide teachers toward increasing symbolic communication demands for their students. Emphasizing sequenced skills might lessen the tendency of teachers to assume that a student's

current level of communicative competence is the highest that can be attained.

3. Explicit inclusion of these skills may be facilitated during the forthcoming transition to Common Core State Standards and development of new types of AA-AAS test items and tasks.

Extended Standards. Many significant strengths of the Kansas Extended Standards were made evident during the alignment review process. One of the positive outcomes is the assurance that all of the content referenced in the extended standards is indeed academic content, as confirmed by the content review panel, with the exception of a single extended indicator in Mathematics. The depth of knowledge of the extended indicators was generally rated at the higher levels of a modified Bloom's taxonomy, which ensures that student performance on the extended indicators will not be circumscribed solely because of the limitations of the extended indictor itself. The KAA demonstrates strong categorical concurrence (i.e., every general standard has a corresponding extended standard), and the requirement of the KAA that selected indicators be distributed among the standards ensures that students taking the KAA have access to all aspects of general curriculum content. Last, reviewers generally found that extended indicators were not grade specific and were therefore age appropriate at each referenced grade.

The Kansas Extended Standards are organized in such a way that there is a multiplicity of references or intended links between the indicators under the extended and general standards. This situation has both positive and negative features. Teachers are guided by the general standards in their selection of indicators and tasks for students taking the KAA, thus increasing their awareness and knowledge of the content of the general standards as well as demonstrating the intended alignment of that content at the indicator level. Unfortunately, there is a downside to the many cross-references that exist in the current extended standards documents. First, when there are multiple links between an extended indicator and a general indicator (or vice versa), not all of those links can be expected to have equally strong fidelity or consistency. This was confirmed by content reviewers who found inconsistent relationships between the extended indicators and their intended matching general indicators on the basis of content and performance fidelity. Second, using a single extended indicator as a reference for several general indicators that increase in difficulty or scope across grades fails to build in expectations of growth for students in the KAA population. It is possible that students could be assessed on the same indicator several times or could return to it after having mastered it in a previous year. This problem is referenced in the Differentiation of Content Across Grades/Grade Bands portion of the alignment study, in which reviewers concluded that most Science and Reading indicators represented identical content at higher grades. The following steps are suggested:

1. The links between a particular extended indicator with a number of general indicators may be examined in order to select or retain only the most consistent links, based on guidance from the content and performance fidelity ratings of the reviewers and their judgment of the primary and secondary indicator matches.

2. Among the comparisons that could be made toward the end of selecting the best aligned extended indicators for general content indicators is a specific comparison of extended and general indicators, or a subset of key indicators, on the basis of the existing depth-of-knowledge ratings. KSDE has previously evaluated the depth-of-knowledge of the general indicators with Bloom's taxonomy or with other, content-specific criteria.

3. Evaluating the differences in the links between extended and general indicators for reading versus math and science will help to explain why the former links are rated more favorably than the latter in terms of content and performance fidelity.

4. Evaluating the appropriateness of links between extended and general indicators is suggested at higher grade levels across all three content areas. Perhaps a smaller set of extended indicators should be identified as adequate references for high school grade-level content than for lower grades.

5. Finally, evaluating changes in general indicator scope across grades could guide development of additional extended indicators to prompt higher performance expectations for KAA students as they move through those grades.

Kansas Alternate Assessment. In terms of the KAA, two additional alignment criteria to those discussed above were evaluated in this study. Range-of-knowledge correspondence addresses the amount of content assessed out of a potentially large number of extended indicators, especially when an AA-AAS requires fewer indicators or test items for assessment than do general tests. The LAL recommends a minimum of 40% of coverage of extended standards on the AA-AAS, whereas the KAA reaches only 10–18% content coverage in reading, math, and science. Balance of representation evaluates the distribution of the KAA tasks or items with respect to general standards and indicators. This index has a somewhat stronger outcome than range of knowledge for the KAA because it references the coverage of each general standard in terms of the number of aligned extended indicators within that standard. Because of the requirement of the KAA that extended indicators be selected from many standards, balance of representation is strongest when there are more standards with fewer indicators, as in Science, and weaker when there are few standards with more indicators, as in Reading. The KAA generally reached recommended levels for balance of representation. The following steps are suggested:

1. Evaluation of the degree to which the number of KAA assessed indicators addresses the content in the extended standards would be beneficial. Though each assessed indicator requires sufficient trials and opportunities for reliable assessment, the concern with respect to range of knowledge is the limited selection of content required by the KAA that serves to classify students into performance categories. Because the KAA often assesses only a single indicator for each standard, the breadth of assessment for that standard may be minimal. This concept of content coverage will be an important aspect of development of the successor to the KAA under the Common Core State Standards.

2. The balance of representation of the general standards by the assessed indicators is strongest when many standards each have few indicators, as in Science. In other content areas, decisions about balance of representation of general curricular content will become crucial when the KAA successor is formulated to assess the Common Core State Standards. In these areas, broader coverage of the standards, whether general or extended, may be a desirable goal.

KAA Administration and Scoring. According to special education reviewers, KAA strong points included the discrimination offered by the 5-point skill performance rubric and the information required by the evidence labels. Another positive feature of the KAA is the opportunity to use any accommodation typically used by the student for instruction. Other aspects of the KAA were rated lower or were a source of confusion for reviewers. Recommendations are offered that may reduce confusion for future users and evaluators of the KAA:

1. Reviewers were concerned about the use of prompts for student responses and the possibility that low levels of correct responses could earn too many points on the performance rubric. Outcomes from the

5-point performance rubric could be analyzed to determine whether too many points can be earned for chance levels of responding. The KAA Teacher's Guide could be more specific on the use of prompts for student responses and guidelines for how prompting should influence scores on the performance rubric.

2. KSDE specifically allows the use of any accommodations necessary for each student on the KAA, and these are to be listed on the evidence labels. However, reviewers found no specific discussion of accommodations and supports in administration materials. Without clearly defined accommodations and knowledge about how and when they are used, consistent performance on extended indicators cannot be assumed and student responses are difficult to interpret. This is especially true for students with minimal communication skills who need significant accommodations or whose responses require high levels of inference on the part of the teacher. Better guidance for teachers on how to use and record accommodations on the KAA is suggested. Evidence labels could be more explicit in requiring accommodations to be listed, perhaps by using a checklist of possible accommodations with several additional places to write in other accommodations.

3. Reviewers had difficulty determining when new learning had occurred given the lack of baseline data or previous performance. This concern would be lessened by requiring performance objectives to increase in symbolic communication skill, cognitive complexity, or content depth or breadth from year to year.

4. The KAA Teacher's Guide requires each piece of evidence to be unique, and professional development materials state that generalization is a goal, but evidence of generalization is not currently required on evidence labels. Categories of generalization (e.g., people, settings, materials, activities) could be required on the evidence labels, which would prompt teachers to include generalization goals when teaching specific academic skills.

Professional Development Resources and Program Quality Indicators. Reviewers generally agreed that teachers were provided with materials and guidance to use the extended standards effectively and to address grade-level academic skills for students within the framework of the linked general and extended indicators. One program quality indicator endorsed by reviewers was the provision for assistive technology by KAA participants. Other areas received somewhat lower ratings by reviewers. For example, professional development materials focused on assessment administration and scoring rather than instructional practices leading to higher performance on the KAA. The following steps are suggested:

1. Professional development materials and resources could be reviewed to determine where information on instructional practices leading to higher performance on the KAA could be included. Instructional practices rated as insufficiently addressed include delivering instruction at various depths of knowledge, promoting student mastery by decreasing prompting and increasing independent responding, increasing expectations across grades or grade bands, increasing generalization of skills, and minimizing barriers for students with sensory and physical challenges and lower levels of symbolic communication.

2. Program quality indicators that were found to be missing from professional development materials include opportunities for instruction in general education classrooms and with general education materials and resources, instruction with typical peers, opportunities for self-determination, continuity of literacy instruction across content areas, and linking of academic skills in functional contexts. Analysis of how these could receive greater focus and attention would be useful for teachers of KAA students.

Enacted Curriculum. The CIS provides insight into the enacted curriculum for students who participate in the KAA. One question was added to the CIS in order to determine whether the Kansas Extended Standards are easy to understand and use. Teachers overwhelmingly responded that the standards are clearly worded and that their instructional purpose is clear. While there are no high or low ratings on the CIS with respect to the KAA, information from teachers can be used to inform professional development activities and materials to improve education for students with significant disabilities. The following steps are suggested:

1. Teachers reported the lowest level of professional development activity in science and the highest in reading. They also reported using more types of materials in reading and math than in science. While this reflects an understandable focus on literacy, teachers may benefit from professional development in science that will enable them to make science concepts more accessible to students who participate in the KAA.

2. Classroom assessment was reported as primarily teacher observation and secondarily performance on demand. Objective questions were used least frequently and by the fewest teachers in all content areas. Classroom assessment is another professional development topic that could lead teachers toward improved assessment techniques and more consistent assessment approaches. While objective questions are not necessarily a better form of assessment, the benefits include greater standardized administration and more repeatable measurements. Performance tasks can also be structured to be standardized, reliable measures.

3. In the CIS, across all three content areas, teachers most frequently rated content-specific concepts as receiving no instruction at all during the current school year. Only a few topics received intensive coverage from a majority of teachers, including *beginning reading* and *number sense.* Teachers may benefit from professional development addressing ways in which the breadth of content taught to KAA students could be improved. Incorporating broader (and possibly deeper) content coverage into classroom lessons would increase the number of extended indicators accessible to students.

4. Though content experts rated the extended indicators at high cognitive levels, teachers rated their performance expectations for students taking the KAA on similar content at low cognitive levels, primarily *attention* and secondarily *performance*. This may indicate a need for professional development regarding performance expectations of students who participate in the KAA. As noted by KSDE staff, teachers who select assessment targets may unintentionally shy away from high but attainable expectations for growth and skill development in favor of easier-to-achieve assessment goals.

References

Atkin, A. (2010). Peirce's theory of signs. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2010 ed.). Retrieved from http://plato.stanford.edu/entries/peirce-semiotics/

Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain*. New York, NY: David McKay Co., Inc.

Browder, D. M., Ahlgrim-Delzell, L., Courtade-Little, G., & Snell, M. (2005). General curriculum access. In M. Snell & F. Brown (Eds.). *Instruction of students with severe disabilities* (6th ed., pp. 489–525). Upper Saddle River, NJ: Merrill Prentice Hall.

Browder, D. M., Flowers, C., & Wakeman, S. Y. (2008). Facilitating participation in assessments and the general curriculum: Level of symbolic communication classification for students with significant cognitive disabilities. *Assessment in Education: Principles, Policy, & Practice, 15,* 137–151. DOI: 10.1080/09695940802164176

Flowers, C., Wakeman, S., Browder, D. & Karvonen, M. (2007). *Links for academic learning: An alignment protocol for alternate assessments based on alternate achievement standards*. Charlotte, NC: University of North Carolina at Charlotte.

Icon. (2011). In *Merriam-Webster online collegiate dictionary* (11th ed.). Retrieved from http://unabridged.merriam-webster.com/cgi-bin/collegiate?va=icon&x=0&y=0

Icon. (2011). In *Oxford dictionaries online.* Retrieved from http://oxforddictionaries.com/view/entry/m_en_us1256872#m_en_us1256872

Individuals with Disabilities Education Act, 20 U.S.C. §1400 *et seq.* (2004).

Kansas State Department of Education. (2010). *Kansas 2010-2011 Alternate Assessment Teacher's Guide*. Topeka, KS: Author. Retrieved from http://www.ksde.org/LinkClick.aspx?fileticket=X80HpN0L90Q%3d&tabid=2384&mid=8879

Karvonen, M., Wakeman, S. Y., Flowers, C., & Browder, D. M. (2007). Measuring the enacted curriculum for students with severe cognitive disabilities: A preliminary investigation. *Assessment for Effective Intervention*, *33*, 29–38.

Kearns, J. & Towles-Reeves, L. (2007). Learner Characteristics Inventory (LCI) report prepared for the Kansas Department of Education. Lexington, KY: National Alternate Assessment Center.

McLaughlin, K., & Cascella, P. W. (2008). Eliciting a distal gesture via dynamic assessment among students with moderate to severe intellectual disability. *Communication Disorders Quarterly, 29,* 75–81. DOI: 10.1177/1525740107311821

No Child Left Behind Act of 2001. Public Law 107–110, 115 Stat. 1425.

O'Toole, C., & Chiat, S. (2006). Symbolic functioning and language development in children with Down syndrome. *International Journal of Language & Communication Disorders, 41,* 155–171. DOI: 10.1080/13682820500221600

Poggio, A., Yang, X., Irwin, P., Glasnapp, D., & Poggio, J. (2007). Kansas Assessments in Reading and Mathematics: 2006 Technical Manual. Retrieved from http://www.ksde.org/Default.aspx?tabid=2374

Poncelas, A. & Murphy, G. (2007). Accessible information for people with intellectual disabilities: Do symbols really help? *Journal of Applied Research in Intellectual Disabilities, 20,* 466–474. DOI: 10.1111/j.1468-3148.2006.00334.x

Rowland, C. & Schweigert, P. (1990). *Tangible symbol systems: Symbolic communication for individuals with multisensory impairments.* Eugene, OR: Oregon State Department of Higher Education. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED319154

Sutton, A., Trudeau, N., Morford, J., Rios, M., & Poirier, M-A. (2009). Preschool-aged children have difficulty constructing and interpreting simple utterances composed of graphic symbols. *Journal of Child Language, 37,* 1–16. DOI: 10.1017/S0305000909009477

Symbol. (2011). In *Merriam-Webster online collegiate dictionary* (11th ed.). Retrieved from http://unabridged.merriam-webster.com/cgi-bin/collegiate?va=symbol&x=0&y=0

Symbol. (2011). In *Oxford dictionaries online.* Retrieved from http://oxforddictionaries.com/view/entry/m_en_us1296461#m_en_us1296461

Tomasello, M., Striano, T., & Rochat, P. (1999). Do young children use objects as symbols? *British Journal of Developmental Psychology, 17,* 563–584.

Towles-Reeves, E., Kearns, J., Kleinert, H., & Kleinert, J. (2009). An analysis of the learning characteristics of students taking alternate assessments based on alternate achievement standards. *The Journal of Special Education, 42,* 241–254. DOI: 10.1177/0022466907313451