Using Webb's Alignment Model to Measure Intended-Enacted Curriculum Alignment: A Brief Treatment

By Bradley C. Niebling, Ph.D.

Introduction

Curriculum alignment, by that name or others (e.g., alignment, opportunity to learn, curriculum articulation, connections, linkages, coordinated), has been brought to light in several education policies, as well as research and evaluation studies. The marketplace is flooded with "aligned" products, as well as processes and tools that will help schools "align their curriculum." What we do know from research is that curriculum alignment can have a both a significant and noticeable impact on student learning (e.g., Gamoran et al., 1997, Cohen, 1987). Unfortunately, there is little available information available to help educators either examine alignment or to determine whether or not information on alignment provided to them is of quality. The purpose of this paper is to provide a brief history of curriculum alignment work, processes and tools available to educators to examine alignment, and to provide an example of how an existing alignment process, the Webb tools, can be extended into collecting enacted curriculum data, an area of curriculum alignment that has few quality resources.

Background Information

Policy Context

Curriculum alignment is a concept that has received increased attention politically since the report, *A Nation at Risk*, came out in 1983. This report claimed that educational goals for all students needed to be identified, because students from the United States were performing poorly on standardized tests compared to students in other countries. Some of this evidence comes from the Second International Mathematics Study (SIMS) and the Third International Mathematics and Science Study (TIMSS). These studies have consistently revealed national differences in the content, depth and breadth of instruction, and the relationship of this instruction to student achievement (e.g., McKnight, Crosswhite, Dossey, Kifer, Swafford, Travers, & Cooney, 1987).

Soon thereafter, federal legislation started mandating the creation of content and performance standards at the state level. President Clinton's *Goals 2000: Educate America Act* (Linn, 2000; McDonnell, McLaughlin, & Morison, 1997) not only set national education standards, it encouraged states to create their own content and performance standards. Later, the reauthorization of the *Elementary and Secondary Education Act* (ESEA) in 1994, known as the Improving America's Schools Act (IASA) (Linn, 2000; McDonnell, McLaughlin, & Morison, 1997) mandated that all states develop content and performance standards for reading, language arts, and mathematics. This also required states to fulfill a number of other requirements, such as reporting on how they were meeting standards and provisions for teachers to learn about and enact curriculum that is aligned with content and performance standards (McDonnell, McLaughlin, & Morison, 1997).

The ESEA was again reauthorized in 2001, becoming the *No Child Left Behind Act of 2001* (NCLB). This reauthorization not only required states to have content and performance

standards, it also required more frequent test administration for accountability purposes. Collectively, these policies called not just for setting standards, but for creating assessments that were "aligned" with them. The expectation was that setting common learning standards and requiring students to take assessments aligned to them, would result in instruction that is tightly aligned to these other two elements.

Today, there continues to be a sentiment that as a nation, we have not done enough to set high standards for students and to hold states, districts, schools, and teachers accountable to making sure all students meet these standards. Achievement data from the Trends in International Mathematics and Science Study (i.e., the new title for TIMSS) and the Programme for International Student Assessment (PISA) indicate that we continue to lag behind many other countries. The theory that we have not set high enough learning standards continues to this day. The development of the Common Core State Standards (CCSS) for English/Language Arts and Mathematics is another step further towards nationalizing what our students are supposed to know and be able to do and addressing the issue of not setting high enough standards.

The implicit theory driving these initiatives regarding revision of standards, as with previous efforts, is that by adopting/improving standards and developing assessments that will align with those standards, the content of instruction will change, thereby aligning with the standards and assessments. In other words, the prevailing hypothesis is that the observed student learning outcomes (that are lagging behind those of other countries) is caused by insufficient standards that in turn are causing insufficient instruction.

Research Context

The following statement is one likely considered common sense by many educators: *students perform better on assessment activities when they have had a chance to learn the things on which they are being assessed*. Evidence from research on this issue seems to support this idea. In general, research evidence indicates that as a student's opportunity to learn increases, so to do student outcomes (e.g., Cohen, 1987; Gamoran, Porter, Smithson, and White, 1997).

These findings from research only tell part of the story. Some research indicates that the effect of opportunity to learn what is assessed ranges from approximately 1.0-3.0 (Cohen, 1987), while other studies indicate an impact around 0.40 (e.g., Gamoran, Porter, Smithson, and White, 1997). Although this is a wide range, controlled research studies in education with effect sizes of these sizes are usually considered practices worth pursuing.

It is helpful to examine multiple factors and practices at the same time when trying to understand what does and does not impact student learning, not just degree of alignment. For example, one study found that opportunity to learn had a positive impact on student learning even when taking into consideration factors typically associated with students do not typically succeed in school (e.g., low prior achievement, low socioeconomic status, membership in a culturally and ethnically diverse group) (e.g., Gamoran, Porter, Smithson, and White, 1997). Put another way, the impact those other factors usually have was almost completely eliminated by alignment between the enacted and assessed curricula.

Yet another reasonable question to ask is "can alignment help close achievement gaps?" What Gamoran and his colleagues discovered was that increased alignment between the enacted and assessed curriculum has shown to play a role in closing the achievement gap (e.g., Gamoran, Porter, Smithson, and White, 1997). That means that the difference between low performers and high performances was decreased. In other words, the rate of learning for lower performers was higher than for high performers.

Overall, research evidence continues to grow supporting the importance of curriculum alignment in positively impacting student learning, regardless of what individual students bring to the table. While the proportion of impact alignment has on student learning may vary from study to study, it continues to demonstrate a noticeable impact on student learning.

What is Curriculum Alignment?

It is important to clearly define the concepts embedded in curriculum alignment, as these terms will be used throughout this discussion to mean very specific things. Defining key concepts and terms has value beyond communicating the meaning of the terms in isolation. The terms below actually reflect an organizing framework for not only helping us better understand curriculum alignment, but helping us to be more critical consumers of alignment work, research, and products. Below are the set of definitions I will use throughout the rest of the document, including visual organizers for the defined terms to represent the curriculum alignment framework.

CURRICULUM	
Curriculum can be divided into four categories: intended, enacted, assessed, and learned curricula (Porter, 2006)	
Intended curriculum	The knowledge and skill targets for the enacted curriculum, often captured in content standards or other similar documents
Enacted curriculum	The knowledge and skills actually delivered during instruction in the classroom and other learning settings, and how it is taught
Assessed curriculum	The knowledge and skills that are assessed to determine achievement
Learned curriculum	The knowledge and skills students actually acquire

Figure 1. Definitions for multi-dimensional curriculum framework.

INSTRUCTION	
Instruction can be divided into two categories: instructional practices and instructional content (Porter & Smithson, 2001)	
Instructional practices	Methods by which instructional content is delivered; *how* content is taught
Instructional content	Enacted curriculum students are exposed to and expected to acquire; *what* is actually taught

Figure 2. Definitions for instruction within multi-dimensional curriculum framework.

ASSESSMENT
A system of processes and tools that are used to determine the extent to which students are

acquiring or have acquired the knowledge and skills listed in the curriculum and delivered via instruction (Niebling, et al., 2008) In general, there are four types of assessment decisions:	
Summative	Comprehensive in nature, provides accountability, and is used to check the level of learning at the end of a unit of study. (RtI Action Network)
Formative	Collection of evidence about student learning that is used to inform instructional decisions in an ongoing manner. *Progress Monitoring*, a type of formative assessment used in RtI systems, is a scientifically-based practice used to assess students' academic performance and evaluate the effectiveness of instruction. It is the process used to monitor implementation of specific interventions. (RtI Action Network)
Screening	A quick check of all students' current levels of performance in a content or skill area. (RtI Action Network) Identifies potential academic and/or behavioral concerns in need of additional assessment. (Midwest Instructional Leadership Council)
Diagnostic	Diagnostic assessments are used to confirm screening results and to inform intervention by determining a student's particular difficulties. (RtI Action Network)

Figure 3. Definitions for assessment within a multi-dimensional curriculum framework.

ADDITIONAL CURRICULUM CONCEPTS/TERMS	
These are some additional items that are often part of curriculum discussions.	
Content Standards	Broad statements that identify the knowledge and skills that students should acquire. (Iowa Department of Education, 2011)
Instructional materials	Instructional resources (e.g., textbooks, teacher-developed activities) that represent the content in the written curriculum and are used to engage students in the learning process (Niebling, et al., 2008)

Figure 4. Definitions for additional curriculum terms within a multi-dimensional curriculum framework.

ALIGNMENT
*The extent to which and how well curricular categories and the elements within them (e.g., content standards, instructional content, and assessment practices) work together to guide instruction and, ultimately, facilitate and enhance **student learning** (e.g., Webb, 1997).*

Figure 5. Definitions for alignment.

DIRECTIONALITY	
The direction in which alignment is examined can be broken down into two approaches (Niebling et al., 2008).	
Horizontal Alignment	Degree of match across two curricular categories (e.g., instructional content with state or national standards) within a single level (e.g., same grade comparisons). Or, degree of match within one curricular category (e.g., instructional content of two teachers of different

	sections of the same course) within a single level (e.g., high school).
Vertical Alignment	Degree of match within one curricular category (e.g., district benchmark assessments) across multiple levels (e.g., across grade levels).

Figure 6. Definitions for directionality within a multi-dimensional alignment framework.

DIMENSIONS	
There are many approaches to examining alignment (e.g., Surveys of Enacted Curriculum, Webb methods), each of which examine different aspects of alignment relationships. In general, these different aspects can be summarized along three dimensions, regardless of the methods used (Niebling et al., 2008).	
Topical/Conceptual Knowledge	Subjects, information, and ideas that students are supposed to learn.
Cognitive Complexity/Demand	What students are expected to do with the topical/conceptual knowledge (e.g., Surveys of Enacted Curriculum Cognitive Expectations dimension, Bloom's Taxonomy, Webb's Depth of Knowledge).
Emphasis	The extent to which topical/conceptual knowledge with accompanying complexity/demand are addressed by the intended, enacted, or assessed curriculum.

Figure 7. Definitions for dimensions within a multi-dimensional alignment framework.

LEVEL OF ANALYSIS	
When engaging in an examination of alignment in either direction, along any dimension(s), the specificity with which alignment is considered can vary along a continuum of less to more specific. This is referred to by Porter (2002) as "grain size."	
Coarse-Grained	Categorical, global or general examinations that tend to yield higher estimates of alignment.
Fine-Grained	Specific, targeted examinations that tend to focus on sub-skills within larger categories that tend to yield lower estimates of alignment (Niebling et al., 2008).

Figure 8. Definitions for level of analysis within a multi-dimensional alignment framework.

The Learning – Centered Curriculum Triangle

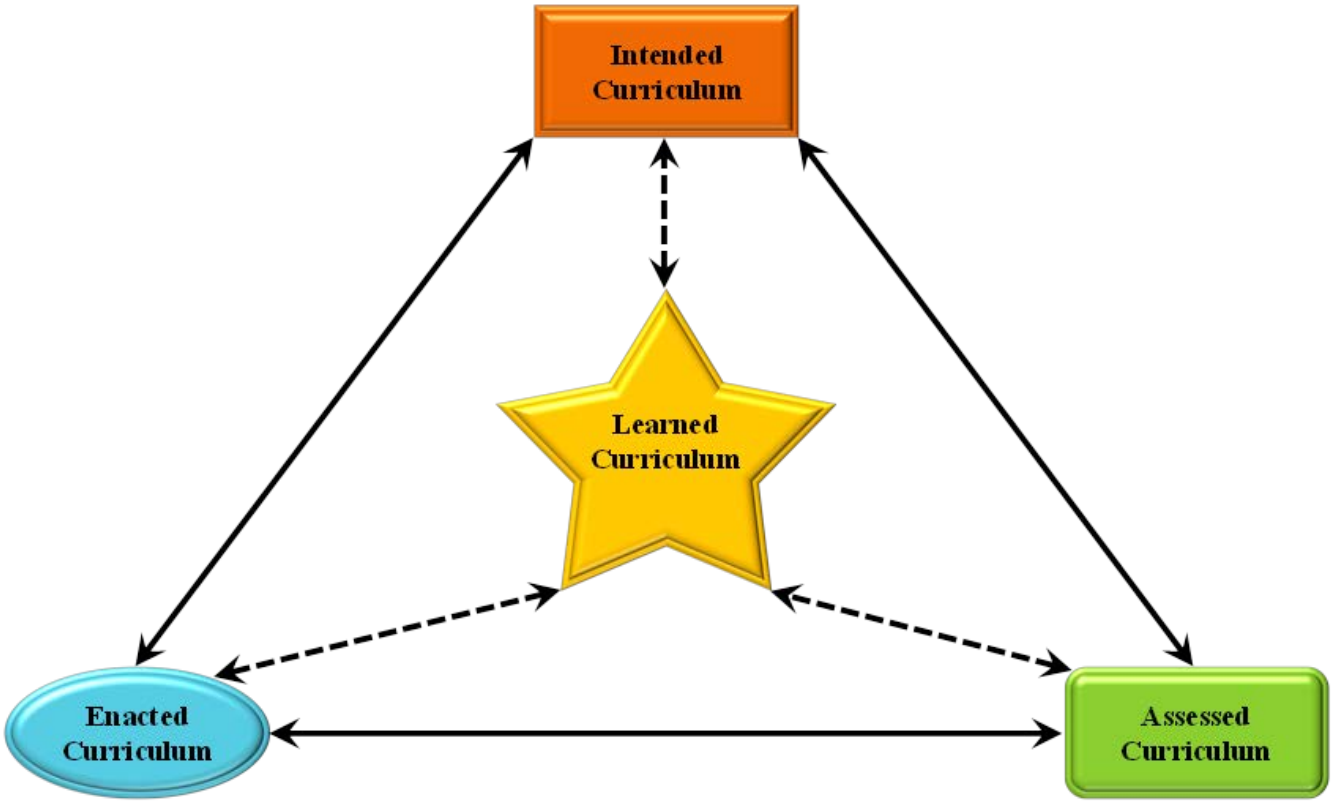

Figure 9. Visual depiction of a multi-dimensional curriculum framework.

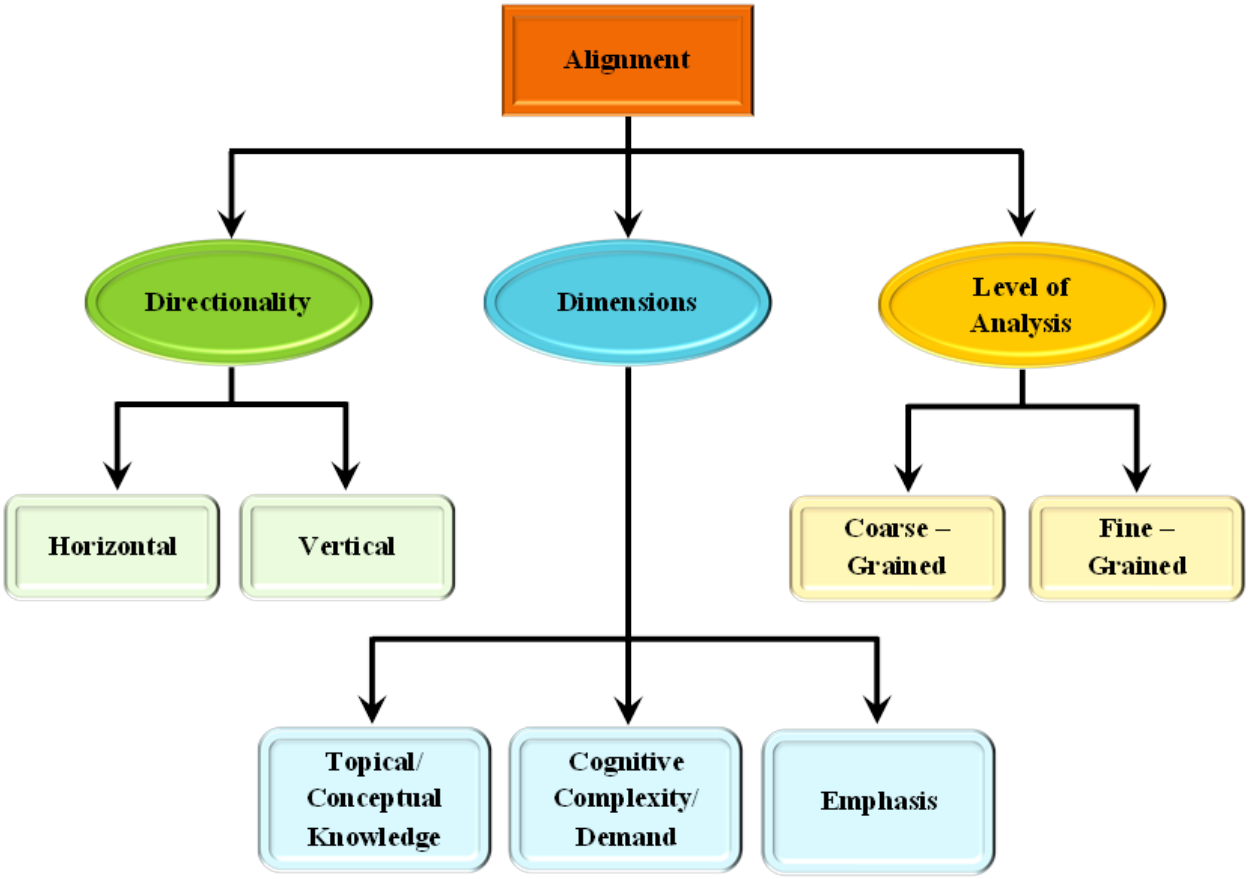

Figure 10. Visual depiction of a multi-dimensional alignment framework.

Methods of Measuring Alignment

Using the frameworks. How do we know if something is "aligned?" Aligned with what? Hopefully, information shared in the previous section defining curriculum alignment has depicted alignment as more than a "yes" or "no" proposition. It is a multi-faceted area of study and practice. Yet, we continue see products on the market today claiming they are "aligned," especially with the Common Core State Standards for English/Language Arts and Mathematics. There are also products on the market that claim to "help you align your curriculum." The purpose here is not to diminish the potential benefits of products on the market. Instead, I hope to illuminate what the science of curriculum alignment has to offer in terms of how we can reliably and validly determine the degree of alignment among curricular elements.

First, we can use the curriculum alignment framework to either design an alignment measurement process, or to examine other frameworks, research, or processes/tools on the market. For illustrative purposes, I will use the following example as the context for describing how to apply the curriculum alignment framework.

The Director of Assessment and School Improvement for Central Unified School District (Central) wants to know the extent to which math teachers at A. H. Miles High School are teaching the Common Core State Standards for Mathematics. In particular, she is curious about

not just about what is being taught, but what type of thinking students are being required to engage in. She also wants teachers to reflect beyond the Domain and Cluster levels, all the way down to the standard level. In the future, she hopes to learn more about how much time is being spent on each of the Common Core standards.

To start, we can use the learning-centered curriculum triangle (Figure 9) to determine which curricular areas the Director wants to compare to examine curriculum alignment. In this scenario, the Director wants to compare the *enacted* curriculum of teachers to the *intended* curriculum of the Common Core. Next, we look to the multi-dimensional alignment framework (Figure 10) to determine which aspects of alignment the Director wants to examine. It is important to note that each of the three characteristics identified on the second level of the framework, Directionality, Dimensions, and Level of Analysis are always in play with alignment. Put another way, these three alignment characteristics may or may not be addressed in an alignment examination or claim, but they are still happening in practice.

In the case of the Director from our scenario, she wants to engage in a horizontal examination that includes both topical/conceptual knowledge and cognitive complexity at a relatively fine-grained level. Let's take each of these characteristics in turn. The alignment comparison is horizontal because it is (a) across two curricular areas (i.e., enacted and intended), and (b) it is within one defined level (i.e., grades 9-12). Although it is possible to engage in a horizontal alignment examination within a single curricular area (e.g., comparing the enacted curriculum of two different teachers of the same course, curricular materials to content standards), the fact that the Director wants to examine alignment across curricular areas within a single level puts this example solidly in the Directionality category of horizontal.

For the Dimensions characteristic, there are a few things the Director is looking for that indicates she is looking for topical/conceptual knowledge and cognitive complexity, but not emphasis. First, the indication of "what" by the Director could be taken more broadly, if it wasn't followed by a desire to look at student thinking as well. As such, we can take that "what" the Director is looking for to be topical/conceptual knowledge. Student thinking then fits best in the category of cognitive complexity. Finally, while the Director does want to know how much time is being spent on different standards, that is future work. As such, emphasis is not a current alignment dimension to be examined in this scenario.

For the Level of Analysis characteristic, there is one aspect of the scenario in particular that points to the Director's desire to dig into alignment at a relatively fine-grained level. By digging down all the way to the standard level of the Common Core, the examination is relatively fine-grained compared to what it would be at the more general levels of Domain and Cluster.

Selecting or developing processes and tools. The next questions to consider should be "what processes and tools can we use to actually engage in the alignment examination? Does the Director have to buy tools, or can she develop some processes and tools locally?" The answer is either approach could work. Factors to consider when making this sort of decision include time, financial resources, technical assistance, and technology and tools available that are needed to either acquire or develop curriculum alignment tools (Niebling, Roach, and Rahn-Blakeslee,

2008). It is also important to consider what it takes to reliably and validly capture curriculum alignment data.

At the heart of measuring curriculum alignment is trying to accurately capture information about two curricular elements so we compare them to determine how similar they are to each other. This is often a matter of comparing two relatively static documents to each other. For example, we can compare test items from a state accountability assessment to the Common Core State Standards. Although there are challenges associated with this type of work, it is still a relatively straightforward process. It is when we try to compare the more dynamic enacted curriculum with the intended or assessed curriculum that the work becomes much more challenging.

In general, there are four different approaches to try and capture the content of the enacted curriculum: (1) review of assessment results, (2) review of course-related documents, (3) observe instruction, and (4) self-report by teachers. The foundation of using assessment results as a proxy for enacted curriculum is that if students do well on the assessments, then they must have had a sufficient opportunity to learn what was tested. This assumption could be problematic given that there could be several reasons why student do or don't do well on any given assessment (e.g., relationship with teacher, sense of safety at school, prior achievement). It can also be problematic to make this assumption without having a method of examining the content of what was assessed, student performance in that context, and how all of that relates to what was taught. In other words, without actual alignment data between the enacted and assessed, and even the learned, curricula we can't say with a great deal of confidence that student performance on the assessment is due to what students had the opportunity to learn. Nor does this approach tell us anything about the degree of alignment between the intended and enacted curricula.

The review of course-related documents as a method of examining the enacted curriculum is a practice that has great appeal, primarily because the approach calls for reviewing static documents that are supposed to represent what is taught. Examples of course-related documents that are typically used for such a review include course syllabi, lesson plans, textbooks and related materials, and curriculum maps. This can be a reasonable approach for capturing the enacted curriculum, if the following conditions are met: (a) the document review occurs after instruction was delivered, and (b) the person(s) delivering the enacted curriculum is the person(s) engaging in the document review. Furthermore, if using course-related is to be used for an alignment examination, the review must be done in the context of the defined intended curriculum. These are conditions that are often overlooked for the sake of getting the work done. Reviewing course-related documents may be more useful as a component of a more comprehensive review of the enacted curriculum that includes self-reflection.

Observing teachers actually enacting curriculum is considered to be the "gold standard" of collecting these type of data (Schmidt & Maier, 2009), particularly if done by an independent, third-party observer. Observation data are viewed as the most accurate form of collecting enacted curriculum data. The primary drawback to this approach is that it is cumbersome and expensive. As such, it would be unrealistic for most school systems to engage in comprehensive observation of teachers to capture the entire enacted curriculum, even for one teacher, for an entire school year.

Finally, self-report by teachers is a method of collecting enacted curriculum data that has received attention for a variety of reasons. Generally speaking, there are two approaches to collecting teacher self-report data: (a) ongoing logs, and (b) surveys. One method to determine the reliability and validity of each of these approaches is to compare each to the other, including matching with observations of instruction (i.e., the "gold standard"). In general, available research indicates that once-a-year surveys are similar enough in content to ongoing teacher logs and observations that they merit strong consideration for collecting enacted curriculum data. This is appealing to many because of the relatively low cost in time and money compared to the other two methods (Schmidt & Maier, 2009). A review of course-related documents (i.e., the second method listed above) is often used to help teachers with their reflections on surveys. The primary drawback of the survey approach is that these type of data are typically not as descriptive as teacher logs or other narrative approaches. Relatedly, as surveys require more specific reflections by teachers, in some cases the accuracy of the reflection decreases.

Ultimately, deciding on one or more of these methods for capturing enacted curriculum data for alignment analyses depends on both the purpose of the alignment work as well as the available resources to do the work. The purpose of this paper is to discuss how to expand the availability of survey options for collecting enacted curriculum data to examine enacted-to-intended curriculum alignment, an area of alignment work that lacks a variety of reliable and valid processes and tools.

Extending Webb's Alignment Model to the Enacted Curriculum

Of particular interest in the curriculum alignment field, is creating efficient, reliable, and valid methods for collecting enacted curriculum data for the purpose of comparing what students actually get the opportunity to learn to what they are supposed to learn (intended curriculum) and what they are assessed on (assessed curriculum). The tools that have perhaps the strongest research support are the Surveys of Enacted Curriculum.

Comparing the SEC and Webb Model

The (SEC) are a set of processes and tools that can be used to examine the content of the intended, enacted, and assessed curriculum, as well as the alignment among them (Porter, 2002). The SEC is unique amongst the standardized, evidence-based approaches to collecting and analyzing curriculum alignment data because they can be used to collect enacted curriculum data. Additional factors lead the SEC to be considered one of the best tools available for measuring curriculum alignment. Not only can the SEC be used to measure all of the components of the learning-centered curriculum triangle (Figure 9), they address all of the characteristics in the multi-dimensional alignment framework (Figure 10). Furthermore, research suggests that the SEC (and its precursors) yield very similar results to observations and teacher logs, and the alignment index generated with the SEC has been found to be a significant predictor of student achievement (Gamoran et al., 1997). This evidence helps solidify the SEC as the state-of-the-art in gathering curriculum alignment measurement data.

Similar to the SEC, Webb's model has been used widely in the United States, especially since the reauthorization of the *Elementary and Secondary Education Act (ESEA)* in 2001, known as

No Child Left Behind (NCLB; Webb, 2007). Unlike the SEC, Webb's model only addresses the intended and assessed curriculum. Furthermore, the Webb model also addresses all of the characteristics described in the multi-dimensional alignment framework (Figure 10).

Impetus for Expanding Webb's Model to the Enacted Curriculum

Most recently, the Common Core State Standards have been coded according to Webb's Depth of Knowledge (DOK), a leveled framework for describing the cognitive complexity of content standards and assessments. It should be noted that the Common Core has also been coded using the SEC language framework (Porter, McMaken, Hwang, & Yang, 2011). States who have adopted the Common Core can use the SEC to examine alignment among the intended and enacted curriculum right now.

With the widespread use of Webb's model, particularly with the coding of the Common Core using Webb's DOK, there is an opportunity to innovate in the field of curriculum alignment by developing a set of processes and tools to use Webb's model to collect and analyze alignment between the enacted and either the intended or assessed curricula. Expanding the Webb model not only provides a natural link to existing intended and assessed alignment data (e.g., Common Core), it is an opportunity to measure and describe enacted curriculum alignment from a different perspective, thereby providing a more comprehensive ability for the field to conceptualize and analyze alignment.

Digging into Webb's Alignment Model

It is important to understand Webb's current alignment model before introducing additions to it. Briefly, Webb's alignment model can be used to examine alignment between the intended and assessed curriculum in four dimensions (Webb, 2005): (a) categorical concurrence, (b) balance of representation, (c) range-of-knowledge correspondence, and (d) depth-of-knowledge consistency.

Table 1. Webb model's alignment dimensions.

Dimension	**Description**
Categorical Concurrence	The assessment in question has at least six items measuring the same content for any given comparison standard.
Range of Knowledge	At least half of the objectives (i.e., finer-grained statements) for any given standard are assessed by at least one test item.
Balance of Representation	Relatively equal emphasis can be found between the number of objectives for any given standard the assessment.
Depth of Knowledge	A measure of cognitive complexity, at least half of the test items for any given objective are at or above the cognitive complexity level of that objective.

Another aspect of the Webb process that is unique is that there are specific criteria set for what constitutes "good enough" alignment between the intended and assessed curricula. Whenever criteria are not met, the coding can be traced back to specific standards, objectives, and

assessment items that are not performing well enough, or to dimensions within the framework to determine which areas need to be further addressed to meet the criteria.

Perhaps the most complicated work involved in using the Webb alignment model is helping coders of standards, objectives, and test items understand and reliably code them according to the DOK framework. There are two aspects to understand about Webb's DOK, in addition to specific coding procedures: (a) the generic leveled framework, and (b) content-specific frameworks for English/language arts, mathematics, science, and social studies. An in-depth review of each content-specific DOK framework is beyond the scope of this paper. The four levels of the generic DOK framework are: Level 1 (Recall), Level 2 (Basic Application), Level 3 (Strategic Thinking), and Level 4 (Extending Thinking) (Roach, Niebling, & Kurz, 2008; Webb, 2005).

Coders are first trained on the DOK framework and the coding process. During the training process, coders practice coding objectives and assessment items to not only learn the coding process, but to calibrate with each other. During the actual coding process, coders independently assign single DOK levels to objectives, and later to individual assessment items. Raters later come back together to have consensus discussions about the objectives and/or assessment items for which they assigned different DOK levels. If raters have a hard time deciding between two different DOK levels for an objective or test item, it is generally recommended to choose the higher of the two levels under consideration. For example, if a rater is deciding whether or not a test item reaches a depth of Level 1 or Level 2, it is recommended in most cases to indicate Level 2. The same holds true for consensus decisions (Webb, 2005).

In the alignment analysis, DOK is considered matching if at least half of the items assigned to an objective are at or above the same DOK level of the objective. In other words, an objective might have several assessment items assigned to it (i.e., categorical concurrence, or a "content" match), but the DOK is only considered to be "sufficiently aligned" if at least 50% of those items are at the same or higher DOK level (Roach, Niebling, & Kurz, 2008; Webb, 2005).

Key Considerations for Expanding Webb's Alignment Model to the Enacted Curriculum

Innovating by adding the enacted curriculum to Webb's alignment model does not mean we should start from scratch. A great deal already exists in Webb's alignment model that we can draw upon. I will focus in particular on the DOK aspect of Webb's model, primarily for this reason: we have little basis at this time to set criteria for "good enough" alignment between the enacted and either the intended or assessed curriculum. Furthermore, there may be simpler methods for capturing the topical/conceptual knowledge aspect of alignment than the existing methods (e.g., categorical concurrence) in the Webb model, particularly for a new innovation. For the sake of illustration, I will focus on an intended-enacted curriculum alignment examination. With that said, there are a number of considerations one could take when trying to add the enacted curriculum to the Webb alignment model.

The intended curriculum needs to be coded for DOK. In order for any alignment comparisons to be made between intended and enacted curricula, the intended curriculum would need to be coded according to the content-specific DOK framework. In general, I would

recommend that Webb's existing process for coding the DOK of standards and objectives be used for this process (e.g., Webb, 2005). It is a well-established process that has a history of positive results. As was previously mentioned, the Common Core State Standards were coded for a variety of features, including Webb's DOK (Sato, Lagunoff, & Worth, 2011). Although a great deal of Webb's training manual (2005) is cited in the Common Core coding report, it is difficult to tell all of the procedures that were followed by the group of coders for this study. It is clear, however, that the group deviated from Webb's protocol in at least two ways: (1) only two coders were used for each content area, and (2) some standards were assigned multiple DOK levels.

Webb's alignment model calls for at least *five coders* to engage in the coding process for the purpose of completing an alignment study (Webb, 2005). While the Common Core coding process utilized only two coders for their process, this may not significantly impact their work in a negative way. The Common Core coding process used a double-rater "read behind" process that allowed ongoing consensus discussions to occur (Sato, Lagunoff, & Worth, 2011).

The approach of assigning multiple DOK levels to single standards was taken because for many of the standards, there were several "sub-standards" within them. That is, standards were often written as compound statements that addressed multiple topics and cognitive expectations for students (Sato, Lagunoff, & Worth, 2011). With that said, standards written as compound statements are still called to be assigned single DOK levels within Webb's alignment model for an alignment study. Although one of the main reasons the Common Core was coded using Webb's DOK was to inform future item development, the work itself was a descriptive not alignment study. While that may cause problems for an alignment study, it does not necessarily cause problems for a descriptive study (N. Webb, personal communication, January 24, 2012).

We are still left with a significant deviation from Webb's alignment model should we try and use the Common Core DOK codes to measure the alignment of the intended and enacted curricula. A proposal for addressing this problem is found below in the section on developing metrics and data displays.

Topical/conceptual knowledge must be recorded by teachers. If teachers are asked to reflect on the cognitive complexity of their enacted curriculum, using Webb's DOK or any other cognitive complexity framework, they should also reflect on the topical/conceptual knowledge of their enacted curriculum. Reflecting solely on the cognitive complexity of the enacted curriculum is not only without context, it also results in data that do not have the ability to accurately predict student outcomes (Gamoran et al., 1997).

Trying to draw a direct parallel between Webb's existing process for comparing alignment of intended and assessed curricula could be difficult, as it is challenging to break enacted curriculum into chunks the same way we can break tests down into items. When teachers use the SEC, they are asked to reflect on the content of their instruction in terms of the time spent on different areas of content defined by the unique language framework developed for the SEC (Porter, 2002). This may be another point of departure from Webb's original alignment model. As such, summarizing alignment results in terms of categorical concurrence, range of knowledge, and balance of representation as defined may be difficult. That does not necessarily

mean teacher reflection on the topical/conceptual knowledge of their enacted curriculum is impossible.

For example, alignment reflections using tools like curriculum mapping rely on making "yes" or "no" indications about the standards themselves, without trying to categorize them further. Another approach may be to simply list the standards, either as whole statements or broken down into smaller chunks, for teachers to make "yes" or "no" indications as to whether or not they enacted those parts of the intended curriculum (e.g., the Iowa Curriculum Alignment Toolkit, or ICAT).

Teachers should be trained on content-specific DOK frameworks. Just like for coders of standards and/or test items, teachers should be trained on what the content-specific DOK frameworks are, and how to use them for reflections on their enacted curriculum. For example teachers who are going to reflect on mathematics should be trained on the math-specific DOK framework. Training could include things like collaborative discussions on what the different DOK levels mean and examples of what that could look like in classrooms, as well as decision making rules for what level to rate their instruction on. Such an approach is conceptually consistent with Webb's alignment model, though the specifics of the training would likely somewhat differ since a different curricular area (i.e., enacted curriculum) is being considered.

Electronic, web-based tools to record and analyze data should be developed and used. The increasing accessibility and power of the internet can greatly enhance data collection and use for engaging in intended-enacted curricula alignment work. Specifically, an online, web-based database would greatly enhance the process of collecting and using alignment data. The SEC is completed this way (www.seconline.org), as is the Webb alignment process (http://wat.wceruw.org/index.aspx).

Using online tools provides accessibility to teachers anytime from any internet-connected device. As such, attention needs to be paid to cross-browser capabilities for such a tool, as well as ongoing and easily accessible technical support should end users have difficulties using the database or if something goes wrong with database functionality. Without such support, end users are likely to become frustrated and disengaged from taking the time, thought, and energy needed to accurately reflect on the content of their enacted curriculum.

Finally, attention will need to be paid to the training and resource materials needed for teachers to easily use the online tools successfully. Examples of such training and resources could include face-to-face or distance (using web-based audio-visual tools like Skype or Adobe Connect Pro) training on how to use the online tools, demonstration videos, cheat sheets with decision making rules, and handouts with step-by-step screen shot directions on how to use various features of the online tools. Having such processes and tools in place can help increase the support teachers will need to successfully use the web-based alignment database.

Metrics and data displays should be developed to summarize alignment data. The power of processes and tools like the SEC and Webb's alignment model is that they provide both numeric and visual summaries of alignment relationships. Although not a replacement for narrative descriptions of teachers' enacted curriculum, having data summarized numerically and

visually provide important additional perspectives on the relationship between intended and enacted curricula. For example, numerical summaries of alignment allow for year-to-year comparisons of teachers' enacted curriculum, and how that changes over time relative to the intended curriculum. Having visual summaries of intended-enacted curricula alignment provide quick and relatively easy-to-interpret information for teachers to collaboratively use to make changes to what they teach to help ensure all students have equity in their opportunity to learn what is defined in the intended curriculum.

Both the SEC and Webb alignment model use somewhat complicated metrics to numerically describe degree of alignment. Although these methods are mathematically sound and ultimately easy to interpret, it may be wise to start with numerical summaries that are easier to compute and explain. One of the most basic metrics that can be used is to calculate the percentage of the intended curriculum addressed in a teacher's enacted curriculum. The formula could look something like this:

$$TCK\ Alignment = \frac{Enacted\ Curriculum}{Intended\ Curriclum},$$

where TCK stands for topical/conceptual knowledge

Adding cognitive complexity using Webb's DOK increases the complexity of any alignment calculation. In general, it may be easier to create additional metrics, instead of trying to come up with one comprehensive metric. We want to know the percentage of standards (i.e., the intended curriculum) with a DOK level match in the enacted curriculum. This could result in two metrics: one for the percentage of DOK matches for taught standards, and a second metric for the percentage of DOK matches for all standards, both taught and untaught.

If we generalized Webb's rule for determining whether or not a test item has a DOK match with an objective to our enacted curriculum, if a teacher's DOK rating of their enacted curriculum meets or exceeds the DOK rating of a standard to which there is a topical/conceptual knowledge match, it would be considered a DOK match. If the DOK rating of their enacted curriculum is lower than for the corresponding DOK rating of the intended curriculum, it would not be considered a match. Using this approach, we could calculate the percentage of DOK alignment for standards that had a topical/conceptual knowledge match as follows

$$DOK\ Matched\ Alignment = \frac{\#\ of\ Enacted\ Curriculum\ DOK\ matches}{\#\ of\ Intended\ Curriclum\ TCK\ matches},$$

where "DOK Matchd" stands for Depth of Knowledge and TCK stands for topical/conceptual knowledge

Calculating the total percentage of DOK matches for all standards in the intended curriculum could look like this:

$$DOK\ Total\ Alignment = \frac{\#\ of\ Enacted\ Curriculum\ DOK\ matches}{\#\ of\ Total\ Intended\ Curriclum},$$

where "DOK Matchd" stands for Depth of Knowledge

The process of calculating intended-enacted curricula DOK alignment gets further complicated when considering Level of Analysis. The potential impact of Level of Analysis is two-fold. First,

we could choose examine standards broken down into smaller chunks instead of whole standards. Or perhaps, we could try to examine DOK alignment around coarser-grained categories on the intended curriculum. For example, for the Common Core State Standards in Mathematics, we may want to examine the degree of DOK alignment not just at the standard level, but at the Cluster or Domain level as well. The former comparison would require both DOK assignments to the smaller chunks, as well as teachers reflecting on the smaller chunks themselves, not just at the coarser-grained standard level. Regardless of the specificity with which the standards were reflected on, the latter option would require the data to be somehow aggregated up to the coarser-grained level.

Another potential approach to digging to a finer-grained level of analysis is with the number of DOK levels assigned to a standard. As was previously mentioned, the coding process completed to assign DOK levels to the Common Core deviated from the standard Webb alignment process by assigning multiple DOK levels to single standards. It is unclear from the report whether multiple DOK levels were assigned to standards in their entirety, or if the coders assigned single DOK levels to the "sub-standards" (or smaller chunks). The available data have multiple DOK levels assigned to the entire standard.

One method of examining cognitive complexity alignment would be to have teachers assign as many DOK levels to their enacted curriculum as they find appropriate. That is, if teachers taught to multiple DOK levels for any given standard (or smaller chunks), they could assign multiple DOK levels to their enacted curriculum reflections. Then, for each standard, you could determine the number of DOK matches for each standard that has multiple DOK levels assigned, not just an overall "yes" or "no" determination for the entire standard. Furthermore, if teachers were allowed to assign multiple DOK levels for their enacted curriculum for each matching standard, two calculations could occur. The first calculation could be to determine if the highest DOK level indicated by the teacher for their enacted curriculum was equal to or higher than the highest DOK level of the standard. If so, then that would be considered a match. A secondary analysis, examining each of the potential DOK matches within that standard could also be calculated.

Once calculations are determined, as was mentioned previously, determining how to display the data can be a very helpful tool for end users. In general, alignment percentages can be easily displayed in tables and bar graphs. Using a method similar to that of the SEC, it may be helpful to have DOK levels across the top of tables, with standard statements down the left-hand side of the tables. That way, the intersection cells can have the DOK assignment(s) for both the standards and the enacted curriculum.

References

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

Gamoran, A., Porter, A. C., Smithson, J. L., & White, P. A. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth *Educational Evaluation and Policy Analysis, 19,* 325-338.

Iowa Department of Education. Standards, *Benchmarks and Grade Level Indicators*. Retrieved from http://educateiowa.gov/index.php?option=com_content&task=view&id=1350&Itemid=2287 .

McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., & Cooney, T. J. (1987). *The underachieving curriculum: Assessing U. S. school mathematics from an international perspective*. Champaign, IL: Stipes Publishing.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. United States Department of Education: Washington, D.C.

Niebling, B. C., Roach, A. T., & Rahn-Blakeslee, A. (2008). Best practices in curriculum, instruction, and assessment alignment. In A. Thomas & J. Grimes (Eds.), Best practices in school psychology, (4)5, 1059-1072. Bethesda, MD: National Association of School Psychologists.

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011). Common Core Standards: The new U.S. intended curriculum. *Educational Researcher, 40*, 103–116.

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

Roach, A. T., Niebling, B. C., & Kurz, A. (2008). Evaluating the alignment among curriculum, instruction, and assessments: Implications and applications for research and practice. Psychology in the Schools, 45, 158-176.

RtI Action Network. *Glossary of Terms*. Retrieved from http://www.rtinetwork.org/glossary.

Sato, E., Lagunoff, R., & Worth, P. (2011). *SMARTER Balanced Assessment Consortium Common Core State Standards analysis: Eligible content for the summative assessment* (Final Report). Retrieved from SMARTER Balanced Assessment Consortium website: http://www.k12.wa.us/SMARTER/pubdocs/SBAC_CCSS_Eligible_Content_Final_Report_030411.pdf.

Schmidt, W.H., & Maier, A. (2009). Opportunity to learn. In G. Sykes, B. L. Schneider, D. N. Plank, & T. G. Ford (Eds.), *Handbook of education policy research.* New York: Routledge.

Webb, N. L. (1997). Criteria for alignment of expectations and assessments in mathematics and science education (Research Monograph No. 8). Madison, WI: National Institute for Science Education, University of Wisconsin–Madison.

Webb, N. (2005). Web Alignment Tool (WAT) Training Manual, Draft Version 1.1. Retrieved from http://wat.wceruw.org/index.aspx.

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education, 20*, 7-25.