WestEd
WestEd.org

Educational Alignment: Past, Present, and Potential Future (Part 1)

Angela D. Bowzer, Elizabeth M. Dyke

July 2023

Suggested citation: Bowzer, A & Dyke, E. (2023). Educational Alignment: Past, Present, and Potential Future (Part 1). WestEd.

WestEd is a nonpartisan, nonprofit agency that conducts and applies research, develops evidence-based solutions, and provides services and resources in the realms of education, human development, and related fields, with the end goal of improving outcomes and ensuring equity for individuals from infancy through adulthood. For more information, visit WestEd.org. For regular updates on research, free resources, solutions, and job postings from WestEd, subscribe to the E-Bulletin, our semimonthly e-newsletter, at WestEd.org/subscribe.

Contents

Abstract

Alignment of educational systems – learning standards, instruction, and assessments – is crucial to ensuring that the inferences made about what students know and can do based on evidence gathered through assessments are accurate (Cizek et al., 2018). With each reauthorization of the Elementary and Secondary Education Act (ESEA, 1965) comes changes to the components of an educational system that need to be shown to be aligned. And while assessments and standards have evolved, the most commonly used methods for assessing their alignment to one another have only modestly been adapted to meet changing needs (Polikoff, 2020). In this paper, we consider the processes and outcomes associated with various alignment methods, modifications to methods that have been made to address challenges of applying them to new assessment areas, and challenges that have yet to be addressed. While the review of literature presented here is extensive, it is not exhaustive. We note places throughout the paper where additional exploration could be beneficial for our alignment work.

Introduction

Alignment of educational systems – learning standards, instruction, and assessments – is crucial to ensuring that the inferences made about what students know and can do based on evidence gathered through assessments are accurate (Cizek et al., 2018). With each reauthorization of the Elementary and Secondary Education Act (ESEA, 1965) comes changes to the components of an educational system that need to be shown to be aligned. And while assessments and standards have evolved, the most commonly used methods for assessing their alignment to one another have only modestly been adapted to meet changing needs (Polikoff, 2020). Research does suggest that some efforts have been made toward advancing alignment methodologies, with a handful of recent studies describing new methodologies given challenges of applying existing methods to new assessments and innovative ways of conducting alignment studies.

In this paper, we consider the processes and outcomes associated with various alignment methods, modifications to methods that have been made to address challenges of applying them to new assessment areas, and challenges that have yet to be addressed. Additionally, we consider the costs – both human and financial – associated with current alignment methods and, in a future report, will consider ways that those costs might be lessened. Lastly, we will consider current WestEd alignment practices and how they might be reframed in light of concerns and innovations presented in recent research. While the review of literature presented here is extensive, it is not exhaustive. We note places throughout the paper where additional exploration could be beneficial for our alignment work.

Alignment Work at WestEd

WestEd has an established line alignment of work, with pamphlets describing the experience and expertise of staff conducting alignment studies for item-to-standards, standards-to-standards, and curriculum-to-standards alignments. WestEd has conducted item-to-standards alignment studies for more than 25 states and national organizations, many in fulfillment of federal peer review requirements. The methodology used largely addresses criteria established by Webb (1997, 2002, 2007), although the Surveys of Enacted Curriculum (SEC, Blank et al., 2001; Porter 2002) has also been used.

WestEd has conducted standards-to-standards alignment and evaluation studies for reading, mathematics, science, and social studies, as well as other academic content areas, such as the arts and humanities. In addition, WestEd has completed several standards-to-standards alignments for career and technical education (CTE) programs. The methodology used largely addresses criteria established by Achieve (2008, 2016).

WestEd has conducted alignment studies for curriculum developers in numerous content areas, including English language arts, mathematics, science, and social studies, as well as career and -

technical education (CTE) courses. The methodology used largely addresses the categorical concurrence criteria established by Webb (1997, 2002, 2007) and results in documentation of any areas of the academic content standards that are not covered by the curricula being analyzed.

WestEd's commitment to supplying rigorous and innovative alignment services provides the motivation behind the analyses of alignment literature. The subsequent sections detail the history of alignment work and current directions that have the potential to inform WestEd's alignment processes in the future.

A History of Alignment

Historically, research considering "alignment" was concerned with learning expectations and assessment items. This is reflected in Webb's (1997) definition of alignment as "the degree to which expectations and assessments are in agreement and serve in conjunction with one another to guide the system toward students learning what they are expected to know and do" (p. 3). Over time, consideration has also been given to the content of instructional materials, instructional practices, and other components key to an effective educational system, including how these components work together to achieve desired outcomes. Currently, state educational systems consist of the components of concern in the 1990s – curriculum, assessment, and instruction – but within each component, the elements that define a coherent, valid, and reliable system have been expanded to include, among other elements, those who are consulted during standards development and the application of universal design for learning (UDL) principles in assessment design. Additionally, the administration of computer-adaptive assessments that serve as accountability measures reflects the increasing role that technology plays in assessment design and administration.

Alignment was not a new area of research in the 1990s as researchers explored the alignment of the intended curriculum, the enacted curriculum, and the assessed curriculum (Cohen, 1987). However, the focus of alignment research shifted with the 1994 reauthorization of the Elementary and Secondary Education Act (ESEA, 1965) – the Improving America's Schools Act (IASA). A goal of IASA was to ensure that the academic standards that served as expectations for student learning aligned to the evidence gathered by the achievement tests taken by students. This goal was addressed through IASA's Title I statute, which required – and still requires – states to develop assessments aligned to state standards and reporting systems that connect student performance to the standards.

To meet requirements under IASA, research supported by the National Science Foundation, the National Institute for Science Education at the University of Wisconsin-Madison, the National Center for Improving Science Education, and the Council of Chief State School Officers (CCSSO) was conducted to develop a system for judging the alignment of academic standards and achievement tests (Webb, 2007). From this research, a set of criteria were developed. These

criteria, attributed to Norman Webb, comprise the most widely used alignment methods applied today.

Criteria for alignment of expectations and assessments: Webb

Norm Webb pioneered the systematizing of alignment work by developing a method to assess the degree of alignment between assessments and standards (1997, 1999). This method, referred to throughout this paper as "Webb," was developed for alignment in mathematics and science but has been adapted for use to assess alignment in other content areas (e.g., Davis-Becker & Wiley, 2017; Impara et al., 2000; Tannenbaum et al., 2015). Implementation of Webb typically involves a two-part process. The first part of the process involves a panel composed of teachers and content specialists individually identifying the degree of match between the content standards and test items. The second part of the process involves analyzing the results of the participants' decisions to determine the sufficiency of the evidence in support of the criteria under consideration (Webb & Smithson, 1999).

The original methodology was designed so that panelists addressed five general categories or dimensions: content focus, articulation across grades and ages, equity and fairness, pedagogical implications, and systems applicability to judge alignment (Webb, 1997). However, modern alignment studies often limit the criteria that are explored given the needs of the study (Martone & Sireci, 2009) and the Webb Alignment Tool (Webb et al., 2005) has been constructed as an online system into which panelists' ratings can be entered and relevant statistics can be calculated.

Currently, four content criteria are most commonly the foci of alignment studies (e.g., Webb & Smithson, 1999; Lombardi, 2006; Webb et al., 2006; Wixson et al., 2002). These criteria are described below.

- Categorical concurrence is the degree to which standards and assessments consist of the same content categories. As initially established, this criterion was met when six or more items aligned to a given standard (Webb & Smithson, 1999; Webb, 2007). This criterion has been debated in the literature and could use further exploration.
- Depth of Knowledge (DOK) consistency considers the relationship between the level of cognitive demand of an item and its associated standard. This criterion is met when the DOK of more than half of the items aligned to a given standard is at or above the DOK level of the associated standard (Webb, 2007). Some researchers adapting Webb substitute Webb's DOK model with a different measure of cognitive demand, such as Bloom's Taxonomy (e.g., Lombardi, 2006) or Cook's (2006) six-stage model of depth of knowledge (e.g., Roach et al., 2010).
- Range of knowledge correspondence addresses whether the breadth of subject matter addressed in the standards is reflected in the assessment. This criterion is met when at least 50% of the standards have at least one item aligned to each (Webb, 2007).

- Balance of representation is the degree to which there is a balance of content coverage across test items (Polikoff et al., 2011). This criterion is met when the difference in the proportion of assessed standards and the proportion of items aligned to a standard is greater than 0.7 (Li & Sireci, 2005).

Webb and adaptations of Webb were used to conduct alignments in the context of IASA and at a time when test publishers were touting their products as standards-aligned when what was really being measured by assessments were the knowledge and skills students developed in the classroom. Adaptations to existing alignment methods were needed when criteria from an original method were not applicable to the content or context being studied. For example, Wixson and colleagues (2002) determined that categorical concurrence was not an appropriate criterion for elementary reading because states do not always consistently apply categories of content across reading standards and assessments. However, they determined that consideration of the coverage of learning objectives by test items would be a way to evaluate whether each objective was measured by at least one item.

Explorations into the impact of reviewer agreement when implementing Webb have also been conducted. For example, Webb and colleagues (2006) examined how alignment results shifted when a minimum level of reviewer agreement on a match was needed for a match to be identified (e.g., agreement that an item aligned to an objective) resulted in different pictures of the alignment of standards and assessments.

Including instruction in the analysis: The Surveys of Enacted Curriculum (SEC)

Around the same time as Webb's alignment method was gaining momentum, a method that considered instruction as a component of evaluation was developed. The Surveys of Enacted Curriculum (SEC, Polikoff et al., 2011; Porter, 2002) helps to gather information about standards, assessments, and instruction (Martone & Sireci, 2009), considering the alignment between what is taught in the classroom, what is assessed, and the level of cognitive demand for relevant tasks (Roach et al., 2005). Since curricula are aligned to the standards to varying degrees and instruction impacts student opportunity to learn, exploring the relationship between curriculum, instruction, and assessments helps to contextualize student achievement (Vockley & Lang, 2009).

The instructional component of the analysis is provided through teacher self-report data about their practice and instructional materials used; therefore, accurate reporting by teachers is essential to the results of the evaluation (Porter, 2002). Each component of the analysis (curriculum, instruction, and standards) is rated by content and cognitive demand, allowing for comparison of the intended curriculum, the enacted curriculum, and the assessed curriculum (Forte, 2017). The results of the analyses can be displayed in graphs similar to heat maps, allowing for visualization of alignment strengths (see Porter, 2002). These maps can provide

information about instructional emphases and opportunity to learn in the evaluation of student achievement (Cizek et al., 2018).

A significant difference between SEC and Webb is that SEC does not have criteria that define adequate alignment (Polikoff et al., 2011). Instead, comparisons are done based on the strengths of the ratings. However, like Webb, SEC has been constructed as an online system into which ratings can be entered and analyzed.[1]

SEC has been used across content areas – in mathematics (e.g., Achieve, 2006, 2010b; Newton & Kasten, 2013), science (e.g., Porter et al., 2007), and English language arts (ELA, e.g., Atchison et al., 2022). Other methods for analyzing the relationship between learning objectives, instructional activities, instructional materials, and test items were developed at this time. For example, the Taxonomy Table (Anderson, 2002) considered student opportunity to learn the content of an assessment by determining the cognitive complexity of the factual, conceptual, procedural, and metacognitive knowledge of each component being analyzed using the Revised Bloom's Taxonomy (Anderson & Krathwohl, 2001). The Taxonomy Table, which consists of a completed table for each component under consideration, is adaptable across content areas.

Considering quality: Achieve

The Achieve methodology ([Achieve], Rothman et al., 2002) was designed specifically to provide information to states about the quality of their educational systems as a path to creating fair and balanced educational systems. The analysis considers the extent to which a test blueprint, item objectives, and test items are related to the standards. The four criteria that are the foci of Achieve alignment studies are described below.

- Content centrality is the degree of match between the content of a test question and the content of its associated standard (Forte, 2017; Rothman et al., 2002).
- Performance centrality is the degree of match between the cognitive demand of an item and the cognitive demand of its associated standard (Achieve, 2006).
- Challenge encompasses both source of challenge (i.e., where the difficulty of the item stems from) and level of challenge (i.e., performance level an item requires). Challenge issues arise when something other than the content focus of the item provides a challenge to students and when there is a lack of variety across items at the same level of difficulty (Achieve, 2006).
- Balance and range encompass the degree of match between the content emphasis of the assessment and the content standards and the representation of the breadth of the

[1] More information about SEC as an online tool can be found at the Center for Curriculum Analysis website: https://curriculumanalysis.org/products-SEC.asp

standards across the content of the assessment, respectively (Martone & Sireci, 2009; Rothman et al., 2002).

Forte (2017) describes an additional criterion for this model – accuracy of the test blueprint. To meet this criterion, each test item must be aligned to at least one content standard.

The result of the analyses is a holistic evaluation of the alignment between items and standards through summaries of evidence for each criterion (Martone & Sireci, 2009; Polikoff et al., 2011). Like SEC, Achieve does not have criteria that define adequate alignment. However, the qualitative and quantitative data that address assessment-to-standards alignment and assessment quality provide useful information for states, districts, and teachers (Martone & Sireci, 2009).

A shift in federal requirements leads to reform

IASA was passed at a time when states were not typically developing assessments aligned to their standards. Most states at the time had standards by grade band and not by grade level. It wasn't until No Child Left Behind (NCLB) – the 2002 reauthorization of ESEA – was passed that all states started to develop more specific standards that often indicated the content that would be assessed on federally-mandated assessments (Reys et al., 2005).

As federal requirements became more stringent, alignment methodologies were adjusted to meet criteria for an aligned educational system (Case & Zucker, 2005). Research during the NCLB era focused on the alignment of test items to newly developed state standards. For example, Lombardi (2006) provided a report for federal peer review that used Webb to conduct an alignment study of Minnesota assessments and standards for mathematics; Polikoff and colleagues (2011) used SEC to explore the relationships between state standards in ELA, math, and science and corresponding state assessments, considering the nature of any misalignments; and Lane (2006) studied the alignment of locally administered mathematics and reading assessments and quality assessment criteria for Nebraska, a state that utilized a school-based, teacher-led assessment and reporting system until 2018 when the state transitioned to state-level assessments.

National Assessment of Educational Progress (NAEP)

The National Assessment of Educational Progress is a congressionally mandated assessment administered to a sample of U.S. students and whose results describe what a typical student knows and can do in the subject area test administered. Under NCLB (2002), states receiving Title I funds for disadvantaged students must administer NAEP at grades 4 and 8 in ELA and mathematics. NAEP results are reported at the state level for states that choose to administer NAEP. For these states, comparisons can be made between NAEP and state assessments (Riddle, 2005).

Research on the relationship between NAEP and state assessments was occurring prior to NCLB. For example, Sanford and Fabrizio (1999) examined the framework, specifications, and test items from North Carolina grade 8 end-of-year assessments and the NAEP grade 8 assessments. This research was conducted as part of a NAEP redesign with a goal of helping states and others link their assessments with NAEP and use NAEP data to improve student outcomes. Post-NCLB, Daro and colleagues (2007) conducted an external validity study that included exploring the coverage and skill content of the NAEP framework in relation to state and international assessments, finding that, in general, NAEP content choices were comparable to those made by states and international organizations.

Research has continued to guide states on ways that they can link their assessments to NAEP. For example, Vockley and Lang (2009) describe a combination of methods (SEC, HumRRO, and the NAEP ESSI Procedural Manual) that can be used to examine the relationship between NAEP and state assessments.

As NAEP is a measure of what U.S. students typically know and are able to do, NAEP research has evolved as the political context within which NAEP is administered has shifted. For example, following the need to show that NAEP is an aligned assessment system, Blank and Smithson (2010) used SEC to examine NAEP items and corresponding item specifications. Additionally, Daro and colleagues (2015) examined the relationship between NAEP mathematics items and the Common Core State Standards (CCSS) for mathematics, considering content alignment for standards at or below the grade level assessed on NAEP. Table 1 presents information about alignment methods and documents analyzed for studies involving NAEP.[2]

Table 1. Alignment Studies Involving NAEP

Source	Alignment Method	Unit(s) of Analysis
Blank & Smithson (2010)	SEC	NAEP student assessment
Daro et al. (2015)	Item-to-Standards	NAEP mathematics item
Sanford & Fabrizio (1999)	Unique [Marzano]	Curriculum frameworks, Test items
Vockley & Lang (2009)	HumRRO; SEC; NAEP ESSI	Standards, Assessment

Text inside brackets indicates the framework used for examining cognitive complexity.

[2] Appendix A presents information about alignment methods and documents analyzed for content area focused alignment studies.

College readiness: Pre-Common Core

Alignment research during the NCLB era extended to exploring the relationship between state standards and student preparedness for college. For example, ACT (2022) linked ACT assessments to Arizona state standards and Achieve (2006) examined the content of college placement tests in relationship to Washington's mathematics standards. Alignment analyses between college entrance exams (e.g., ACT, SAT) and state standards continue to be conducted (e.g., Christopherson & Webb, 2018; Clough & Montgomery, 2015; Nemeth et al., 2016).

National Standards

The Common Core State Standards (CCSS) for English language art/literacy and mathematics were released in 2010 and subsequently adopted by 46 states. Two assessment consortia were established to develop assessments measuring the CCSS–the Smarter Balanced Assessment Consortium (SBAC) and the Partnership for Assessment of Readiness for College and Careers (PARCC)–and states that adopted the CCSS aligned themselves with one of the two consortia. The process for developing the CCSS-aligned assessments took many years, and before the first operational administration of the assessments, many states had left their originally chosen consortia.

Prior to operational administrations of the SBAC and PARCC assessments, new lines of alignment research were underway that explored the relationship between existing standards/benchmarks and the CCSS (Achieve, 2010a, 2010b; College Board, 2010). Researchers continue to explore the relationship between existing assessments and the CCSS (Achieve, 2018; Dynamic Measurement Group, Inc., 2018), the consortia assessments and the CCSS (HumRRO, 2016; Schultz et al., 2016), and the content of Advanced Placement courses and the CCSS (Hart et al., 2011). Additionally, in anticipation of a call for alignment research focused on the CCSS, Newton and Katsten (2013) analyzed standards and assessments using Webb and SEC, comparing the results from each.

Even as states have moved away from the 2010 CCSS as their state standards, many standards documents are still closely aligned to the CCSS. Assessment vendors who want to present their products to states as aligned to state standards, and thereby for use by states to meet their assessment needs, have brought external evaluators in to examine the alignment of standards and test items (e.g., edMetric, LLC., 2020). Researchers are also exploring the relationship between the CCSS and international assessments (e.g., PISA, OECD, 2013), and Porter and colleagues (2011) used SEC to compare the CCSS to existing state standards, international standards, state assessments, and teacher reports of practice.

The Next Generation Science Standards (NGSS) were released in 2013. Dickinson and colleagues (2020) adapted Webb to examine the relationship between test items and the NGSS. Results from the study suggest that the California Science Assessment (CAST) meets Link to Standards, Range Adequacy, and Multidimensional Adequacy criteria at each assessed grade level, but only

partially meets Depth of Knowledge Adequacy criteria at grade 8 and high school and Balance of Knowledge criteria at grade 8.

Fulmer and colleagues (2018) noted that the foci of science alignment studies in the NGSS-era vary (e.g., performance expectations, dimensions, adapted rubrics and frameworks) making it difficult to compare curriculum materials and science assessments. A move to aid the evaluation of instructional materials stems from a joint effort by WestEd, BSCS Science Learning, and Achieve, who developed an online professional learning program to support science materials evaluation.[3]

The Current Political Climate

Along with the passing of IASA came the first guidelines for establishing alignment within an educational system, which have evolved to meet changes to requirements by NCLB and ESSA. Currently, the 2015 reauthorization of ESEA – the Every Student Succeeds Act (ESSA) – requires each state to administer: (1) reading/language arts and mathematics assessments and corresponding alternate assessments based on alternate academic achievement standards (AA-AAS) annually in grades 3-8 and once in high school; (2) science assessments and corresponding AA-AAS at least once in each of these grade spans: 3-5, 6-9, and 10- 12; and (3) English language proficiency (ELP) assessments annually in grades K-12 (USDoE, 2018).

As part of the requirements of ESSA, states are required to undergo federal peer review– "the process through which a State demonstrates the technical soundness of its assessment system" (USDoE, 2018, p. 4). Submissions for federal peer review must include a report showing the alignment of tests to their test blueprints and evidence that test blueprints reflect the depth and breadth of the associated standards. For CAT assessments, alignment reports also need to address the item pool's sufficiency to meet the test blueprint, the balance of content represented by the item pool, and the cognitive complexity and range of item difficulty for each standard assessed.

Apart from the required assessments listed above, other assessments to be submitted for peer review, should states choose to administer them, are: (1) alternate English language proficiency assessments (AELPAs) for English learners (ELs) who are students with the most significant cognitive disabilities in grades K-12; (2) locally selected, nationally recognized high school academic assessments, (3) assessments used for 8th-grade mathematics students taking a high school credit-bearing course, (4) assessments in an EL student's native language, and (5) assessments in a Native American language. Additionally, states must show that their standards align with entry-level state public college coursework requirements and relevant state career and technical education standards.

[3] See Virtual NextGen Time: https://www.nextgentimepl.org/home

The established methodologies of Webb (1997), SEC (Polikoff et al., 2011; Porter, 2002), and Achieve (Rothman et al., 2002) have been recommended by the CCSSO (2002) for use in alignment studies.[4] However, these early and most frequently used methods for determining the alignment of educational systems were not developed to account for the variety of assessments required to be submitted for federal peer review. In the section that follows, we describe in greater detail adapted methods and new methods developed to meet the needs of states submitting their assessments for federal peer review.

New Methods

Newer alignment methods developed independently of existing methods have been developed and used to explore the alignment of modern educational systems (e.g., Cizek et al., 2018; Davis-Becker & Buckendahl, 2013; Kulm et al., 2005; Peltier et al., 2021). Additionally, existing methods continue to be adapted to show alignment of educational systems developed for special populations of students.

Students with significant disabilities

Federal legislation impacts all students, including those with the most significant disabilities, as ESSA requires the assessment of all students. However, students with significant disabilities (SWSD) are not properly assessed with traditional assessment methods (e.g., pencil and paper examinations). Addressing the need to fairly assess SWSD led to the development of alternate assessments–assessments with accommodations appropriate to a test taker's individual needs–as a means of school reform and increased accountability (Flowers et al., 2007; Roach et al., 2005). Alternate assessments are more flexible than large-scale assessments, allowing for greater opportunities to capture evidence of the knowledge and skills of SWSD (Flowers et al., 2007). For example, an alternate assessment may be a modified version of the general education assessment, data collected from observations, or examples of student performance over time (Thompson et al., 2005; Tindal, 2005).

Though these tests assess the same grade level content as general education assessments, alternate assessments align to alternate achievement standards (AA-AAS), which present expectations at a lower level of complexity than the general education standards. The characteristics of alternate assessments require additional considerations when assessing alignment. Consequently, alignment methodologies have been developed to specifically address the alignment between alternate assessments and content standards (Flowers et al., 2007).

[4] A fourth alignment method – Council for Basic Education (CBE) –was also recommended by the CCSSO (2002). However, a detailed description of the CBE method could not be located.

Links for Academic Learning (LAL) is an alignment methodology designed for use with alternate assessments and alternate content standards. This method was developed to assess the degree of match between alternate assessments and their corresponding standards and to address gaps that existed when applying traditional alignment methods given the unique characteristics of SWSD and AA-AAS (Flowers et al., 2007).

LAL is an adaptation of Webb and SEC that examines criteria from Webb (e.g., DOK, range and balance) and employs the Curriculum Indicators Survey (CIS), which is a modified version of SEC that collects information about teacher- and school-level factors that impact student learning. Additional foci of analyses include equity, fairness, impact of pedagogical choices, and differentiation. The result of an LAL alignment study is detailed information on each criterion that describes the alignment of alternate assessments, alternate content standards, and teacher instructional practice (Flowers et al., 2007).

Researchers have used LAL by itself or in conjunction with another alignment method to explore the relationship between alternate assessments, alternate achievement standards, and even the CCSS. Table 2 presents information about alignment methods and documents analyzed for studies involving AA-AAS.

Table 2. Alignment Studies Involving AA-AAS

Source	Alignment Method	Unit(s) of Analysis
Nemeth et al. (2011)	LAL	Test items, alternate content standards
Peltier et al. (2021)	Unique, but includes LAL	Assessment items, extended content standards
Roach et al. (2005)	Webb	Alternate assessment, content standards
Thacker et al. (2018)	Adapted Webb and LAL	Alternate achievement standards, CCSS, test items
Tindal (2005)	Webb	Performance tasks, portfolios, alternate standards

English language learners

Researchers have used Webb to explore the alignment between English Language Proficiency (ELP) standards and corresponding assessments. However, the DOK criteria from Webb cannot be applied during the alignment process due to the linguistic complexity associated with ELP standards. CCSSO (2012) describes this complexity as a connection between the content of language standards and cognitive complexity. To address the "linguistic hierarchy that forms the foundation of ELP standards" (Johnson, 2005, p. 3), DOK is often replaced by an examination of Linguistic Difficulty Levels (LDLs, Cook, 2006), which correspond to the ways that non-native speakers of English acquire language. Table 3 presents information about alignment methods and documents analyzed for studies involving ELP standards.

Table 3. Alignment Studies Involving ELP Standards

Source	Alignment Method	Unit(s) of Analysis
Johnson (2005)	Adapted Webb [Linguistic Difficulty Levels]	ELP standards and assessments, content standards
Kaplan (2016)	Unique - Descriptive	World language standards and assessments
Papageorgiou et al. (2020)	Webb	Language assessments (TOEFL Primary tests), English as a foreign language (EFL) curricula
Thacker et al. (2021)	Adapted Webb [Linguistic Difficulty Levels]	California ELPA Standards and assessments

Text inside brackets indicates the framework used for examining cognitive complexity.

As ESSA requires that state ELP standards have comparable language demands to the state mathematics and science standards (USDoE, 2018), researchers have examined the link between ELP standards and general education standards (e.g., Bailey et al. 2022; CCSSO, 2012; Christopherson & Webb, 2015; WestEd, 2015). Support in the development and review of ELP standards and assessments has been provided by the Assessment and Accountability Comprehensive Center (2009) which used research on quality standards and assessments in general, and research on how students learn English as a second language specifically, to create a framework that can be implemented at any point in the development or review process.

Early Childhood

Achieving greater alignment in education systems, including early childhood systems, is important for promoting continuity across educational settings (Shuey et al., 2019). Increasingly, efforts have been focused on including early childhood education in accountability initiatives (Roach et al., 2010). The National Association for the Education of Young Children (NAEYC) and the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE, 2003) recommend that early childhood education systems be comprehensive systems of curriculum and assessment supported by appropriate learning standards and program standards that improve young children's school readiness. However, early childhood educational systems are often misaligned (Whitaker et al., 2022), presenting a great need for examples of aligned systems and methods for aligning systems.

As with the variability of early childhood educational systems, there is not a common method for examining the alignment of an early childhood system. While Roach and colleagues (2010) adapted Webb to examine the relationship between early childhood assessments and kindergarten standards, Davidson and Egan (2020) adapted LAL to show strong alignment between components of a state-level early childhood educational system. Additionally, Litkowski and colleagues (2020) used statistical methods to examine the alignment between preschool students' numeracy performance in relation to CCSS and state-level early learning standards. However, the results of this study suggested misalignment between evidence of student knowledge and skills and those measured by the standards. See Table 4 for more information on these studies.

Table 4. Alignment Studies Involving Early Childhood Standards

Source	Alignment Method	Unit(s) of Analysis
Davidson & Egan (2020)	LAL	Test items (SKBs, SKB Levels), standards (Foundations)
Litkowski et al. (2020)	Unique - Statistical analyses	Standards, measures of counting and cardinality, numeracy subtests
Roach et al. (2010)	Adapted Webb [Cook's]	Early childhood assessment, state kindergarten standards

Text inside brackets indicates the framework used for examining cognitive complexity.

Other Alignment Foci

Alignment is important to the coherence of an educational system (Vockley & Lang, 2009), with existing, modified, and new alignment procedures applied to a multitude of systems. The following list shows a variety of contexts in which alignment has been studied.

- State high school assessments to standards for success in entry-level university courses (Brown & Conley, 2007)
- The Praxis Performance Assessment for Teachers with the Interstate Teacher Assessment and Support Consortium (InTASC) model core teaching standards (Reese et al., 2015)
- World language standards and assessments (Kaplan, 2016)
- Content of state assessments to each other through the CCSS (Blank & Adams, 2018)
- The Praxis Performance Assessment for Teachers with Praxis content knowledge test specifications for social studies (Educational Testing Service [ETS], 2019)
- Learning objectives and assessments in college courses (Bae et al. (2019)
- Curricula to Revised Bloom's (Yilmaz & Oner Sunker, 2021)

Challenges

While work is being done to ensure that the requirements of peer review can be met and that educational systems are valid and coherent, there are still challenges associated with conducting alignment studies that have yet to be overcome, some of which cannot be overcome due to constraints in the political processes associated with alignment. In the subsections that follow, we present several challenges posed by researchers.

Changing education landscape

Beginning with Webb and continuing through today, the policy contexts within which alignment of educational systems are examined is ever changing. Federal requirements frequently shift, requiring a change in the alignment review process or a component of the educational system needing to be reviewed. State policies also shift, with some states requiring rigorous high school exit exams and others administering less rigorous tests of college and career readiness (Ananda, 2003). Additionally, advancements in our understandings of how people learn, expansion of content areas being taught or assessed, and the evolution of technology–both as a tool for learning and assessment and as a tool for analyzing assessment data–have changed the processes involved in evaluating alignment (Fulmer et al., 2018; Webb 1997; Wixson et al., 2002).

Interpretation of policies

The lack of a common language when talking about or implementing policies poses a challenge to the evaluation of educational systems. At the most basic level, interpretation of standards is often inconsistent and the operationalization of the content of the standards in assessments often differs across systems. (Martone & Sireci, 2009; Rothman et al., 2002; Webb, 1997; Wixson et al., 2002). Complicating the review of systems is a lack of clarity on what is needed to meet federal peer review guidelines. For example, ESSA requires that assessments "address the depth and breadth" of the corresponding state standards (USDoE, 2018, p. 24), and while some clarification of what is meant by "depth and breadth" is provided, greater specificity is needed to guide the evaluation process.

Limitations of evidence

While ESSA requires assessment to "cover all of the knowledge and skills [of the standards] over a period of time" (USDoE, 2018, p. 25), some content standards are not easily measured, if at all measurable, forcing states to make policy decisions about what to include in their standards and on their assessments (Ananda, 2003; Martone & Sireci, 2009; Resnick & Zurawsky, 2002; Way & Croft, 2020). Researchers tend to agree that "[i]t is improbable that a single assessment instrument will provide the breadth of coverage necessary for an aligned system" (La Marca et al., 2000, p. 18). Others suggest the need to supplement traditional alignment evidence with information about the design and development of the educational system (Wixson et al., 2002). Additionally, shifting test forms might result in some test forms that are more aligned to the standards than others (Fulmer, 2010), leading to additional challenges, such as a need to consider how many test forms to include in an evaluation to achieve the level of alignment necessary to meet federal peer review.

Study designs and alignment methodologies

In general, the design of an alignment study and implementation of the design is challenging. Differences in design can lead to different results, which is especially concerning when the same assessment is being evaluated using different alignment methodologies (Herman et al., 2005; Kaira, 2010; Polikoff, 2020; Way & Croft, 2020; Wixson et al., 2002). Additionally, changes in the testing landscape (e.g., computer-adaptive testing) present a need for new methods to capture the full picture of alignment (Gallagher, 2016).

Components under review

Frequently, several documents need to be reviewed to gather a complete picture of the educational system under evaluation (Webb, 1997; Wixson et al., 2002). The complexities of some components (e.g., language in ELP assessments, Standards for Mathematical Practice) call for new and more complex review processes (Organisation for Economic Co-operation and Development [OECD], 2013; Porter et al., 2007). Additionally, the quality of these documents

influences the result of the analysis, and results showing a system is aligned do not equate to the quality of the system (Newton & Kasten, 2013; Porter et al., 2007).

The human factor

Among the challenges associated with alignment studies presented in the literature, the most commonly mentioned, and likely the most costly, is the involvement of humans in the evaluation process. Current methodologies require educational judgments, necessitating training and calibration to support a common understanding of the processes and procedures by participants. The knowledge and skills of the participants are just as important as the design of the methodology used during the study in achieving consistent results (Case et al., 2004; Herman et al., 2005; Martone & Sireci, 2009; Polikoff, 2020; Tannenbaum et al., 2015; Webb, 1997). Additionally, the complexity of some methods results in a great deal of time and cognitive effort on the part of participants (Kaira, 2010; Porter, 2002).

Innovations

Some progress has been made to support the development of an aligned system and expand the ways that alignment studies are conducted or that evaluation of alignment studies occurs. While further examination into innovations is needed, and more detailed descriptions of methods are needed here, several methods suggested for use are listed below.

- Matching standards to student performance (Kaira, 2010)
- Automated classifications (Sheehan et al., 2010)
- Automated construction of aligned assessments (Porter et al., 2013)
- Using machine learning to investigate textual congruence (Anderson et al., 2020)
- Automated data mining or text scraping (Polikoff, 2020)

Next Steps

Forte (2017) reminds us that, as researchers, the lens through which we view alignment should not be the peer review process but The Standards for Educational and Psychological Testing (The Standards; AERA/APA/NCME, 2014). The Standards define alignment as "[t]he degree to which the content and cognitive demands of test questions match targeted content and cognitive demands described in the test specifications" (p. 216). In the next part of our review, we will explore what it means to take a Standards view of alignment, including how a Standards lens has the potential to address issues of cognitive complexity in the alignment process and how a Standards lens might support a coherent alignment methodology that is defensible while also being efficient.

References

Achieve, Inc. (2006). *An alignment analysis of Washington State's college readiness mathematics standards with various local placement tests.*

Achieve, Inc. (2008). *Out of many, one: Toward rigorous Common Core Standards from the ground up.*

Achieve, Inc. (2010a). *A comparison of the Common Core State Standards for mathematics to the American Diploma Project Benchmarks for mathematics.*

Achieve, Inc. (2010b). *Comparing the Common Core Standards in Mathematics to California and Massachusetts standards.*

Achieve, Inc. (2016). *A review of the Oklahoma January 2016 English Language Arts and Mathematics Academic Standards.*

Achieve, Inc. (2018). *Independent analysis of the alignment of the ACT to the Common Core State Standards.*

ACT. (2022). *Linking ACT assessments and Arizona standards to drive student success*.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* American Educational Research Association.

Ananda, S. (2003). *Rethinking issues of alignment under No Child Left Behind.* WestEd.

Anderson, L. W. (2002). Curricular Alignment: A re-examination. *Theory into Practice*, *41*(4), 255-260.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Anderson, D., Rowley, B., Stegenga, S., Irvin, P. S., & Rosenberg, J. M. (2020). Evaluating content-related validity evidence using a text-based machine learning procedure. *Education Measurement: Issues and Practice, 39*(4), 53-64.

Assessment and Accountability Comprehensive Center. (2009). *Framework for high-quality English Language Proficiency standards and assessments*. WestEd.

Atchison, D., Garet, M. S., Smith, T. M., & Song, M. (2022). The validity of measures of instructional alignment with state standards based on Surveys of Enacted Curriculum. *AERA Open*, *8*(1), 1-17.

Bae, Y., Tunstall, S. L., Knowles, K. S., & Matz, R. L. (2019). Alignment between learning objectives and assessments in a quantitative literacy course. *Numeracy: Advancing Education in Quantitative Literacy*, *12*(2). https://doi.org/10.5038/1936-4660.12.2.10

Bailey, A. L., Wolf, M. K., & Ballard, L. (2022). *English language proficiency standards and assessment alignment: Issues, insights, and innovations for guiding peer review*. Educational Testing Service.

Blank, R. K., & Adams, E. (2018). *How well aligned are state assessments with state standards or Common Core/Next Generation science standards?* NORC.

Blank, R. K., Porter, A., & Smithson, J. (2001). *New tools for analyzing teaching, curriculum and standards in mathematics & science.* Council of Chief State School Officers.

Blank, R. K., & Smithson, J. (2010). *Alignment Content Analysis of NAEP 2009 Science Assessments with the methodology of Surveys of Enacted Curriculum.* Council of Chief State School Officers.

Brown, R. S., & Conley, D. T. (2007). Comparing state high school assessments to standards for success in entry-level university courses. *Educational Assessment, 12*(2), 137-160. https://doi.org/10.1080/10627190701232811

Case, B., & Zucker, S. (2005, July). *Methodologies for alignment of standards and assessments* [Paper presentation]. Conference on Alignment of Assessments and Instruction, Beijing, China.

Case, B. J., Jorgensen, M. A., & Zucker, S. (2004). *Alignment in educational assessment.* Pearson Education, Inc.

Center for Assessment. (2016). *Guide to evaluating assessments using the CCSSO criteria for high quality assessments: Focus on test content*. https://www.nciea.org/wp-content/uploads/2021/11/CFA-Guide-FocusOnTestContent-R1_0.pdf

Christopherson, S. C., & Webb, N. L. (2015). *Final report: Alignment analysis of key practice language functions from the framework for English Language Proficiency Development Standards corresponding to the Common Core State Standards for English Language Arts and Mathematics and the WIDA English Language Proficiency Standards.* Wisconsin Center for Education Products & Services.

Christopherson, S. C., & Webb, N. L. (2018). *Alignment analysis of the ACT and SAT with the Georgia standards of excellence for American literature and composition, Algebra I, geometry, and biology.* Wisconsin Center for Education Products & Services.

Cizek, G. J., Kosh, A. E., & Toutkoushian, E. K. (2018). Gathering and evaluating validity evidence: The generalized assessment alignment tool. *Journal of Educational Measurement, 55*(4), 477-512.

Clough, S., & Montgomery, S. (2015). *How ACT assessments align with state college and career readiness standards.* ACT.

Cohen, S. A. (1987). Instructional alignment: Searching for a magic bullet. *Educational Researcher, 16*(8), 16-20.

Cook, H. G. (2006). *Aligning English language proficiency test to English language learning standards. Aligning assessment to guide the learning of all students*. Council of Chief State School Officers.

Council of Chief State School Officers. (2002). *Models for alignment analysis and assistance to states*.

Council of Chief State School Officers. (2012). *Framework for English Language Proficiency Development Standards corresponding to the Common Core State Standards and the Next Generation Science Standards*.

Council of Chief State School Officers. (2014). *Criteria for procuring and evaluating high-quality assessments*.

D'Agostino, J. V., Welsh, M. E., Cimetta, A. D., Falco, L. D., Smith, S., VanWinkle, W. H., & Powers, S. (2008). The rating and matching item-objective alignment methods. *Applied Measurement in Education, 21,* 1–21.

Daro, P., Hughes, G. B., & Stancavage, F. (2015). *Study of the alignment of the 2015 NAEP mathematics items at grades 4 and 8 to the Common Core State Standards (CCSS) for mathematics*. American Institutes for Research.

Daro, P., Stancavage, F., Ortega, M., DeStefano, L., & Linn, R. (2007). *Validity study of the NAEP mathematics assessment: Grades 4 and 8.* American Institutes for Research.

Davidson, A. H., & Egan, K. (2020). *Indiana Student Performance Readiness and Observation of Understanding Tool (ISPROUT) Early Learning Assessment*. EdMetric.

Davis-Becker, S. L., & Buckendahl, C. W. (2013). A proposed framework for evaluating alignment studies. *Educational Measurement: Issues and Practice, 32*(1), 23-33.

Davis-Becker, S., & Wiley, A. (2017). *Alignment of the ACT to the Wisconsin Academic Standards in ELA and mathematics: Final report*. ACS Ventures.

Dickinson, E., Thacker, A., & Hardoin, M. (2020). *California Assessment of Student Performance and Progress (CAASPP): California Science Test (CAST) alignment study report.* Human Resources Research Organization.

Doorey, N., & Polikoff, M. (2016). *Evaluating the content and quality of next generation assessments*. Thomas B. Fordham Institute.

Dynamic Measurement Group, Inc. (2018). *How Acadience Reading aligns with the Common Core State Standards in English Language Arts*.

edMetric, LLC. (2020). *Evaluation of the alignment of Transcend item pool to Common Core State Standards*.

Educational Testing Service. (2019). *PPAT assessment: Alignment with Praxis test specifications for social studies: Content knowledge*.

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

Forte, E. (2017). *Evaluating alignment in large-scale standards-based assessment systems.* Council of Chief State School Officers.

Forte, E., Towles, E., Greninger, E., Buchanan, E., & Deters, L. (2017). *Evaluation of the alignment quality in the Georgia Milestones Assessment System in ELA, mathematics, science, and social studies.* edCount, LLC.

Fulmer, G. W. (2010). *Estimating critical values for strength of alignment among curriculum, assessments, and instruction* [Paper presentation]. AERA, Denver, CO, United States.

Fulmer, G. W., Tanas, J., & Weiss, K. A. (2018). The challenges of alignment for the Next Generation Science Standards. *Journal of Research in Science Teaching, 55,* 1076-1100. https://doi.org/10.1002/tea.21481

Gallagher, C. (2016). *Evaluating content alignment in the context of computer-adaptive testing: Guidance for state educational agencies*. WestEd: The Center on Standards & Assessment Implementation.

Hart, B., Carman, E., Luisier, D., Vasavada, N. (2011). *Common Core State Standards alignment: Advanced Placement* (Research Report 2011-8). College Board.

Herman, J., L., Webb, N. M., & Zuniga, S., A. (2005). *Measurement issues in the alignment of standards and assessments: A case study* (CSE Report 653). National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Human Resources Research Organization. (2016). *Smarter Balanced Consortium: Alignment study report.*

Impara, J. C., Plake, B. S., & Beckendahl, C. W. (2000, June). *The comparability of norm-referenced achievement tests as they align to Nebraska's language arts content standards* [Paper presentation] Large Scale Assessment Conference, Snowbird, UT.

Iowa Testing Program. (2017). *Overview of the Iowa tests. -to-the* https://www.centralriversaea.org/wp-content/uploads/2017/03/E2_Overview-of-the-Iowa-Tests-Alignment -Common-Core-Revised-6.8.17.pdf

Johnson, D. F. (2005). *Aligning ELP assessments to ELP standards*. Harcourt Assessment, Inc.

Kaira, L. T. (2010). *Using item mapping to evaluate alignment between curriculum and assessment* (Publication No. 318) [Doctoral dissertation, University of Massachusetts Amherst]. Open Access Dissertations. https://scholarworks.umass.edu/open_access_dissertations/318?utm_source=scholarworks.umass.edu%2Fopen_access_dissertations%2F318&utm_medium=PDF&utm_campaign=PDFCoverPages

Kaplan, C. S. (2016). Alignment of world language standards and assessments: A multiple case study. *Foreign Language Annals*, *49*(3), 502-529.

Kulm, G., Wilson, L. D., & Kitchen, R. (2005). Alignment of content and effectiveness of mathematics assessment items. *Educational Assessment*, *10*(4), 333-356. http://dx.doi.org/10.1207/s15326977ea1004_2

La Marca, P. M., Redfield, D., & Winter, P. C. (2000). *State standards and state assessment systems: A guide to alignment. Series on standards and assessments.* Council of Chief State School Officers.

Lane, S. (2006). *Evaluation of the six quality assessment criteria used in the Nebraska school-based teacher-led assessment and reporting System (STARS)*. Nebraska Department of Education.

Li, S., & Sireci, S. G. (2005). *Evaluating the fit between test content, instruction, and curriculum frameworks: A review of methods for evaluating test alignment* (MCAS Validity Report No. 9). Center for Educational Assessment.

Litkowski, E. C., Duncan, R. J., Logan, J. A. R, & Purpura, D. J. (2020). Alignment between children's numeracy performance, the kindergarten Common Core State Standards for Mathematics, and state-level early learning standards. *AERA Open, 6*(4), 1–11.

Lombardi, T. J. (2006). *An alignment study of the Minnesota Comprehensive Assessment-II with state standards in mathematics for grades 3-8 and 11*. Minnesota Department of Education.

Lopez, A. A. (2013). Alignment between standardized assessments and academic standards: The case of the Saber Mathematics Test in Colombia. *RELIEVE*, *9*(2), doi: 10.7203/relieve.19.2.3026

Martone, A., & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment, and instruction. *Review of Educational Research, 79*(4), 1332-1361. https://doi.org/10.3102%2F0034654309341375

National Association for the Education of Young Children (NAEYC), & the National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE). (2004). *Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8.*

Nemeth, Y., Michaels, H., Wiley, C., & Chen, J. (2016). *Delaware system of student assessment and Main comprehensive assessment system: SAT alignment to the Common Core State Standards.* Human Resources Research Organization (HumRRO).

Nemeth, Y. M., Thacker, A., Deatz, R., Buckland, W., Fry, S., Hardoin, M., & Wiley, C. (2011). *Independent Alignment Review of the Mathematics Minnesota Test of Academic Skills (MTAS)*. Human Resources Research Organization (HumRRO).

Newton, J. A., & Kasten, S. E. (2013). Two models for evaluating alignment of state standards and assessments: Competing or complementary perspectives? *Journal for Research in Mathematics Education, 44*(3), 550-581. https://doi.org/10.5951/jresematheduc.44.3.0550

Organisation for Economic Co-operation and Development. (2013). *Lessons from PISA 2012 for the United States.* OECD Publishing. http://dx.doi.org/10.1787/9789264207585-en

Papageorgiou, S., Xu, X., Timpe-Laughlin, V., & Dugdale, D. M. (2020). *Exploring the alignment between a curriculum and a test for young learners of English as a foreign language* (Research Memorandum No. RM-20-08). Educational Testing Service.

Peltier, H., Snipes, A., Konort, K., & Forte, E. (2021). *Alignment evaluation for the GAA 2.0 assessments – Executive summary.* edCount, LLC.

Pemberton, J. B., Tyler-Wood, T., & Cereijo, V. P. (2006). Aligning assessments with state curriculum standards and teaching strategies. *Intervention and School Clinic, 41*(5), 283-289. https://doi.org/10.1177%2F10534512060410050501

Polikoff, M. S. (2020). The present and future of alignment. *Educational Measurement: Issues and Practice, 39*(2), 18-20.

Polikoff, M. S., & Porter, A. C. (2014). Instructional alignment as a measure of teaching quality. *Educational Evaluation and Policy Analysis, 36*(4), 399-416.

Polikoff, M. S., Porter, A. C., & Smithson, J. (2011). How well aligned are state assessments of student achievement with state content standards? *American Educational Research Journal*, *48*(4), 965-995. https://doi.org/10.3102%2F0002831211410684

Polikoff, M. S., Zhou, N., & Campbell, S. E. (2015). Methodological choices in the content analysis of textbooks for measuring alignment with standards. *Educational Measurement: Issues and Practice*, *34*(3), 10-17.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, *31*(7), 3-14.

Porter, A. C., McMaken, J., Hwang, J., & Yang, R. (2011). Common Core State Standards: The new U.S. intended curriculum. *Educational Researcher, 40*(3), 103-116. http://dx.doi.org/10.3102/0013189X11405038

Porter, A., Polikoff, M. S., Barghaus, K. M., & Yang, R. (2013). Constructing aligned assessments using automated test construction. *Educational Researcher, 42*(8), 415-423. https://doi.org/10.3102%2F0013189X13503038

Porter, A. C., & Smithson, J. L. (2002, March 7). *Alignment of assessments, standards and instruction using curriculum indicator data* [Conference presentation]. AERA, New Orleans, LA, United States. https://curriculumanalysis.org/Reference/PorterSmithson-AlignmentNCME02.pdf

Porter, A. C., Smithson, J., Blank, R., & Zeidner, T. (2007). Alignment as a teacher variable. *Applied Measurement in Education*, *20*(1), 27-51. https://doi.org/10.1080/08957340709336729

Reese, C. M., Tannenbaum, R. J., & Kuku, B. (2015). *Alignment between the Praxis® Performance Assessment for Teachers (PPAT) and the Interstate Teacher Assessment and Support Consortium (InTASC) model core teaching standards* (Research Memorandum No. RM-15-10). Educational Testing Service.

Renaissance Learning. (2011). *STAR Math Enterprise: Tennessee standards alignment: Grades three through twelve.*

Resnick, L. B., & Zurawsky, C. (2002). Standards and tests: Keeping them aligned. *Essential Information for Education Policy, 1*(1), 1-4.

Reys, B., Dingman, S. W., Sutter, A., & Teuscher, D. (2005). *Development of state-level mathematics curriculum documents: Report of a survey*. Center for the Study of Mathematics Curriculum.

Riddle, W. C. (2005). *Educational testing: Implementation of ESEA Title I-A requirements under the No Child Left Behind Act* (CRS Report for Congress). The Library of Congress.

Roach, A. T., Elliott, S. N., & Webb, N. L. (2005). *Alignment of an alternate assessment with state academic standards: Evidence for the content validity of the Wisconsin alternate assessment. The Journal of Special Education*, *38*(4), 218-231. https://doi.org/10.1177%2F00224669050380040301

Roach, A. T., McGrath, D., Wixson, C., & Talapatra, D. (2010). Aligning an early childhood assessment to state kindergarten content standards: Application of a nationally recognized alignment framework. *Educational Measurement: Issues and Practice*, *29*(1), 25-37.

Rothman, R., Slattery, J. B., Vranek, J. L., & Resnick, L. B. (2002). *Benchmarking and alignment of standards and testing* (CSE Technical Report 566). National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Sanford, E. E., & Fabrizio, L. M. (1999, April). *Results from the North Carolina--NAEP Comparison and what they mean to the end-of-grade testing program* [Paper presentation] Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Schultz, S. R., Michaels, H. R., Dvorak, R. N., & Wiley, C. R. H. (2016). *Evaluating the content and quality of next generation high school assessments*. Human Resources Research Organization.

Sheehan, K. M., Kostin, I., Futagi, Y., & Flor, M. (2010). *Generating automated text complexity classifications that are aligned with targeted text complexity standards* (ETS RR-10-28). Educational Testing Services (ETS).

Shuey, E. A., Kim, N., Cortazar, A., Poblete, X., Rivera, L., Lagos, M. J., Faverio, F., & Engel, A. (2019). *Curriculum alignment progression between early childhood ad care and primary school: A brief review and case studies* (Working Paper No. 193). Organisation for Economic Co-operation and Development (OECD).

Tannenbaum, R. J., Baron, P. A., & Kannan, P. (2015). *Alignment between innovative summative assessment prototypes and the Common Core State Standards: An exploratory investigation* (Research Memorandum No. RM-15-07). Educational Testing Service.

Thacker, A., Dickinson, E., Paulsen, J., & Buckland, W. (2021). *Summative English Language Proficiency Assessments for California (ELPAC) alignment study report.* Human Resources Research Organization.

Thacker, A., Nemeth, Y. M., Hardoin, M., Namniek, R., Longabach, T., Deatz, R., Wiley, C., & Banjanovic, E. (2018). *California Alternate Assessment alignment study: Mathematics and English Language Arts/Literacy final report*. Human Resources Research Organization.

Thompson, S. J., Johnstone, C. J., Thurlow, M. L., & Altman, J. R. (2005). *2005 state special education outcomes: Steps forward in a decade of change.* University of Minnesota, National Center on Educational Outcomes.

Tindal, G. (2005). *Alignment of alternate assessments using the Webb system*. Council of Chief State School Officers.

U.S. Department of Education. (USDoE, 2018). *A state's guide to the U.S. Department of Education's assessment peer review process.*

Vockley, M., & Lang, V. (2009). *Three approaches to aligning the National Assessment of Educational Progress with state assessments, other assessments, and standards.* The Council of Chief State School Officers.

Way, W. D., & Croft, M. (2020). Alignment issues facing testing organizations. *Educational Measurement: Issues and Practice, 39*(2), 28-30.

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education* (Research monograph no 6). Council of Chief State School Officers.

Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states* (Research monograph no 18). Council of Chief State School Officers.

Webb, N. L. (2002). *Alignment study in language arts, mathematics, science, and social studies of state standards and assessments for four states.* The Council of Chief State School Officers.

Webb., N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education, 20*(1), 7-25. https://doi.org/10.1080/08957340709336728

Webb, N. L., Alt, M., Ely, R., Cormier, M., & Vesperman, B. (2005). The WEB Alignment Tool: development, refinement, and dissemination. Council of Chief State School Officers.

Webb, N., & Herman, J., & Webb, N. (2006). *Alignment of mathematics state-level standards and assessments: The role of reviewer agreement* (CSE Report 685). National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Webb, N. L., & Smithson, J. (1999). *Alignment between standards and assessments in mathematics for grade 8 in one state* [Paper presentation]. AERA, Monteal, Quebec, Canada.

Welch, M. J., Davis, S., Isaia, R. G., Johnston, W. T., Stein, L. B., Jenuwine, H., & Macdonald, K. (2016). *Aligning evaluation: How much do teacher evaluation rubrics emphasize Common Core instruction?* American Institutes for Research.

WestEd. (2015). *California English Language Development Standards, Common Core State Standards for Mathematics, and Next Generation Science Standards: Correspondence study and augmentation document project.* Retrieved from https://www.cde.ca.gov/ta/tg/ep/documents/corspndstdyrptab899.pdf

Whitaker, A. A., Jenkins, J. M., & Duer, J. K. (2022). Standards, curriculum, and assessment in early childhood education: Examining alignment across multiple state systems. *Early Childhood Research Quarterly*, *58*, 59-74.

Wixson, K. K., Fisk, M. C., Dutro, E., & McDaniel, J. (2002). *The alignment of state standards and assessments in elementary reading*. Center for the Improvement of Early Reading Achievement.

Yilmaz, F., & Oner Sunkur, M. (2021). A curriculum alignment analysis: A sample of life sciences course curriculum (2018) for 3rd-grade students. *Journal of Qualitative Research in Education, 27*, 279-297. https://doi.org/10.14689/enad.27.13

Appendix. Content Area Focused Alignment Studies

Source	Alignment Method	Unit(s) of Analysis	Content Area(s)
Achieve (2006)	Achieve	Washington state mathematics standards, college placement tests	Mathematics
Achieve (2018)	Achieve	ACT and CCSS	ELA and Mathematics
Achieve (2010b)	Achieve	California and Massachusetts state math standards and the CCSS	Mathematics
ACT (2022)	Unique	Assessment, AZ standards	Any
Anderson (2002)	Unique [Bloom's]	Instructional objectives, instructional activities, formal and informal assessments	Any
Atchison et al. (2022)	Adapted SEC	Content of instruction, state standards	ELA and Mathematics
Bae et al. (2019)	Unique	Learning objectives, course assessment items	Mathematics
Brown & Conley (2007)	Adapted Webb	State HS exams, college readiness standards	ELA and Mathematics

Source	Alignment Method	Unit(s) of Analysis	Content Area(s)
D'Agostino et al. (2011)	Unique	Arizona math standards and corresponding test items	Mathematics
Davis-Becker & Wiley (2017)	Adapted Webb	WI standards, ACT standards, ACT test items	ELA
Dickinson et al. (2020)	Adapted Webb	California science test and NGSS	Science
Doorey & Polikoff (2016)	Center for Assessment (2016) [Webb]	Assessment items, test development documents, CCSS	ELA and Mathematics
Dynamic Measurement Group, Inc. (2018)	Achieve – content centrality	Acadience Reading assessment and CCSS	ELA
EdMetric (2020)	Webb	Transcend item pool and CCSS	ELA and Mathematics
Forte et al. (2017)	Forte (2017)	Georgia assessment system	ELA, Mathematics, Science, and Social Studies
Hart et al. (2011)	Borrows from Webb, Porter, SEC, Achieve, WestEd, the Education Policy Improvement Center (EPIC), the Center for Assessment	CCSS, AP course materials	ELA and Mathematics
Human Resources Research Organization (HumRRO, 2016)	HumRRO	Smarter Balanced test blueprints and specifications, test items, and the CCSS	ELA and Mathematics

Source	Alignment Method	Unit(s) of Analysis	Content Area(s)
Impara et al. (2000)	Adapted Webb	Norm-referenced test items, NE's language arts content standards	ELA
Iowa Testing program (2017)	Adapted Webb	The Iowa tests and the CCSS	ELA and Mathematics
Lombardi (2006)	Adapted Webb [Bloom's]	Minnesota state mathematics standards and state tests	Mathematics
Lopez (2013)	Webb	International standards and assessment	Mathematics
Newton & Kasten (2013)	Webb and SEC	State mathematics standards and assessments	Mathematics
Nemeth et al. (2016)	HumRRO	SAT, CCSS	ELA and Mathematics
Organisation for Economic Co-operation and Development (OECD, 2013)	Unique	PISA test items and the CCSS-M	Mathematics
Pemberton et al. (2006)	Powerful writing	Formative writing assessments	ELA
Polikoff & Porter (2014)	SEC	Instruction, standards, tests	ELA and Mathematics
Polikoff et al. (2015)	Adapted SEC [Revised Bloom's]	Mathematics textbook topics and CCSS	Mathematics

Source	Alignment Method	Unit(s) of Analysis	Content Area(s)
Porter & Smithson (2002)	SEC	Content standards, assessment, instruction	Mathematics
Porter et al. (2007)	SEC	Mathematics and science standards, assessments, and instruction	Mathematics and Science
Porter et al. (2011)	SEC	State and international standards, CCSS, reports of teacher practice	ELA and Mathematics
Reese et al. (2015)	Adapted Davis-Becker & Buckendahl (2013) framework for evaluating alignment studies	Praxis performance tasks and InTASC standards	Any
Renaissance Learning (2011)	Unique	TN Standards and Product Skills	Mathematics
Stone & Wylie (2019)	Achieve	Test design documentation and test items	ELA and Mathematics
Tannenbaum et al. (2015)	Adapted Webb	CBAL tasks (through-year assessment items) and CCSS ELA standards	ELA
Vasavada et al. (2010)	Borrows from Webb, Porter, SEC, Achieve, WestEd, the Education Policy Improvement Center (EPIC), the Center for Assessment	Skill categories for tests, CCSS	ELA and Mathematics

Source	Alignment Method	Unit(s) of Analysis	Content Area(s)
Webb (1999)	Webb	State science and math standards and assessments	Mathematics
Webb et al. (2006)	Webb and Adapted Webb	State mathematics standards and assessments	Mathematics
Webb & Smithson (1999)	Webb	State mathematics standards and assessments	Mathematics
Welch et al. (2016)	AIR Protocol (found in Appendix A of the report)	Teacher evaluation rubrics, teaching practices associated with CCSS-aligned standards	Any
Wixson et al. (2002)	Adapted Webb	Elementary reading state standards and assessments	ELA
Yilmaz & Oner Sunkur (2021)	Revised Bloom's Taxonomy	Science objectives, instructional activities, assessment items	Science

Text inside brackets indicates the framework used for examining cognitive complexity.