

# Alignment Between the Kentucky Core Content for Assessment and the WIDA Consortium English Language Proficiency Standards

**October 9, 2007**



H. Gary Cook, Ph.D.  
Carsten Wilmes, Ph.D.

Report prepared by the WIDA Consortium  
Wisconsin Center for Education Research  
1025 W. Johnson Street  
University of Wisconsin-Madison  
Madison, WI 53706

## Table of Contents

Executive Summary.....	1
1. Introduction.....	2
Background.....	2
Alignment Methods.....	3
Standards-to-Standards Alignment Criteria.....	4
Linking.....	6
Correspondence.....	7
Standards Aligned in this Study.....	8
Participants and Review Process.....	11
2. Results .....	13
Reading Alignment Results .....	13
Mathematics Alignment Results .....	16
Science Alignment Results .....	19
Reliability among Reviewers.....	22
3. Summary .....	23
References .....	25
Appendix A	
General Comments by Reviewers.....	27
Appendix B	
Examples of Linking and Non-Linking Standards.....	49

## List of Tables

Table 1	NCTM Standards.....	7
Table 2	English Language Proficiency to Academic Content Standard Standard-to-Standard Alignment Criteria.....	8
Table 3	Alignment Study Participants.....	11
Table 4	Alignment for Reading across Grades 3-12.....	13
Table 5	Summary of Alignment for Reading across Grades 3-5.....	15
Table 6	Summary of Alignment for Reading across Grades 6-8.....	16
Table 7	Summary of Alignment for Reading across Grades 9-12.....	16
Table 8	Alignment for Mathematics across Grades 3-12.....	17
Table 9	Summary of Alignment for Mathematics across Grades 3-5.....	18
Table 10	Summary of Alignment for Mathematics across Grades 6-8.....	18
Table 11	Summary of Alignment for Mathematics across Grades 9-12.....	19
Table 12	Alignment for Science across Grades 3-12.....	19
Table 13	Summary of Alignment for Science across Grades 3-5.....	21
Table 14	Summary of Alignment for Science across Grades 6-8.....	22
Table 15	Summary of Alignment for Science across Grades 9-12.....	22
Table 16	Reliability among Reviewers.....	23

## List of Figures

Figure 1	Standards-to-Standards Alignment of Highly Similar Constructs.....	5
Figure 2	Standards-to-Standards Alignment of Associated Constructs.....	6
Figure 3	The Format of the English Language Proficiency Standards for Large-Scale State and Classroom Frameworks.....	10

## Executive Summary

This report presents the results of an alignment study conducted in Frankfort, Kentucky on September 7 and 8, 2006 using the Web Alignment Tool (WAT). The alignment protocol is based on Cook's (2005, 2006, 2007) adaptation of Webb's (1997) alignment framework; for this study Cook's framework was used to examine the relationship between Kentucky's Core Content for Assessment (academic content standards) in Reading, Mathematics, and Science and the model performance indicators (MPIs) within the WIDA English Language Proficiency (ELP) Standards.

### What is alignment?

Federal guidance refers to two criteria to evaluate the relationship between English language proficiency standards and a state's academic content standards: linking and alignment (U.S. Department of Education, Office of English Language Acquisition, February 2003). Linking is required as a minimum criterion; alignment, the higher criterion, is encouraged. In our conceptualization, alignment is the combination of linking (match between standards) and correspondence (comprised of depth and coverage). *Depth* refers to similarity of cognitive complexity and *coverage* to similarity in dispersion. Each aspect of the alignment has associated statistics: *Link*, *Depth of Knowledge (DOK) Consistency* (depth), and *Coverage* (breadth). Alignment is a higher criterion as it not only examines whether there is a match between standards (linking), but also establishes whether there is strong cognitive correspondence between standards and whether a state's content goals within a content standard have corollary English proficiency expectations (correspondence).

### Linking

Results suggest strong linkage across all grade clusters between the MPIs in the WIDA English Language Proficiency Standards and the three academic content standards investigated in this study. We therefore conclude that the relationship between Kentucky's Core Content for Assessment in Reading, Mathematics, and Science and MPIs within the WIDA ELPs meets NCLB requirements, with some limitations in the 3-5 and 9-12 grade clusters.

### Correspondence

As stated above, federal guidance encourages states to meet a higher standard, i.e., alignment. Our analyses indicate that the Depth criterion is largely met for Reading and Mathematics. This is not the case for Science. In addition, Coverage tends to be somewhat limited for all three content areas. Overall, we conclude that while the alignment criteria as defined here are not entirely met, Kentucky's Core Content for Assessment in Reading and Mathematics align moderately with the MPIs within the WIDA ELPs; alignment in Science is rather limited.

## 1. Introduction

### Background

This study was an evaluation of the alignment between the Kentucky Core Content for Assessment and the WIDA English Language Proficiency Standards in the areas of Reading, Mathematics, and Science. Webb's (1997) alignment methodology, which has traditionally been used to evaluate the alignment between academic content standards and academic content assessments, has recently been adapted to study the alignment between different sets of standards (e.g., English language proficiency and academic content). Cook (2005) explains that more of a one-to-one correspondence is expected when aligning two sets of standards than when examining the alignment between a set of standards and an assessment. Thus, the criteria for acceptable levels of key alignment statistics are different for standards-to-standards alignment than for test-to-standards alignment.

The text below is drawn from federal non-regulatory guidance as it relates to English language proficiency standards and the issue of alignment.

#### **B-3. What is the relationship between English language proficiency standards, English language proficiency annual measurable achievement objectives, and English language proficiency assessment?**

English language proficiency standards *must, at a minimum, be linked* [highlighting not in original] to the State academic content and achievement standards. States *are encouraged, but not required, to align* [highlighting not in original] English language proficiency standards with academic content and achievement standards. Annual measurable achievement objectives for English language proficiency serve as targets for achievement of the English language proficiency standards. English language proficiency assessments must be aligned with English language proficiency standards and provide a means of demonstrating progress towards meeting the English language proficiency annual measurable achievement objectives. (U.S. Department of Education, Office of English Language Acquisition, February 2003, pp.9, 10).

Note the italicized, highlighted phrases in the text above. Herein the federal government has expanded upon the notion of alignment, traditionally seen as a relationship between standards and assessments, to include the relationship between a state's English language proficiency standards and its academic content standards. Guidance sets forth a

minimum criterion of linking student expectations and offers the “gold standard” as alignment. While little research is available describing the nature and scope of linking one set of standards to another, there has been work examining alignment between standards.

## **Alignment Methods**

The alignment of assessment systems to state standards (test-to-standards alignment) has gained prominence in recent years. The No Child Left Behind Act of 2001 (NCLB) requires alignment of state assessments to state standards. The notion of alignment is not new. Alignment is and has been a mechanism for assuring a test’s content validity. In years past, however, alignment was often evaluated in a very ad hoc fashion. Typically, alignment activity was conducted during a test’s item review. Content experts reviewed assessment items and determined if items matched test specifications, test framework documents, or standards. The primary purpose in this type of alignment was to assure that a test item matched a specification, framework or standard. Researchers have argued that there is more to alignment than just matching (see La Marca, et al., 2001; Webb 1997, 2002; and Rothman, et al., 2002). Alignment refers not only to matching items to standards but also to ascertaining the breadth and the cognitive depth of items relative to standards.

A variety of alignment strategies and methodologies exist (see CCSSO, 2002 & 2007). One of the most prominent methods used today is that created by Dr. Norman Webb of the Wisconsin Center for Educational Research. The Webb approach to alignment evaluates item match, cognitive complexity (or depth), and breadth of coverage. Each alignment component (match, depth, breadth) has associated statistics.

To evaluate match, the statistic Categorical Concurrence is used. Categorical Concurrence refers to the average number of items raters assign to specific standards or curricular goals. Raters select specific standards, goals or objectives that match to individual test items on rated tests. The numbers of coded items are averaged across all raters and reported as Categorical Concurrence. Think of this statistic as a proxy for average numbers of items raters believe address a specific standard or objective. With this methodology, items can address more than one standard, and raters are allowed to code accordingly.

To evaluate depth, raters judge the Depth of Knowledge (DOK) of standards, goals and/or objectives and the DOK of test items. Depth of knowledge can be defined in a variety of ways. Webb argues that,

Standards vary on the complexity of what students are expected to know and do. Some standards simply expect students to reproduce a fact or complete a sequence of steps while others expect students to reason, extend their thinking, synthesize information from multiple sources, and produce significant work over time. Alignment on depth-of-knowledge is achieved when the assessment and standards

agree on the cognitive level students are expected to demonstrate and are asked to perform. (Webb, 2001).

Webb identifies four DOK levels. They are as follows:

Level 1 Recall and Reproduction,  
Level 2 Skills and Concepts,  
Level 3 Strategic Thinking, and  
Level 4 Extended Thinking.

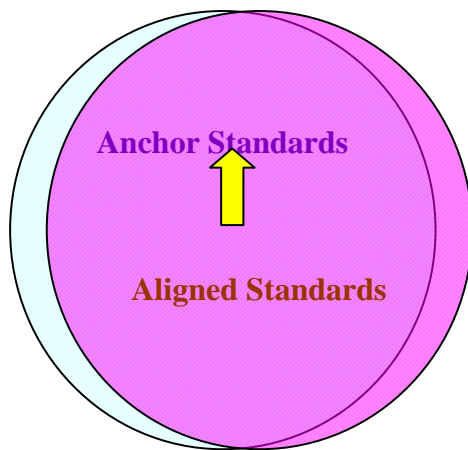
During the alignment process test items and standards are assigned unique DOK levels, and these levels are compared to identify their correspondence. The final component analyzed in a Webb alignment is breadth. Two statistics are associated with breadth: Range and Balance. The Range “criterion is met if a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the assessment items/activities” (Webb, 2001). If test items are identified with most, if not all, relevant objectives in a standard, then it is said that there is good Range. In essence, Range examines whether all objectives within a goal or standard are adequately covered. The second statistic examining breadth is Balance. Balance refers to the “degree to which one objective is given emphasis on the assessment is comparable to the emphasis given to the other objectives within a standard” (Webb, 2001).

## **Standards-to-Standards Alignment Criteria**

Webb alignments focus on state tests and state academic content standards, usually in the areas of reading and mathematics. Federal linking or alignment guidance described above differs. Instead of examining test-to-standards (i.e., Webb’s approach), requirements suggest conducting standards-to-standards investigations, be they linking or alignment. A variety of procedures have been developed to “align” curriculum in education (Anderson, 2002). A very prominent example is the Surveys of Enacted Curriculum (Porter and Smithson, 2001 and Blank, 2002). With this approach, researchers examine relationships between standards, instructional practices, and assessments. The power of this approach is to unveil how standards-based, assessment evaluated systems are realized in the classroom. This approach is very comprehensive and informative. It does not solely focus on examining two sets of standards per se. Undoubtedly, it could be adapted to accomplish this. Another approach to examine standard-to-standard relationships has been applied to sets of standards using a modified version of the Webb alignment procedure (Cook, 2005). With this method, Cook aligned a state’s academic framework to a district’s learning targets. The goal of this alignment was to communicate the association between the district’s standards and the state’s standards for assessment. The district’s learning targets were developed to support the state’s assessment framework, as such good alignment was anticipated between these two sets of student expectations. Close correspondence, however, might not always be the

expectation in a standards-to-standards alignment. This distinction is highlighted by the figures below.

**Figure 1: Standards-to-Standards Alignment of Highly Similar Constructs**



In Figure 1, the anchor standards are defined as expectations that one aligns to, e.g., state standards/ assessment frameworks, and aligned standards are expectations to be aligned, e.g., learning targets. For example, one might align one set of mathematics standards at 4<sup>th</sup> grade to another set of mathematics standards at 4<sup>th</sup> grade. A high degree of overlap (i.e., match, depth and breadth) would represent good alignment. Note, however, that Figure 1 portrays alignment between highly similar constructs—in our example 4<sup>th</sup> grade mathematics. Would this be the expected alignment between associated constructs, say between elementary, mathematics academic language standards for grades 3 through 5 and 4<sup>th</sup> grade mathematics content standards? Probably not. Continuing this line of reasoning, alignment between language proficiency standards and academic content standards is best reflected in Figure 2. Were Figure 1 the target, why have different standards?



**Figure 2: Standards-to-Standards Alignment of Associated Constructs**

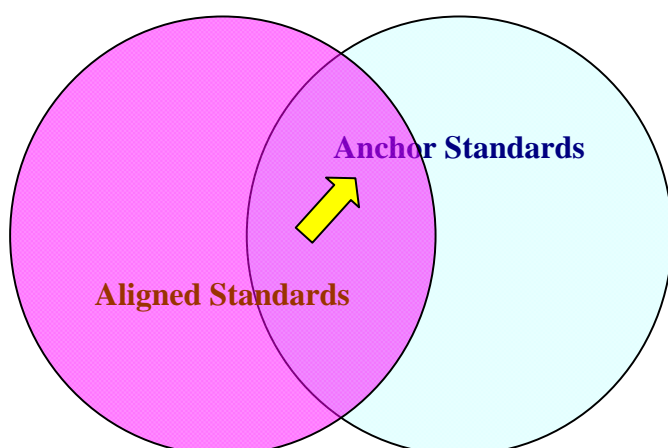


Figure 2 portrays association between two sets of standards—the association of related but not identical expectations. The distinction between academic content standards-to-standards alignment and English language proficiency standards-to-standards alignment is what is being compared. In content alignment, subject matter expectations are being compared. In an English language proficiency alignment, content register relationships are being compared. The register used in subject areas like mathematics, science or language arts are subsets of the content domain. As a result, the criterion for alignment should differ.

As stated earlier, federal guidance identifies two notions related to academic content and language proficiency standards alignment: link and align. We interpret the term *alignment* mentioned in federal guidance to be that reflected by Figure 2. That is, strong alignment between English language proficiency standards and academic content standards *ARE NOT* one-to-one correspondences. What then does alignment mean?

## **Linking**

First, a state’s English language proficiency standards must be, at a minimum, linked to its academic content standards. **BY LINKED, WE MEAN THAT AT LEAST ONE ALIGNED CONTENT STANDARD IN EACH ASSESSED SUBJECT MUST BE REPRESENTED IN THE ENGLISH LANGUAGE PROFICIENCY STANDARDS AT EACH GRADE BAND.** An example will help clarify this criterion. Table 1 displays elements of the National Council of Teachers of Mathematics (NCTM) standards. Let us assume that Table 1 reflects a state’s mathematics standards at a particular grade. To be appropriately linked, linguistic elements (i.e., phonological, lexical, grammatical, sociolinguistic) associated with Number Operations, Algebra, Geometry, Measurement, and Data Analysis and Probability would need to be reflected in the English language proficiency standards for speaking, listening, reading or writing at the grade span associated with this standard. A language proficiency standard requiring students to orally describe groups of and/or sequences of objects, figures or numbers would be consistent with Number and Operations. Another standard might have students read a graph or figure representing numeric relationships. This standard could be linked to Algebra and possibly Data

Analysis and Probability. Linking assures that register elements associated with the language of mathematics are included in language proficiency standards.

**Table 1: NCTM Standards**

<b>Standards</b>	<b>Goals</b>
Number and Operations	<ol style="list-style-type: none"> <li>1. Understand numbers, ways of representing numbers, relationships among numbers, and number systems;</li> <li>2. Understand meanings of operations and how they relate to one another;</li> <li>3. Compute fluently and make reasonable estimates;</li> </ol>
Algebra	<ol style="list-style-type: none"> <li>1. Understand patterns, relations, and functions;</li> <li>2. Represent and analyze mathematical situations and structures using algebraic symbols;</li> <li>3. Use mathematical models to represent and understand quantitative relationships;</li> <li>4. Analyze change in various contexts;</li> </ol>
Geometry	<ol style="list-style-type: none"> <li>1. Analyze characteristics and properties of two- and three-dimensional geometric shapes and develop mathematical arguments about geometric relationships;</li> <li>2. Specify locations and describe spatial relationships using coordinate geometry and other representational systems;</li> <li>3. Apply transformations and use symmetry to analyze mathematical situations;</li> <li>4. Use visualization, spatial reasoning, and geometric modeling to solve problems;</li> </ol>
Measurement	<ol style="list-style-type: none"> <li>1. Understand measurable attributes of objects and the units, systems, and processes of measurement;</li> <li>2. Apply appropriate techniques, tools, and formulas to determine measurements;</li> </ol>
Data Analysis and Probability	<ol style="list-style-type: none"> <li>1. Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them;</li> <li>2. Select and use appropriate statistical methods to analyze data;</li> <li>3. Develop and evaluate inferences and predictions that are based on data;</li> <li>4. Understand and apply basic concepts of probability;</li> </ol>

## Correspondence

Federal guidance states that linking is a minimum criterion. Alignment is encouraged. **ALIGNMENT, IN OUR CONCEPTUALIZATION, IS THE COMBINATION OF LINKING AND CORRESPONDENCE.** Table 2 shows this relationship. Linking describes the match between standards. Correspondence includes Depth and Breadth. For Depth, we adopt a criterion of 40%. That is, 40% of linked English language proficiency standards should be at or above the Depth of Knowledge (DOK) level of the content standards to reflect strong cognitive correspondence between standards. The DOK criterion associates with Scarcella's (2003) cognitive dimension, including higher-order thinking, strategic competence, and metalinguistic awareness. A 40% DOK criterion establishes challenging but attainable expectations.

**Table 2: English Language Proficiency to Academic Content Standard Stanard-to-Standard Alignment Criteria**

Scope		Criterion	
Alignment	Link	Match	At least one aligned content standard across skill domains, as agreed upon by a <b>majority</b> of raters
	Correspondence	Depth	At least a 40% DOK across skill domains
		Breadth	At least moderate Coverage of goals across domains where: <b>Limited</b> $\leq 1$ goal aligned for each standard, <b>Moderate</b> $> 1$ goal aligned for each standard, <b>Strong</b> = a majority of goals aligned for each standard

The second aspect of Correspondence is Breadth. The Breadth criterion relates to the number of goals within a standard that are aligned. In Table 1, we see there are 3 goals for Number and Operations, 4 goals for Algebra, 4 goals for geometry, 2 goals for Measurement, and 4 goals for Data Analysis and Probability. Moderate breadth would mean that more than one goal in the math standards is associated with the language proficiency standards. Strong breadth would mean a majority of a state's content goals within a content standard have corollary English language proficiency expectations. As with the DOK criterion, this is an aggressive but obtainable expectation.

For adequate alignment, we suggest that a state's English language proficiency standards should meet the linking criterion, the DOK criterion, and have moderate or greater breadth of coverage. Were language proficiency standards to have this degree of alignment, we believe greater attention would be given to Academic English in the classroom and on language proficiency assessments. Given Gottlieb's (2006) conviction that Academic English language proficiency is a precursor to academic achievement, good alignment would promote students' progress in English, which could directly affect annual measurable achievement objective (AMAO) goals. This type of alignment would move states toward best practice in language instruction and assessment.

### **Standards Aligned in this Study**

The following are brief descriptions of the two sets of standards aligned in this study:

#### Kentucky Core Content for Assessment

The *Core Content for Assessment* (version 4.1) reflects the content that all students in Kentucky are expected to know and that will be included on the state assessment. These content standards provide the basis for the *Kentucky Core Content Test* (KCCT). Within each content domain, standards are organized by grade level (end of primary, fourth, fifth, sixth, seventh, eighth and high school) and include several subdomains. Within each subdomain are specific topics (organizers) that categorize the standards.

The *Core Content for Mathematics Assessment* includes the following five subdomains:

- Number Properties and Operations
- Measurement
- Geometry
- Data Analysis and Probability
- Algebraic Thinking

The *Core Content for Reading Assessment* is also comprised of five subdomains:

- Forming a Foundation for Reading
- Developing an Initial Understanding
- Interpreting Text
- Reflecting and Responding to Text
- Demonstrating a Critical Stance

The *Core Content for Science Assessment* includes four subdomains:

- Physical Science
- Earth/Space Science
- Biological Science
- Unifying Concepts

#### WIDA English Language Proficiency Standards

The WIDA English Language Proficiency Standards (WIDA, 2004) are comprised of the following five standards:

1. English language learners communicate in English for SOCIAL AND INSTRUCTIONAL purposes within the school setting.
2. English language learners communicate information, ideas, and concepts necessary for academic success in the content area of LANGUAGE ARTS.
3. English language learners communicate information, ideas, and concepts necessary for academic success in the content area of MATHEMATICS.
4. English language learners communicate information, ideas, and concepts necessary for academic success in the content area of SCIENCE.
5. English language learners communicate information, ideas, and concepts necessary for academic success in the content area of SOCIAL STUDIES.

Each standard covers four language domains: listening, speaking, reading, and writing. The model performance indicators for each standard are organized into four grade-level clusters (K-2, 3-5, 6-8, and 9-12) and two frameworks: classroom assessment and large-scale assessment. Within each framework, grade cluster and language domain, there are model performance indicators for each language proficiency level. The model performance indicators are functional, measurable indices of the language domains (listening, speaking, reading, and writing) and aimed at the targeted age/developmental levels of English language learners. As their label implies, model performance indicators

are merely examples that have been drawn from a myriad of English language proficiency and state academic content standards. There are three components of a model performance indicator: 1). function (how the students use language), 2). content (what the students are expected to communicate), and 3). modality (how the students process the input either through oral or written language). For some indicators, there are suggested topics that add clarity or specificity; these ideas are introduced by the phrase “such as.” Other indicators have “e.g.,” followed by an example of an expected language pattern that students may use in their response.

At times, there are two strands of model performance indicators within a grade level cluster; reviewers of the document felt that these additions were necessary to create a closer alignment with state academic content standards. A visual layout of the components of the standards is displayed in Figure 5. The English language proficiency levels head each column and the grade level clusters begin each row. The remaining cells contain at least one model performance indicator, creating a strand or strands across proficiency levels within a grade level cluster. (Figure 5 points to an example of a strand of performance indicators for grade level cluster 3-5.)

**Figure 3. The Format of the English Language Proficiency Standards for Large-scale State and Classroom Frameworks**

STANDARD

LANGUAGE DOMAIN

**English Language Proficiency Levels**

Grade Level Cluster	Level 1 Entering	Level 2 Beginning	Level 3 Developing	Level 4 Expanding	Level 5 Bridging
K-2					
3-5					
6-8					
9-12					

**Model Performance Indicator(s)**

**A Strand of Model Performance Indicators (through grade level cluster)**

The five language proficiency levels covered in the Standards are Level 1 – Entering, Level 2 – Beginning, Level 3 – Developing, Level 4 – Expanding, and Level 5 – Bridging.

### Participants and Review Process

The alignment workshop was conducted in Frankfort, Kentucky on September 7 and 8, 2006. Forty-one Kentucky educators with expertise in the *Core Content* expectations (CCE) or ESL served as alignment reviewers. The participants were grouped based on their grade level experience into three panels: one for Grades K to 5, another for Grades 6 to 8 and the other for 9 to 12—for each content area. The following are the names of the participants, their grade cluster, and area of expertise:

<b>Table 3: Alignment Study Participants</b>		
<b>Grade Cluster</b>	<b>Participant Name</b>	<b>Expertise Area</b>
K-5	Laura Pinkerton	CCE (Reading)
K-5	Elizabeth Lewis	ESL
K-5	Kathy Gordon	ESL
K-5	Vongmany Edmonds	ESL
K-5	Beth Gniot	ESL
K-5	Maria Scherrer	ESL
K-5	Margaret Hill	CCE
K-5	Patty Allen	CCE
K-5	Sonia James	ESL
K-5	Karen Botts	ESL
K-5	Mary Morgan	ESL
K-5	Beverly Stevens	ESL
K-5	Daniele Novak	ESL
K-5	Greg Howell	CCE (Science)
K-5	Heather Johnson	ESL
K-5	Susan Reed	ESL
K-5	Stella Loveland	ESL
K-5	Leesa Moman	CCE and ESL
K-5	Vanessa Dials	CCE
K-5	Angela Gabbard	CCE
6-8	Cathy Fernandez	ESL
6-8	Katy Stephens	CCE (Science)
6-8	Shannon Lindsey	CCE (Science and Math)
6-8	Marti Kinney	ESL
6-8	Jayne Kraemer	ESL
6-8	Julester Bennett	ESL
6-8	Sean Elkins	CCE (Science)

<b>Table 3: Alignment Study Participants</b>		
<b>Grade Cluster</b>	<b>Participant Name</b>	<b>Expertise Area</b>
6-8	Danna Morrison	ESL
6-8	Lisa Hillenbrand	ESL
6-8	Kathy Holland	CCE
9-11	Cindy Parker	CCE (Reading)
9-11	Danielle Burke	CCE
9-11	Phyllis Shuttleworth	CCE (Science)
9-11	Robin Hill	CCE (Math)
9-11	Latisha Sparks	CCE
9-11	Nichole Neuhard	ESL
9-11	Sandy Byrd	ESL
9-11	Scott Kremer	ESL
9-11	Chris Brady	CCE (Math)
9-11	David Gibson	CCE (Math)
9-11	Ivonne Beagle	ESL

To facilitate the alignment workshop, external consultants from the Wisconsin Center for Education Research (WCER) and the University of Illinois at Urbana-Champaign facilitated each of the four groups of Kentucky reviewers who reviewed the assessments.

An intensive training was provided to all reviewers, explaining Webb’s alignment model and the three alignment criteria and the use of the web-based alignment tool. The general training included an overview of the alignment process and a brief description of the standards that would be reviewed. After the general session, the reviewers broke into subject area groups to learn how to apply the *DOK* levels to standards in their respective grade levels. All participants reviewed the definitions of the four levels of *DOK* and sample standards at each level during the content-related training. Following the content-related training, reviewers split into the grade-level groups to continue the alignment process. The process involved five steps:

**Step One** – Reviewers read the KY Reading, Math and Science standards and reached consensus on the appropriate *DOK* level for each objective.

**Step Two** – As training for the review process, each team of reviewers independently coded a sample of model performance indicators drawn from the WIDA standards and then discussed the *DOK* levels and the KY standards that they had assigned to each of the WIDA standards. Reviewers were encouraged to assign only one KY content standard to each WIDA MPI unless the WIDA MPI clearly assessed more than one standard. In cases where a WIDA MPI did not adequately describe the knowledge and skills assessed, reviewers could assign secondary and tertiary standards. Reviewers were not required to reach agreement on the *DOK* assigned to a WIDA MPI. Instead, they discussed the rationale for the assignments to help each other reach a clearer understanding of *DOK* levels and the Reading, Math and Science model performance indicators (MPIs) of the WIDA ELP Standards.

**Step Three** – Reviewers independently coded the WIDA model performance indicators in Reading, Math, and Science for each grade level, identified a KY standard to which each one most closely matched, and noted any issues or sources of challenge related aligning the KY and WIDA standards. Reviewers coded the WIDA MPIs in a different sequence to avoid the order of test review affecting the results of the alignment study.

**Step Four** – After all of the WIDA MPIs in one grade level were reviewed, the reviewers discussed the results as a group. Reviewers discussed MPIs for which fewer than 50% of them agreed on the *DOK* level. Again, reviewers were not required to reach agreement on the *DOK* level assigned to a MPI. Instead, they discussed the rationale for their assignments and changed their assignments only if they felt they had assigned the wrong *DOK* level to a WIDA MPI. Reviewers did not know whether other reviewers kept or changed their ratings.

**Step Five** – Reviewers participated in a debriefing session for each grade level. They had been encouraged to complete a debriefing questionnaire for each test as they reviewed it and to use their notes in the discussion session. During this session, the reviewers provided their impressions about the degree of alignment between the two sets of standards.

The same process was applied to each grade level. At the conclusion of the alignment workshop, reviewers were asked to complete an evaluation questionnaire to provide feedback about the alignment review process.

## 2. Results

### Reading Alignment Results

Based on the alignment criteria specified above, Table 4 below presents findings from the alignment between the Kentucky Reading standards and the WIDA ELP model performance indicators (MPIs) in Reading. The first set of columns presents alignment statistics and the second displays alignment findings based on the criteria set forth in the previous section.

**Table 4: Alignment for Reading Across Grades 3-12**

Subdomains	(Standards-to-Standards) Alignment Criteria					
	Alignment Statistics			Alignment Findings		
	Linked	Correspondence		Linked	Correspondence	
		DOK	Coverage		DOK	Coverage
<b>Grade 3 (with 5 Panelists)</b>		66%				
3.1 - Foundation	0	25%	0 of 5	NO	NO	LIMITED
3.2 - Initial	13	74%	3 of 5	YES	YES	STRONG



Subdomains		(Standards-to-Standards) Alignment Criteria				
		Alignment Statistics		Alignment Findings		
		Linked	Correspondence	Linked	Correspondence	
			DOK	Coverage	DOK	Coverage
Understanding						
3.3 - Interpreting Text	8	60%	3 of 4	YES	YES	STRONG
3.5 - Critical Stance	2	80%	1 of 1	YES	YES	STRONG
<b>Grade 4 (with 4 Panelists)</b>			67%			
4.1 - Foundation	4	47%	1 of 5	YES	YES	LIMITED
4.2 - Initial Understanding	15	65%	4 of 6	YES	YES	STRONG
4.3 - Interpreting Text	12	70%	2 of 7	YES	YES	MODERATE
4.5 - Critical Stance	3	78%	1 of 2	YES	YES	STRONG
<b>Grade 5 (with 4 Panelists)</b>			72%			
5.1 - Foundation	3	44%	1 of 5	YES	YES	LIMITED
5.2 - Initial Understanding	9	78%	2 of 5	YES	YES	MODERATE
5.3 - Interpreting Text	12	73%	4 of 7	YES	YES	STRONG
5.5 - Critical Stance	2	75%	1 of 4	YES	YES	LIMITED
<b>Grade 6 (with 5 Panelists)</b>			76%			
6.1 - Foundation	1	79%	1 of 4	YES	YES	LIMITED
6.2 - Initial Understanding	3	100%	1 of 7	YES	YES	LIMITED
6.3 - Interpreting Text	0	0%	0 of 7	NO	NO	LIMITED
6.5 - Critical Stance	6	62%	2 of 3	YES	YES	STRONG
<b>Grade 7 (with 5 Panelists)</b>			79%			
7.1 - Foundation	1	83%	1 of 2	YES	YES	STRONG
7.2 - Initial Understanding	3	100%	1 of 7	YES	YES	LIMITED
7.3 - Interpreting Text	0	50%	0 of 6	NO	YES	LIMITED
7.5 - Critical Stance	6	61%	2 of 3	YES	YES	STRONG
<b>Grade 8 (with 5 Panelists)</b>			72%			
8.1 - Foundation	1	79%	1 of 3	YES	YES	LIMITED
8.2 - Initial Understanding	3	100%	1 of 6	YES	YES	LIMITED
8.3 - Interpreting Text	0	50%	0 of 6	NO	YES	LIMITED
8.5 - Critical Stance	4	22%	1 of 3	YES	NO	LIMITED
<b>Grade 9 (with 4 Panelists)</b>			73%			
9.1 - Foundation	1	100%	1 of 4	YES	YES	LIMITED
9.2 - Initial Understanding	10	70%	3 of 7	YES	YES	MODERATE
9.3 - Interpreting Text	6	80%	2 of 9	YES	YES	MODERATE
9.4 - Reflecting/Responding	2	100%	2 of 2	YES	YES	STRONG

Subdomains	(Standards-to-Standards) Alignment Criteria				
	Alignment Statistics			Alignment Findings	
	Linked	Correspondence		Linked	Correspondence
		DOK	Coverage	DOK	Coverage
9.5 - Critical Stance	7	47%	3 of 9	YES	MODERATE

The Linking criterion was met for the overwhelming number of subdomains in each grade. The only exceptions are the *Interpreting Text* subdomain for grades 6-8 and the *Foundation* subdomain for grade 3. At least one subdomain was linked at each grade cluster (3-5, 6-8, 9-12). The only exception to this is the *Interpreting Text* subdomain for grades 6-8. Further investigation revealed that two out of the five raters in this area consistently linked 2 WIDA standards (Grade band 6-8, Reading, Language Arts, Level 3 and Grade band 6-8, Reading, Science, Level 4). While not meeting the specified Linkage criterion, this suggests some rater agreement between WIDA and KY reading standards in the area of *Interpreting Text* at grade band 6-8. Plausible causes for the lack of linkage might be: (1) non-traditional interpretation of the subdomain by raters, (2) true misalignment between standards, and (3) the limited number of raters. For the correspondence criterion, both Depth and Coverage need to be investigated. Based on the data presented above, the cognitive complexity of both standards was quite similar in that only *Foundation* for grade 3, *Interpreting Text* for grade 6, and *Critical Stance* for grade 8 did not meet the 40% Depth of Knowledge requirement. The Depth criterion was met for each of the three grade clusters. In terms of Coverage, the representation of the WIDA model performance indicators in the KY Reading standards appears to be strongest for early and late grade clusters (3-5 and 9-12). Notably, the *Foundation* subdomain exhibited only limited Coverage for all grades except grade 7.

Tables 5-7 summarize reading alignment results across grade clusters. Again, to meet Linkage criteria at least 1 linked WIDA standard should be identified for each subdomain across grades. To meet correspondence criteria DOK should be  $\geq 40\%$  across each subdomain, and there should be moderate or strong coverage across subdomains. Adequate alignment would be represented by acceptable Linking and Correspondence.

**Table 5: Summary of Alignment for Reading across Grades 3-5**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Foundation	7	39%	0	0
2 - Initial Understanding	37	72%	1	2
3 - Interpreting Text	32	68%	1	2
5 - Critical Stance	7	78%	0	2

For the 3-5 grade cluster, Linking was met for all four subdomains. In addition, 3 of the 4 subdomains reached the 40% DOK criterion (with *Foundation* coming within 1%).

Coverage was particularly weak for the *Foundation* subdomain, with adequate dispersion for the other subdomains.

**Table 6: Summary of Alignment for Reading across Grades 6-8**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Foundation	3	80%	0	1
2 - Initial Understanding	9	100%	0	0
3 - Interpreting Text	0	33%	0	0
5 - Critical Stance	16	48%	0	2

As noted above, the *Interpreting Text* subdomain did not link for the 6-8 grade cluster. It also failed to meet the Depth criterion. Remaining subdomains met both the Linking and DOK criteria. However, Coverage was limited across 2 of the 4 subdomains.

**Table 7: Summary of Alignment for Reading across Grades 9-12**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Foundation	1	100%	0	0
2 - Initial Understanding	10	70%	1	0
3 - Interpreting Text	6	80%	1	0
4 - Reflecting/Responding	2	100%	0	1
5 - Critical Stance	7	47%	1	0

Linkage was met for all subdomains at the high school level. For alignment, all subdomains reached the required DOK criterion, while falling short in Coverage (again, in the *Foundation* subdomain).

## Mathematics Alignment Results

Table 8 below presents the summary of the alignment for Math across Grades 3-12. Again, the first set of columns present alignment statistics and the second displays alignment findings based on the criteria set forth in the previous section. As with reading, the Linking criterion was met for the vast majority of subdomains. However, Linking was lacking for the *Measurement* and *Data Analysis and Probability* subdomains for grades 3 and 5, the *Algebraic Thinking* subdomain for grade 7, and the *Geometry* and *Algebraic Thinking* subdomain for grade 11. Overall, Linking was met for each grade cluster, except for 9-12. This is likely due to the fact that the 9-12 alignment was only a single grade (grade 11). The Depth of Knowledge criterion was generally met. Coverage, however, was rather limited and appeared strongest for the *Number Properties and Operations* subdomain, which aligned strongly at all grade levels. Of all grade levels, grade 4 exhibited the strongest Coverage, with 4 of 5 subdomains aligning strongly.

**Table 8: Summary of Alignment for Mathematics Across Grades 3-12**

Subdomains	(Standards-to-Standards) Alignment Criteria					
	Alignment Statistics			Alignment Findings		
	Linked	Correspondence		Linked	Correspondence	
		DOK	Coverage		DOK	Coverage
<b>Grade 3 (with 4 Panelists)</b>		75%				
3.1 - Number Properties and Operations	9	63%	2 of 4	YES	YES	STRONG
3.2 - Measurement	0	33%	0 of 1	NO	NO	LIMITED
3.3 - Geometry	8	75%	1 of 2	YES	YES	STRONG
3.4 - Data Analysis and Probability	0	0%	0 of 1	NO	NO	LIMITED
3.5 - Algebraic Thinking	4	100%	2 of 2	YES	YES	STRONG
<b>Grade 4 (with 3 Panelists)</b>		72%				
4.1 - Number Properties and Operations	10	78%	3 of 4	YES	YES	STRONG
4.2 - Measurement	2	88%	2 of 2	YES	YES	STRONG
4.3 - Geometry	8	77%	1 of 3	YES	YES	LIMITED
4.4 - Data Analysis and Probability	1	50%	1 of 2	YES	YES	STRONG
4.5 - Algebraic Thinking	3	77%	1 of 2	YES	YES	STRONG
<b>Grade 5 (with 4 Panelists)</b>		77%				
5.1.1 - Number Sense	8	53%	2 of 4	YES	YES	STRONG
5.1.2 - Estimation		77%			YES	
5.1.3 - Number Operations		83%			YES	
5.1.5 - Properties of Numbers and Operations		0%			NO	
5.2 - Measurement	0	75%	0 of 2	NO	YES	LIMITED
5.3 - Geometry	8	84%	2 of 3	YES	YES	STRONG
5.4 - Data Analysis and Probability	0	100%	0 of 3	NO	YES	LIMITED
5.5 - Algebraic Thinking	2	83%	1 of 3	YES	YES	LIMITED
<b>Grade 6 (with 5 Panelists)</b>		80%				
6.1 - Number Properties and Operations	5	98%	3 of 5	YES	YES	STRONG
6.2 - Measurement	1	100%	0 of 1	YES	YES	LIMITED
6.3 - Geometry	6	66%	1 of 3	YES	YES	LIMITED
6.4 - Data Analysis and Probability	2	75%	1 of 3	YES	YES	LIMITED
6.5 - Algebraic Thinking	3	50%	2 of 3	YES	YES	STRONG
<b>Grade 7 (with 4 Panelists)</b>		64%				
7.1 - Number Properties and Operations	6	67%	3 of 5	YES	YES	STRONG
7.2 - Measurement	2	100%	0 of 1	YES	YES	LIMITED
7.3 - Geometry	6	67%	1 of 3	YES	YES	LIMITED
7.4 - Data Analysis and Probability	1	50%	1 of 2	YES	YES	STRONG
7.5 - Algebraic Thinking	0	0%	0 of 3	NO	NO	LIMITED
<b>Grade 8 (with 4 Panelists)</b>		76%				
8.1 - Number Properties and Operations	5	79%	3 of 5	YES	YES	STRONG
8.2 - Measurement	1	100%	0 of 2	YES	YES	LIMITED
8.3 - Geometry	7	98%	1 of 3	YES	YES	LIMITED

Subdomains	(Standards-to-Standards) Alignment Criteria					
	Alignment Statistics			Alignment Findings		
	Linked	Correspondence		Linked	Correspondence	
		DOK	Coverage		DOK	Coverage
8.4 - Data Analysis and Probability	2	25%	1 of 3	YES	NO	LIMITED
8.5 - Algebraic Thinking	1	67%	1 of 3	YES	YES	LIMITED
<b>Grades 9-12 (with 5 Panelists)</b>		19%				
11.1 - Number Properties and Operations	2	8%	2 of 2	YES	NO	STRONG
11.3 - Geometry	0	67%	0 of 3	NO	YES	LIMITED
11.4 - Data Analysis and Probability	1	8%	0 of 4	YES	NO	LIMITED
11.5 - Algebraic Thinking	0	50%	0 of 3	NO	YES	LIMITED

Tables 9-11, below, summarize results across the three grade clusters.

**Table 9: Summary of Alignment for the Mathematics across Grades 3-5**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Number Properties and Operations	27	65%	0	3
2 - Measurement	2	65%	0	1
3 - Geometry	24	79%	0	2
4 - Data Analysis and Probability	1	58%	0	1
5 - Algebraic Thinking	9	87%	0	2

For grades 3-5, the linkage criterion was met (even if barely so for the *Data Analysis and Probability* subdomains). DOK consistency was generally high and well above the 40% criterion (between 58% for *Data Analysis and Probability* and 87% *Algebraic Thinking*). Coverage was strong for all subdomains.

**Table 10: Summary of Alignment for the Mathematics across Grades 6-8**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Number Properties and Operations	16	81%	0	3
2 - Measurement	4	100%	0	0

3 - Geometry	19	77%	0	0
4 - Data Analysis and Probability	5	28%	0	1
5 - Algebraic Thinking	4	39%	0	1

The 6-8 grade clustered exhibited healthy linkage. For the DOK criterion, the first 3 subdomains showed strong DOK consistency (>77%), whereas *Data Analysis and Probability* and *Algebraic Thinking* did not meet the 40% criterion. Coverage was strong, except for 2 of 5 subdomains (*Measurement* and *Geometry*).

**Table 11: Summary of Alignment for the Mathematics across Grades 9-12**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Number Properties and Operations	2	8%	0	1
3 - Geometry	0	67%	0	0
4 - Data Analysis and Probability	1	8%	0	0
5 - Algebraic Thinking	0	50%	0	0

Grade 11 exhibited rather weak alignment. Linkage was not evident for two of the four subdomains (*Geometry* and *Algebraic Thinking*), while DOK consistency was met for only those two unlinked subdomains. Coverage was extremely weak.

## Science Alignment Results

Table 12 below presents the summary of the alignment of the WIDA Science model performance indicators to the KY Science standards. Again, the first set of columns present alignment statistics and the second displays alignment findings based on the criteria set forth in the previous section.

**Table 12: Alignment for Science Across Grades 3-12**

Subdomains	(Standards-to-Standards) Alignment Criteria					
	Alignment Statistics			Alignment Findings		
	Linked	Coverage		Linked	Coverage	
		DOK	e		DOK	Coverage
<b>Grade 3 (with 5 Panelists)</b>		55.00%				
3.1.1 - Structure and Transformation of Matter	11	33%	2 of 2	YES	NO	LIMITED
3.1.2 - Motion and Forces	2	0%	0 of 3	YES	NO	STRONG
3.2.3 - Earth and the Universe	11	79%	1 of 4	YES	YES	STRONG
3.3.4 - Unity and Diversity	15	94%	2 of 3	YES	YES	STRONG
3.3.5 - Biological Change	0	0%	0 of 1	NO	NO	LIMITED
3.4.6 - Energy Transformations	1	0%	1 of 4	YES	NO	LIMITED
3.4.7 - Interdependence	3	100%	1 of 1	YES	YES	STRONG
<b>Grade 4 (with 4 Panelists)</b>		24%				
4.1.1 - Structure and Transformation of Matter	10	17%	1 of 1	YES	NO	STRONG
4.1.2 - Motion and Forces	0	0%	0 of 3	NO	NO	LIMITED

Subdomains	(Standards-to-Standards) Alignment Criteria					
	Alignment Statistics			Alignment Findings		
	Linked	Correspondence		Linked	Correspondence	
			Coverage			
		DOK	e	DOK		Coverage
4.2.3 - Earth and the Universe	9	20%	0 of 4	YES	NO	LIMITED
4.3.4 - Unity and Diversity	5	20%	2 of 3	YES	NO	STRONG
4.3.5 - Biological Change	5	0%	0 of 1	YES	NO	LIMITED
4.4.6 - Energy Transformations	5	36%	1 of 5	YES	NO	MODERATE
4.4.7 - Interdependence	5	40%	1 of 2	YES	YES	STRONG
<b>Grade 5 (with 4 Panelists)</b>		59%				
5.1.1 - Structure and Transformation of Matter	6	88%	1 of 1	YES	YES	STRONG
5.1.2 - Motion and Forces	0	0%	0 of 1	NO	NO	LIMITED
5.2.3 - Earth and the Universe	14	31%	1 of 5	YES	NO	MODERATE
5.3.4 - Unity and Diversity	15	85%	1 of 2	YES	YES	STRONG
5.3.5 - Biological Change	6	62%	0 of 1	YES	YES	LIMITED
5.4.6 - Energy Transformations	0	0%	0 of 3	NO	NO	LIMITED
5.4.7 - Interdependence	3	28%	1 of 1	YES	NO	STRONG
<b>Grade 6 (with 5 Panelists)</b>		55%				
6.1.1 - Structure and Transformation of Matter	4	64%	0 of 2	YES	YES	LIMITED
6.1.2 - Motion and Forces	1	50%	0 of 1	YES	YES	LIMITED
6.2.3 - Earth and the Universe	9	63%	1 of 3	YES	YES	MODERATE
6.3.4 - Unity and Diversity	3	33%	0 of 2	YES	NO	LIMITED
6.3.5 - Biological Change	1	80%	0 of 1	YES	YES	LIMITED
6.4.6 - Energy Transformations	7	35%	2 of 2	YES	NO	STRONG
6.4.7 - Interdependence	1	89%	0 of 1	YES	YES	LIMITED
<b>Grade 7 (with 5 Panelists)</b>		32%				
7.1.1 - Structure and Transformation of Matter	5	44%	0 of 2	YES	YES	LIMITED
7.1.2 - Motion and Forces	1	0%	0 of 1	YES	NO	LIMITED
7.2.3 - Earth and the Universe	1	28%	1 of 3	YES	NO	MODERATE
7.3.4 - Unity and Diversity	1	70%	0 of 2	YES	YES	LIMITED
7.3.5 - Biological Change	0	100%	0 of 1	NO	YES	LIMITED
7.4.6 - Energy Transformations	7	17%	2 of 2	YES	NO	STRONG
7.4.7 - Interdependence	0	0%	0 of 1	NO	NO	LIMITED
<b>Grade 8 (with 5 Panelists)</b>		48%				
8.1.1 - Structure and Transformation of Matter	2	68%	0 of 2	YES	YES	LIMITED
8.1.2 - Motion and Forces	3	8%	0 of 1	YES	NO	LIMITED
8.2.3 - Earth and the Universe	1	69%	1 of 2	YES	YES	STRONG
8.3.4 - Unity and Diversity	3	43%	0 of 3	YES	YES	LIMITED
8.3.5 - Biological Change	0	33%	0 of 1	NO	NO	LIMITED
8.4.6 - Energy Transformations	7	49%	2 of 4	YES	YES	STRONG
8.4.7 - Interdependence	1	33%	0 of 2	YES	NO	LIMITED
<b>Grade 9 (with 4 Panelists)</b>		60%				
9.1.1 - Structure and Transformation of Matter	8	78%	1 of 2	YES	YES	STRONG
9.1.2 - Motion and Forces	2	19%	0 of 3	YES	NO	LIMITED

Subdomains	(Standards-to-Standards) Alignment Criteria					
	Alignment Statistics			Alignment Findings		
	Linked	Correspondence		Linked	Correspondence	
		<b>Coverage</b>				
		<b>DOK</b>	<b>e</b>		<b>DOK</b>	<b>Coverage</b>
9.2.3 - Earth and the Universe	2	50%	0 of 4	YES	YES	LIMITED
9.3.4 - Unity and Diversity	6	75%	2 of 3	YES	YES	STRONG
9.3.5 - Biological Change	3	100%	0 of 1	YES	YES	LIMITED
9.4.6 - Energy Transformations	7	29%	1 of 4	YES	NO	MODERATE
9.4.7 - Interdependence	5	58%	1 of 1	YES	YES	STRONG

The Linking criterion was generally met across grades, except for: the *Biological Change* subdomain for grades 3 and 7-8, the *Motion and Forces* subdomain for grades 4-5, the *Energy Transformations* subdomain for grade 5, and the *Interdependence* subdomain for grade 7. Results indicate that Linking was met for each grade cluster. Results for Depth were weaker than for reading or math. For grades 3, 4, 5, and 7 the 40% DOK criterion was not met for a majority of subdomains. Coverage was generally limited as well, with the early (3-5) and later (9-12) grades exhibiting the most Coverage.

Tables 13-15 summarize this alignment information across grade clusters.

**Table 13: Summary of Alignment for Science across Grades 3-5**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Structure and Transformation of Matter	27	46%	0	2
2 - Motion and Forces	2	0%	0	1
3 - Earth and the Universe	34	43%	1	1
4 - Unity and Diversity	35	66%	0	3
5 - Biological Change	11	21%	0	0
6 - Energy Transformations	6	12%	1	0
7 - Interdependence	11	56%	0	3

For grades 3-5, Linking was clearly evident in the data. The Depth criterion, however, was only met for 4 of the 7 subdomains. Coverage was adequate, except for the *Biological Change* subdomain.



**Table 14: Summary of Alignment for Science across Grades 6-8**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Structure and Transformation of Matter	11	59%	0	0
2 - Motion and Forces	5	19%	0	0
3 - Earth and the Universe	7	58%	2	1
4 - Unity and Diversity	7	49%	0	0
5 - Biological Change	1	71%	0	0
6 - Energy Transformations	21	34%	0	3
7 - Interdependence	2	41%	0	0

The linking criterion is met at Grade 6-8 for all subdomains. Overall, neither the Depth criterion nor Coverage were met (5 of 7 subdomains exhibited limited Coverage).

**Table 15: Summary of Alignment for Science across Grades 9-12**

Subdomains	Linked	DOK	Coverage	
			Moderate	Strong
1 - Structure and Transformation of Matter	8	78%	0	1
2 - Motion and Forces	2	19%	0	0
3 - Earth and the Universe	2	50%	0	0
4 - Unity and Diversity	6	75%	0	1
5 - Biological Change	3	100%	0	0
6 - Energy Transformations	7	29%	1	0
7 - Interdependence	5	58%	0	1

Summary statistics in Table 15 indicate good Linking for the 9-12 grade cluster. 5 out of the 7 subdomains met the 40% DOK consistency criterion. Coverage was moderate or strong for 4 of the 7 subdomains.

### Reliability among Reviewers

The following table shows the intraclass correlations for each grade level, which indicate the degree of agreement among reviewers in each group. Values larger than 0.7 indicate a good level of reliability among reviewers; this criterion has been met for all of the groups in this alignment study. It should, however, be noted that these values are highly dependent on the number of reviewers in each group. The pairwise comparison values represent the average agreement for pairs of reviewers in each group. A result of 0.6 or higher reflects reasonable agreement, 0.7 or higher demonstrates good agreement, and a result of less than 0.5 reflects poor agreement among reviewers.

Table 16: Reliability among Reviewers							
Grade(s)	Standards	Number of Standards	Number of Reviewers	Intraclass Correlation	DOK Pairwise Comparison	Objective Pairwise Comparison	Standard Pairwise Comparison
<b>READING</b>							
3	WIDA Language Arts Grades 3-5	25	5	0.87	0.59	0.51	0.65
4				0.86	0.65	0.46	0.62
5				0.89	0.61	0.37	0.55
6	WIDA Language Arts Grades 6-8	25	5	0.89	0.62	0.67	0.78
7				0.87	0.61	0.66	0.79
8				0.85	0.6	0.84	0.87
9-12	WIDA Language Arts Grades 9-12	25	4	0.98	0.9	0.92	0.95
<b>MATH</b>							
3	WIDA Math Grades 3-5	20	4	0.82	0.63	0.37	0.63
4				0.94	0.83	0.66	0.83
5				0.92	0.74	0.41	0.55
6	WIDA Math Grades 6-8	20	6	0.9	0.67	0.58	0.78
7			4	0.89	0.63	0.55	0.84
8			5	0.87	0.75	0.68	0.9
9-12	WIDA Math Grades 9-12	20	6	0.94	0.74	0.51	0.76
<b>SCIENCE</b>							
3	WIDA Science Grades 3-5	20	4	0.97	0.9	0.97	0.99
4				0.96	0.94	0.84	0.9
5				0.98	0.95	0.83	0.9
6	WIDA Science Grades 6-8	20	5	0.86	0.59	0.25	0.51
7				0.93	0.72	0.25	0.43
8				0.89	0.66	0.16	0.36
9-12	WIDA Science Grades 9-12	20	4	0.93	0.8	0.58	0.71

### 3. Summary

Findings from this alignment study generally suggest that there is strong *linkage* between the WIDA model performance indicators in Reading, Mathematics and Science and the KY Core Content for Assessment in Reading, Mathematics and Science. Federal

guidance on the association between ELL and state content standards directs that, at a minimum, ELL Standards must be *linked* to state academic content standards. In terms of *alignment*, the reviewers' ratings indicate that there is moderate alignment between the WIDA MPIs and the KY standards in Reading and Mathematics. This is primarily due to limited Coverage. Reviewers found only limited alignment between the WIDA Science MPIs and the KY Science standards.

## References

- Anderson, L.W. (2002). Curricular Alignment: A Re-Examination. *Theory into Practice*, 41:4, pp. 255-260.
- Blank, R.K. (2002). Using surveys of enacted curriculum to evaluate quality of instruction and alignment with standards. *Peabody Journal of Education*, 77:4, 86-121.
- Cook, H.G. (2005). Research Report #0504: Milwaukee Public Schools Alignment Study of Milwaukee Public Schools' Learning Targets in Reading and Math to Wisconsin Student Assessment System Criterion-Referenced Test Frameworks in Reading and Math. Milwaukee, WI: Milwaukee Public Schools Office of Assessment and Accountability.
- Cook, H.G. (2006). Aligning English Language Proficiency Tests to English Language Learning Standards. In *Aligning Assessment to Guide the Learning of All Students*. Council of Chief State School Officers; Washington, D.C.
- Cook, H.G. (2007). Some thoughts on English Language Proficiency Standards to Academic Content Standards Alignment. Working Draft.
- Council of Chief State School Officers. (September 2002). *Models for Alignment Analysis and Assistance to States*. Washington, D.C.: Author.
- Council of Chief State School Officers. Alignment Analysis. Retrieved July 9, 2007. Author's website: [www.ccsso.org/projects/Alignment\\_Analysis/](http://www.ccsso.org/projects/Alignment_Analysis/).
- Gottlieb, M. (2006). *Assessing English Language Learners*. Thousand Oaks, CA: Corwin Press.
- LaMarca, P.M., Redfield, D., Winter, P.C., Bailey, A., and Hansche Despriet, L. (2001). *State Standards and State Assessment Systems: A Guide to Alignment*. A study of the State Collaborative on Assessment & Student Standards (SCASS) Comprehensive Assessment Systems for ESEA Title I (CAS). Washington, D.C.: Council of Chief State School Officers.
- Porter, A.C. and Smithson, J.L. (December 2001). CPRE Research Report Series RR-048: Defining, Developing, and Using Curriculum Indicators. Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education.
- Rothman, R., Slattery, J.B., Vranek, J.L., Resnick, L.B. (2002). *Benchmarking and Alignment of Standards and Testing*, CSE Technical Report 566. Los Angeles, CA: Center for Student and Evaluation, National Center for Research on Evaluation, Standards and Student Testing.

- Scarcella, R. (2003). *Academic English: A conceptual framework*. The University of California Linguistic Minority Research Institute Technical Report 2003-1. Retrieved July 3, 2007 from [http://www.lmri.ucsb.edu/publications/03\\_scarcella.pdf](http://www.lmri.ucsb.edu/publications/03_scarcella.pdf).
- U.S. Department of Education, Office of English Language Acquisition. (February 2003). Final Non-Regulatory Guidance on the Title III State Formula Grant Program – Standards, Assessments, and Accountability. Washington, DC: U.S. Department of Education.
- Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education*. Council of Chief State School Officers and National Institute for Science Education Research Monograph No. 6. Madison: University of Wisconsin, Wisconsin Center for Education Research.
- Webb, N. L. (2002). Alignment study in language arts, mathematics, science, and social studies of state standards and assessments for four states. A study of the State Collaborative on Assessment & Student Standards (SCASS) Technical Issues in Large-Scale Assessment (TILSA). Washington, D. C.: Council of Chief State School Officers.
- WIDA Consortium (2004). WIDA English Language Proficiency Standards. Madison, WI: State of Wisconsin.

## Appendix

### *General Comments by Reviewers*

This section includes reviewer responses to the general debriefing questions listed in Part II of the WAT Training Manual as well as any generalizations or comments by the group leaders or program administrators. The following table provides a summary of these comments:

<b>Table 17: Reviewers' Perceptions of Alignment Between KY Core Content in Reading, Math, and Science and the WIDA ELP Standards</b>		
Acceptable Alignment	Needs Slight Improvement	Needs Major Improvement
6%	52%	42%
<b>Summary of Reviewer Comments by Content Area</b>		
<b>READING</b>	For the lower grade levels, all DOK levels were represented across the two sets of standards. DOK Level 1 was less represented in the standards for higher grade levels. Some of the WIDA Standards require content knowledge in addition to academic language. The KY standards emphasize fictional texts more than the WIDA standards. The two sets of standards use different verbs to depict student performance expectations.	
<b>MATH</b>	The WIDA standards mostly covered language related to geometry and number operations and did not cover as much language related to measurement and data analysis. The DOK levels for the KY standards generally appeared to be higher than the DOK levels for the WIDA standards. The WIDA standards are broader and the KY standards are specific, which made it somewhat difficult to align the two.	
<b>SCIENCE</b>	The topic of energy was covered the most within the WIDA standards. The DOK levels represented in the WIDA standards tended to be low. A challenge in aligning the standards was that the WIDA standards are broader in their objectives and the KY standards are quite specific.	

## **Reading**

### **Kentucky WIDA Alignment Grade 3**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- There were many important standards in reading that were not addressed. For example, there was nothing about characters, plot and setting, antonyms and synonyms, and sequencing.
- There were some important standards that were not assessed. Most of these fall under the umbrella of fiction: retelling a story, identifying a character's feelings, and sequencing were not addressed. These content items are: 3.3.1, 3.2.2, 3.2.3. In addition, I was not able to determine if some phonics skills were addressed, because I did not have access to the test items themselves. Only by reading the text would I know that these skills were assessed. These items are: 3.1.2, 3.1.3, 3.1.4, and 3.1.5.
- None of the "Forming a Foundation for Reading" standards were reached explicitly within the WIDA standards. It may be that the students will use these standards while trying to complete a question that addresses another standard, but the match at hand is not obvious.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- Yes.
- I believe that the DOK levels were well represented.
- Yes
- The DOK levels were low.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- I thought the standards were too vague to really work with. I sat down with the standards this summer in an attempt to make plans for the next school year and it was very difficult to do.
- A few of the items I felt could have been interpreted differently by different people. In addition, some of the items addressed content rather than language.
- Some of the standards were not as specific as they could have been; it seems that some of them could be taken in different ways depending on the person reading them. For example: Science--Associate descriptive phrases with visually supported scientific objects or terms.
- Many of the standards in Science and Social Studies require content knowledge.

**D. What is your general opinion of the alignment between the standards and assessment:**

- ii. Acceptable Alignment (1) : 20%
- iii. Needs slight improvement (3) : 60%
- iv. Needs major improvement (1) : 20%

**E. Comments**

- One serious issue is the fact that students are required to have a certain level of content knowledge to answer the questions in science, math and social studies. That makes the test a good measure of content but a poor and inaccurate measure of English language proficiency.
- Many of the standards were covered, with a strong focus in non-fiction texts. The test would match KY standards more if more items addressed skills with fictional texts. In addition, some items need to be strongly considered because a child with good language, but a lack of knowledge of the content would not be able to answer the questions. In order for this test to be a tool for teachers, it needs to focus on language, not content.
- The math section was completely wrong--it was showing the speaking domain standards instead of those in the reading domain. Multiple questions do require students to have a background knowledge of the content before they can even answer the question using one of the language domains despite the fact that this is not supposed to be a content test.
- I think there is need for more information/discussion about the relationship of DOK to ELP Standards.

**Kentucky WIDA Alignment Grade 4**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- There were many standards that were not addressed, such as character, setting and plot, synonyms and antonyms, etc.
- It covered some of the most important ones. However, much of the content is not covered at all. Much of it falls under fictional text: retelling, sequencing, identifying the character's feelings. In addition, it is impossible to tell if some of the decoding strategies are covered without seeing the test.
- Only a select few were met consistently.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- Yes.
- Yes.
- yes



**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- I felt like the standards were too vague to work with.
- Some are subject to interpretation.
- yes

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (3) : 75%
- iv. Needs major improvement (1) : 25%

**E. Comments**

- I feel that many of the questions require a certain level of content knowledge to successfully meet that standard. That is an inaccurate way to measure their level of English language proficiency.
- Some content addressed content rather than language. These need to be reviewed and changed. A teacher cannot know how her students are progressing language-wise if they are being tested on content!

**Kentucky WIDA Alignment Grade 5**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- There were many reading standards from Kentucky that were not addressed, such as character, plot and setting, synonyms and antonyms and sequence.
- There were some important topics covered. However, much content was left out that pertains to fictional text. These include: retelling, character's feelings, and sequencing. In addition, it is impossible for me to determine if phonics skills were assessed without seeing the actual test.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- Yes.
- Yes.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- I felt that the standards are too vague. There were a few instances of inappropriate material at certain grade levels.
- Some were up for interpretation.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (3) : 75%
- iv. Needs major improvement (1) : 25%

**E. Comments**

- Many of the items require a certain degree of content knowledge in order to meet the standard. I feel that we do not need this aspect in our English language proficiency test. Our students participate in KCCTS and that measures their content knowledge.
- Some items measure content rather than language. These items need to be evaluated and changed.

**Kentucky WIDA Alignment Grade 6**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- Not consistently applicable. Areas not addressed were main idea, supporting details, characterization, plot, point of view, etc. that the student will need for open response writing.
- It seemed that the same standards were applicable over and over and many of the standards did not apply. Reading content for language arts purposes (e.g., literary elements) were missing. To really work with and analyze Kentucky content standards, one must consider the Program of Studies (POS) document as well as the Core Content for Assessment. I have a sense that there would have been more of a match to the POS because speaking and listening are not assessed but are part of the POS.
- REading as language arts was minimal.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- I felt that DOK levels 2 and 3 were focused on heavily without allowing for students who can only participate in English language tasks at the DOK level 1 to participate in the assessment at all.

- Most of the standards addressed DOK levels 2 and 3. Students that are learning English should be assessed at DOK 1 as well.
- I expected more DOK 3. DOK 2 levels were more consistently addressed although not always clearly stated.
- Honestly, I needed more examples. The DOK indicated in the Kentucky content standard did not seem to match those indicated in the WIDA examples.
- Most items seemed to be DOK 1 and 2.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- 
- Yes, at times. The math, science, and social studies related areas often used verbs unrelated to Ky. standards.
- Not enough of the standards were covered. It was difficult to separate the math and science content from the language standards.
- Some were difficult to interpret. The use of the verb "classify" in mathematics was confusing.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (2) : 40%
- iv. Needs major improvement (3) : 60%

**E. Comments**

- Maybe the verbs are what made the alignment difficult. WIDA standards uses verbs like use and match. KY uses specific verbs like identify and analyze. It's hard to align standards with such discrepancies in verbs.
- I really think that the alignment needs SOME improvement.

**Kentucky WIDA Alignment Grade 7**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- No, main idea, supporting details, etc. that are need for open response and on demand writing were not comprehensively addressed.
- See Grade 6
- Same as grade 6.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- No, there were more 2 and 3 levels of DOK. Level 1 is needed by some students.
- See Grade 6
- Same as grade 6.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- The reference to other content areas is important, but the verbs did not facilitate the connection to the standard(s) correlation.
- See Grade 6
- Same as grade 6.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (2) : 40%
- iv. Needs major improvement (3) : 60%

**E. Comments**

- SOME improvement

**Kentucky WIDA Alignment Grade 8**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- A few of the standards that were applicable for 6 and 7 were supporting standards for grade 8 and caused a few more "noncodable." I found myself using the same standards over and over or stretching to make the standard fit. I think that the KDE content standards are broader in nature and while the WIDA standard may "fit" within the standard, it didn't feel right matching it to the standard.
- NO, main idea, supporting details and etc. needed for open response and on demand writing were not sufficiently covered.
- See grade 6.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- Again, the depth of knowledge for EIDA standards did not seem to match the intensity of the content standards.
- No, for DOK level 1
- See grade 6.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- Verb use was often difficult (match, differentiate, etc.)
- See grade 6.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (2) : 40%
- iv. Needs major improvement (3) : 60%

**E. Comments**

- SOME improvement

**Kentucky WIDA Alignment High School**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

**D. What is your general opinion of the alignment between the standards and assessment:**

**E. Comments**

## **Math**

### **Kentucky WIDA Alignment Grade 3**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- Too many items covered core content for geometry but lacked coverage in many areas. It did not equally represent the mathematics sub-domains. A student may score quite highly on this test but not be able to perform basic mathematical operations, solve math problems or develop a solid understanding of mathematics. Students might develop the language of "Geometry" but little else.
- No Data Analysis

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- Nearly all the items on the WIDA were very low level on the DOK (1) versus KY core content (DOK of 2 and 3) DOK levels are not a good match.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- Some of these standards were too vague. The standards were far below the grade level expectation.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (1) : 33%
- iv. Needs major improvement (2) : 67%

**E. Comments**

- Some items matched but most of the items missed the mark -
- Could use less emphasis on three-dimensional objects and add some analysis

#### **Kentucky WIDA Alignment Grade 4**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- We found that these standards lean heavily toward number sense and geometry and touch very lightly on measurement and data analysis. Estimation, multiples, factors, probabilities, lines & 2 dimensional shapes were ignored for the most part.
- estimation, multiples and factors of numbers, 2 dimensional shapes, symmetry, graphing a coordinate system, probability, function rules and output, and modeling of simple number sentences

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- There were no DOK Level 4 and very few Level 3.
- most of the items fell in DOK 1 or 2. There were few level 3 and no level 4 DOKs

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- The language was very general.
- but they were grade appropriate Kentucky standards are far more specific than WIDA standards, b

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (2) : 67%
- iv. Needs major improvement (1) : 33%

**E. Comments**

- We could cover some neglected areas and eliminate some repetition.
- need to eliminate some of the more repetitive standards and cover some of the standards that were overlooked

#### **Kentucky WIDA Alignment Grade 5**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- Items were heavy on the number sense, vocabulary, and Geometry concepts, but it lacked many of the other standards, subdomains, etc.
- The items were very heavy on the subdomains of number sense and geometry but did not seem to cover estimation, probability, and algebraic thinking.
- This section is not representative of all math sub-domains
- No, I did not see any items that addressed the areas of symmetry, factoring, and ordered pairs or those that required performing conversions or predicting probabilities.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- The majority of the items seemed to be 1, with only a few reaching the level of 2 or 3. The KY standards, however, have mostly 2 or above. The skills of application and solving problems using math concepts needed to be more evident.
- Most of the performance levels for the WIDA standards were at a level 1 while most of the Ky. standards are at a level 2. Would these standards (WIDA) prepare these students to take the Kentucky Core Content Assessment?
- lower DOK levels than expected
- Too few level 3&4

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- Most items seemed to be written below the grade level
- Many of these standards were below grade level.
- standards are too vague

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (1) : 25%
- iv. Needs major improvement (3) : 75%

**E. Comments**

- It seems difficult to align a language evaluating test to the math content area.
- A few of the standards matched but only slightly. None of the standards matched closely. This could be because we were trying to match language objectives to core content standards.



- Too much emphasis on 3 dimensional objects. 1 or 2 would have been enough. If a student can answer that many, we can assume they have that skill mastered. If not, the unproportionate number could skew the overall picture. Some of these could be eliminated to make room to assess other skills that were not included.

### **Kentucky WIDA Alignment Grade 6**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- everything was assessed but not equally so
- Somewhat - there were some missing standards such as graphing, congruency, properties, and prime and composite numbers.
- not covered were 07-1.5.1 and .2 as well as 07-3.3.1 and 07-4.2.1

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- for the most part
- Yes

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- no, this was the biggest problem; was not specific enough
- Some - some standards were very general, many interpretations are possible
- yes

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (3) : 75%
- iv. Needs major improvement (1) : 25%

### **E. Comments**

### **Kentucky WIDA Alignment Grade 7**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- everything was assessed but not equally so
- Somewhat - there were some missing standards such as graphing, congruency,

properties, and prime and composite numbers.

- not covered were 07-1.5.1 and .2 as well as 07-3.3.1 and 07-4.2.1

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- for the most part
- Yes

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- no, this was the biggest problem; was not specific enough
- Some - some standards were very general, many interpretations are possible
- yes

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (3) : 75%
- iv. Needs major improvement (1) : 25%

**E. Comments**

### **Kentucky WIDA Alignment Grade 8**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- Mostly.
- Somewhat - properties, measurement, formulas and prisms, transformation of shapes, and central tendencies were not addressed.
- didn't seem to cover coordinate geometry, transformation of shapes, characteristics of data sets, experiments and samples or probability very well

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- No 4's and a few 3's.
- Yes
- DOK 4 was not covered at all and DOK 3 was seldom used

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- A few were not, majority were.
- Some were too general - needs specificity on general problems to narrow the task.
- no, this was one of the main problems

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (2) : 50%
- iv. Needs major improvement (2) : 50%

**E. Comments**

- The generality of this test seemed to make the 8th grade test align less than the 6th and 7th grade portion did. 8th grade content seems to get much more specific. Thus, the test does not seem to align as well.

**Kentucky WIDA Alignment High School**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- No. No specific high school mathematics content was covered by these items. The items would only assess a student's ability to use mathematical language, but would not assess the student's ability to perform the specific mathematics content objectives. The items are designed to test mathematics language, not mathematics content.
- Measurement was covered more than any other.
- Measurement was covered more than any other item.
- The items presented were testing language abilities which do not match the content specific objectives to a 't'.
- The only standard that was "covered" was measurement. The others were not assessed well.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- DOK level 4 was not assessed. Most items covered DOK level 2.
- Although all were covered, I expect to see more DOK 2 because of its description.
- All were assessed, but I would expect more from DOK2.

- Yes, there seemed to be various DOK levels that would be expected.
- Yes, I would expect more DOK 2s and less DOK 3s and 1s.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- Yes. The standards are very specific, but the items were very general.
- At least one seems to relate to MS content.
- One of the standards may have been from middle school content.
- There was some confusion over the use of the word, 'function' since it was not clear if it was making reference to a mathematical operation or not.
- "specificity" I did not see specifics that were needed to make the necessary connections: content and grade level.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (1) : 20%
- iv. Needs major improvement (4) : 80%

**E. Comments**

- I found it extremely difficult and very frustrating to attempt to align these items which are based on language objectives to the standards which are based on specific content. The language objectives are clearly necessary for successful mastery of the content objectives, but language and content objectives are not the same and the language objectives are too general to align to specific content objectives.
- The process of aligning WIDA standards to KY core content standards is impossible. They are too general to align with KY's very specific core content.
- It is impossible to align WIDA standards with KCCT standards using the 20 items listed.
- If you are trying to match the WIDA standards to any content area, then they need to be more specific. However, as an ESL instructor, I think they represent a thorough range of linguistic abilities.

**Science**

**Kentucky WIDA Alignment Grade 3**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- Only 1 question addressed standard 3.4.6.
- The standard 3.4.6 was not covered well at all.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- Yes.
- There was a good sampling of the DOK levels present in the standards.
- There was a good variety of DOK

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- No, several of the standards were extremely general and did not address specific content.
- The science standards were too broad for the ky standards to match.
- No, many standards were vague and nonspecific.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (3) : 100%

**E. Comments**

**Kentucky WIDA Alignment Grade 4**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- Yes.
- Standards appear to well covered in items.
- Yes. It was a good coverage of the main concept topics.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- No, there were very few questions at a high DOK.
- The DOK was fairly low, not many questions that were a Level 3
- These questions seemed to be very low in their DOK. We found very few DOK 3 and no DOK 4.

- No, DOK questions were basically low.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- No, some of the standards were too vague.
- Some standards were very vague and applied to many different expectations.
- Still many holes in terms of Ky CC. Not much correlation between the standards. It became hard to match the WIDA to KY.
- Some standards were too vague.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (4) : 100%

**E. Comments**

**Kentucky WIDA Alignment Grade 5**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- No, 5.4.6 was not addressed. There were only 2 questions that addressed 5.4.7. The other standards were well covered.
- Standard 5.4.6 was not covered. No items matched to this standard.
- Standard 5.4.6 was not covered.
- Standard 5.4.6 not covered and 5.4.7 was weakly represented.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- Yes, DOK was well represented across levels.
- DOK was well represented
- Performance DOK levels were evenly matched.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- A few items could not be coded because they were too vague.
- Some standards were very vague and not focussed on content
- Most items were good, but some were vague and could apply to many skills.

**D. What is your general opinion of the alignment between the standards and assessment:**

- ii. Acceptable Alignment (1) : 25%
- iii. Needs slight improvement (3) : 75%

**E. Comments**

**Kentucky WIDA Alignment Grade 6**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- Motions and Forces, and Unifying Concepts were not assessed.
- yes
- The items tended to be largely language objectives and not specifically aligned with core content
- This is a difficult question to answer, since most of the 20 items were written in a way that allows them to be marginally related to the individual content standards. KY science standards are very concept specific where the WIDA standards are more process/language oriented. This makes for a difficult match to such detailed content standards.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- The DOK levels were represented well.
- Very few dok Level 3 were found
- No, there seemed to be several DOK 1 and only a few at a higher level and there seemed to be one at a 4 which for the CCA is a bit high.
- There were few items at DOK 3.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- Yes, they were level appropriate.
- yes
- no, too broad and not focused on content directly
- No- see response to question A.

**D. What is your general opinion of the alignment between the standards and assessment:**

- ii. Acceptable Alignment (1) : 25%
- iii. Needs slight improvement (1) : 25%
- iv. Needs major improvement (2) : 50%

**E. Comments**

- Kentucky's science standards are written to such specificity of content that generalized standards like those used by WIDA can only be aligned in a very general and loose way.

**Kentucky WIDA Alignment Grade 7**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- THE ONLY CONTENT AREA THAT WAS SPECIFICALLY COVERED WAS ENERGY THE REST WERE REALLY LANGUAGE OBJECTIVES
- 7.3.5.1 (fossils) were not assessed---although this topic could potentially fit into several of the WIDA standards
- Motions and Forces could be covered better. There were not very many items that addressed these standards.
- I wish that there were as many items for the other areas as there were for energy.
- yes

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- MOST MATCHING AND IDENTIFYING ARE DOK 1.... "DEFEND" IS A DOK 4
- DOK 3 items were scarce.
- The items covered a good range of DOK level questions.
- Very few DOK level 3 - no 4
- yes

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- THE STANDARD MATCHING ENERGY WAS APPROPRIATE, THERE WERE NO OTHER GOOD MATCHES.
- KY has science standards that are extremely content specific, while the WIDA standards are much more generalized--it is difficult to reconcile the difference.



- Good Grade Level specificity.
- yes

**D. What is your general opinion of the alignment between the standards and assessment:**

- ii. Acceptable Alignment (2) : 40%
- iv. Needs major improvement (3) : 60%

**E. Comments**

- See question C
- The tools items need to be addressed, these are not specifically addressed in the core content, but would be considered pre-requisites.

**Kentucky WIDA Alignment Grade 8**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- CONTENT AREA COVERED ENERGY OTHERS ARE NOT COVERED
- yes
- Life science topics (Unity, Bio Change) seem to potentially have less coverage than other topics in 8th grade.
- yes
- Motion and Forces is not addressed very thoroughly.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- THE RANGE FOR THE 8TH GRADE MORE BALANCED THAN FOR 6TH AND 7TH
- not enough of DOK level 3
- Only one DOK 3 item (in energy)
- For the most part.
- The DOK levels were all covered appropriately.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- NO, they need work
- yes
- WIDA standards are very general, while KY standards are quite specific. It is hard

to make a cast that the two are well-aligned.

- yes
- Written at an appropriate grade level.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (3) : 60%
- iv. Needs major improvement (2) : 40%

**E. Comments**

- I feel the tools aspect should be addressed. Since the tools are not in the core content, maybe it could be written differently, so that it could align more efficiently.

**Kentucky WIDA Alignment High School**

**A. For each standard, did the items cover the most important topics you expected by the standard? If not, what topics were not assessed that should have been?**

- Energy and motion were not address very fully. The space concepts were middle school concepts for KY.
- No. Reading in the Content area. If you are asking learners to read and interpret tables then this is a reading component area.

**B. For each standard, did the items cover the most important performance (DOK levels) you expected by the standard? If not, what performance was not assessed?**

- Yes.
- Yes. There was some differences in interpretation because most of us are "bloom based" in our thinking. I would suggest to color coding the thinking levels. I have a great model and use it with my teachers. Maybe doing some kind of hands on materials and then talk with members and then retalk about DOK levels.

**C. Were the standards written at an appropriate level of specificity and directed towards expectations appropriate for the grade level?**

- Certain words limited the match (structure, experiments).
- Some were a little vague. Would have like to have seen the bolded and upbolded to get the larger picture from 9 - 12. Seeing the larger picture would have helped.

**D. What is your general opinion of the alignment between the standards and assessment:**

- iii. Needs slight improvement (2) : 67%
- iv. Needs major improvement (1) : 33%

**E. Comments**

· This is a great tool to learn more about how we think about what students need to learn. I thoroughly enjoyed this! :) Thanks.

## Appendix B

### *Example of Linking and Non-Linking Standards*

The following illustrates cases of linking and non-linking of standards for mathematics at the 6-8 grade cluster.

---

---

**Kentucky's Core Content for Assessment (Mathematics) MA-06-2.1.1**

Students will measure lengths (to the nearest eighth of an inch or the nearest centimeter) and will determine and use in real-world and mathematical problems:

- area and perimeter of triangles;
  - area and perimeter of quadrilaterals (rectangles, squares); (using the Pythagorean theorem will not be required as a strategy) and
  - **area and perimeter of compound figures composed of triangles and quadrilaterals.**
- 
- 

**WIDA Level 2, Grade Cluster 6-8, Reading**

“Classify written examples supported visually of math procedures used in real world problems (such as perimeter or area)”

**State A's ELP Standard, Level 2, Grade Cluster 6-8, Reading**

“Recognize math symbols and terms”

The above two expectations are taken from actual states' ELP standards. Both are associated with mathematics; both are at level 2 and address the domain of reading. The WIDA standard focuses on classification of real world mathematics procedures, with perimeter and area given as examples. As can be seen, this standard closely associates (or links) with Kentucky's mathematics standard MA-06-2.1.1.

State A's ELP mathematics standard is addressing recognition, but this is vague and unfocused. What math symbols and terms are to be addressed? Number Operations? Addition/subtraction symbols? Geometry? The lack of clarity in this standard would make it difficult to link to any particular standard; hence, this ELP standard is NOT linked to Kentucky's content standard MA-06-2.1.1.

The goal in linking ELP standards to content expectations deals with both specificity and discourse appropriacy. The WIDA ELP example is specific and provides an appropriate discourse function: classify. While State A's ELP example does provide an appropriate language function (recognize) it is too vague.