The Validity of Measures of Alignment with State Standards Based on Surveys of Enacted Curriculum

Authors: Drew Atchison, Michael S. Garet, Toni M. Smith, and Mengli Song   **June 2021**

Standards-based reforms have been a core element of state and federal efforts to improve education for the past three decades and continue to occupy the center stage in state and federal education policy arenas. These reforms are based on the premise that specification of challenging standards for student learning will drive the content of instruction and assessment, which in turn would lead to increased student learning (Hannaway, 2003; Porter, 2000). Research suggests, however, that teachers' instruction is generally is not well-aligned with state standards, and teachers need support to understand and align their instruction to new state standards (Polikoff, 2012, 2021).

To help improve teachers' instructional alignment to standards, the Center on Standards, Alignment, Instruction, and Learning (C-SAIL), funded by the U.S. Department of Education's Institute of Education Sciences, developed the Feedback on Alignment and Support for Teachers (FAST) program, a virtual coaching program designed to help teachers better align their instruction to their state standards.

During the 2017–18 and 2018–19 school years, a research team at the American Institutes for Research (a C-SAIL partner) conducted a school-level randomized controlled trial in 56 elementary schools in five districts to test the impact of the FAST program, focusing on Grade 4 math and Grade 5 English language arts (ELA). For both subjects, the study found that the FAST program had a positive impact on the alignment of teachers' instruction to state standards, but not on student achievement as anticipated.

Prompted by the perplexing impact findings about the FAST program, the study team explored the validity of measures of instructional alignment used in the study, which were based on Surveys of Enacted Curriculum (SEC; Porter & Smithson, 2001) and served as the foundation for the FAST program as well as key outcome measures for the impact study. This brief highlights key findings from our validity analyses.

Measures of Instructional Alignment

The FAST Instructional Survey was used to assess the alignment of teachers' instruction with state standards. It was developed by the study team and based on the SEC. The survey uses topics (e.g., equivalent fractions, adding whole numbers) paired with cognitive demands (e.g., demonstrating understanding, performing procedures) to describe the content of instruction. See Exhibit 1 for a sample page from the online survey.

Exhibit 1. A Sample Page from the FAST Instructional Survey

Number Concepts

Instruction can focus on different topics at different levels of cognitive demand. The table below provides a list of math topics and associated cognitive demands. Please indicate the level of emphasis you gave to each topic and cognitive demand pair in your math instruction over the log period and use the following scale:

| 0 | 1 | 2 | 3 | 4 | 5 | 6 |

No Emphasis   Moderate Emphasis   Major Emphasis

For example, a "6" for "Decimals" and "Recall/Perform Procedures" would indicate that you placed major emphasis on the Decimals topic at the Recall/Perform Procedures level of cognitive demand.

To see definitions for the cognitive demands click here: Cognitive Demand Definitions

To save the information you entered, click "Submit". If you click "Cancel" you will lose the information.

Topic		Level of Emphasis
Place Value	Recall/Perform Procedures	○0 ○1 ○2 ●3 ○4 ○5 ○6
	Demonstrate/ Communicate Understanding	○0 ○1 ○2 ●3 ○4 ○5 ○6
	Justify/Evaluate/Generalize	●0 ○1 ○2 ○3 ○4 ○5 ○6
	Solve and Interpret Findings from Word Problems	●0 ○1 ○2 ○3 ○4 ○5 ○6
Whole numbers	Recall/Perform Procedures	○0 ○1 ●2 ○3 ○4 ○5 ○6
	Demonstrate/ Communicate Understanding	○0 ○1 ●2 ○3 ○4 ○5 ○6
	Justify/Evaluate/Generalize	●0 ○1 ○2 ○3 ○4 ○5 ○6
	Solve and Interpret Findings from Word Problems	●0 ○1 ○2 ○3 ○4 ○5 ○6

To calculate overall level of instructional alignment with state standards for each teacher, we compared (a) the proportion of emphasis that a teacher reported giving to each topic and cognitive demand pair out of the sum of their reported emphasis across all pairs with (b) the proportion of emphasis that expert coders gave to each topic and cognitive demand pair out of the sum of emphasis given across all pairs. This comparison resulted in an overall alignment index with a value between 0 (complete lack of alignment) and 1 (perfect alignment).

In addition to the overall alignment as typically measured (based on the intersection of topics and cognitive demands), we also examined separate measures of alignment by topic and by cognitive demand. The topic alignment index captured the extent to which a teacher's instruction was aligned with the state standards in terms of emphasis allocated to different topics, regardless of the cognitive demand. The cognitive demand alignment index captured the extent to which a teacher's instruction was aligned with the state standards in terms of emphasis on different cognitive demands, regardless of the topic.

Associations Between Alignment and Value-Added Scores

Our first set of analyses explored the predictive validity of the SEC-based alignment measures by examining the associations between teachers' instructional alignment with state standards and their contribution to student learning as measured by teacher value-added scores. Relying primarily on student-level data from district administrative records, we estimated the value-added scores for study teachers following typical procedures, accounting for both student and teacher characteristics as well as measurement errors (Aaronson et al., 2007; Koedel et al., 2015).

We found that neither overall alignment nor alignment by cognitive demand was statistically significantly associated with value-added for either math or ELA. However, for both subjects, alignment by topic was positively associated with value-added, and the association was statistically significant (see Exhibit 2). Specifically, we found that an increase in the topic alignment index of 0.1 was associated with a 0.09 standard deviation increase in value-added in math and a 0.05 standard deviation increase in value-added for ELA ($p < .05$).

Exhibit 2. Change in Value-Added Associated With an Increase of 0.1 in the Alignment Index
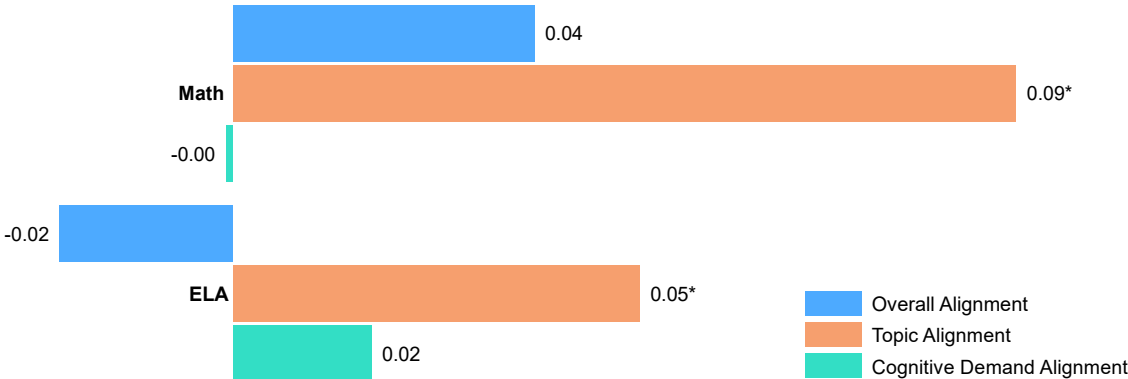
Notes: Sample size for math analyses = 192 teacher-year observation, for 135 unique teachers, in 51 schools. Sample size for ELA analyses = 128 teacher-year observations, for 93 unique teachers, in 46 schools.
* p<.05.

Teachers' Differentiation of Topics and Cognitive Demands

Our second set of analyses examined the construct validity of the SEC-based alignment measures used in the FAST study. Specifically, we examined the extent to which teachers properly differentiated topics and cognitive demands in reporting the content of their instruction. If for example, teachers reported different levels of emphasis for topics but not cognitive demands, it may mean that teachers (a) did, indeed, give equal emphasis to different cognitive demands in their instruction or (b) did not differentiate between cognitive demands in their reporting because they did not understand the distinctions among the cognitive demands or did not make an effort to carefully recall and accurately report the extent to which they emphasized different cognitive demands for each topic. The latter has implications for the validity of SEC-based instruments that require teachers to report on the emphasis they gave to different cognitive demands when teaching a specific topic.

To examine the extent to which teachers differentiated among topics and cognitive demands in their reporting, we first calculated the average proportion of emphasis that teachers reported on each topic category across cognitive demands and the average proportion of emphasis that teachers reported on each cognitive demand across topic categories.[1] To provide a point of reference, we also calculated the corresponding average proportion of emphasis for each topic category and cognitive demand in state standards. We found that teachers reported different levels of emphasis for different topic categories but similar levels of emphasis for all cognitive demands. Furthermore, the levels of emphasis that teachers reported were generally consistent with those in the state standards for topic categories, but not for cognitive demands.

Our next investigation focused on the extent to which teachers differentiated different topics for a given cognitive demand and differentiated different cognitive demands for a given topic. Again, we compared the patterns of findings based on teachers' reporting with those in state standards. We found that teachers differentiated topics within cognitive demands but did not differentiate cognitive demands within topics when reporting on their instruction. In fact, teachers generally reported emphasizing different cognitive demands evenly within topics. Overall, teacher-reported instructional emphasis reflected less differentiation than the state standards, particularly for emphasis on cognitive demands.

The lack of differentiation by cognitive demand as reported by teachers raises concern about how well the FAST Instructional Survey can elicit accurate responses about teachers' instructional emphasis on cognitive demands. The survey is based on the assumption that teachers are able to properly differentiate topics as well as cognitive demands when reporting their instruction. If teachers did not differentiate the different cognitive demands and, essentially, made random selections for cognitive

---

[1] Topic categories are made up of the fine-grained topics on the SEC. In Exhibit 1, for example, "number concepts" is a topic category and "place value" and "whole numbers" are fine-grained topics within that category.

demands when responding to the FAST survey, it would call into question the validity of both the overall alignment index and the cognitive demand alignment index that rely on accurate reporting of teachers' instructional emphasis on different cognitive demands.

Given this concern, in our final set of analyses, we examined the extent to which the alignment measures based on teachers' reports differed from measures that we would have observed had teachers randomly reported their emphasis by topic and cognitive demand on the FAST Instructional Survey. To conduct this analysis, we created 1,000 simulated teachers and randomly assigned levels of emphases to each topic and cognitive demand pair for each teacher. We then computed the overall alignment, topic alignment, and cognitive demand alignment index values for each simulated teacher and compared those values to the observed alignment index values of the teachers in our study; Exhibit 3 depicts the results.

Exhibit 3. Distributions of Alignment Indices as Reported by Teachers Versus Distributions Based on Randomly Simulated Data
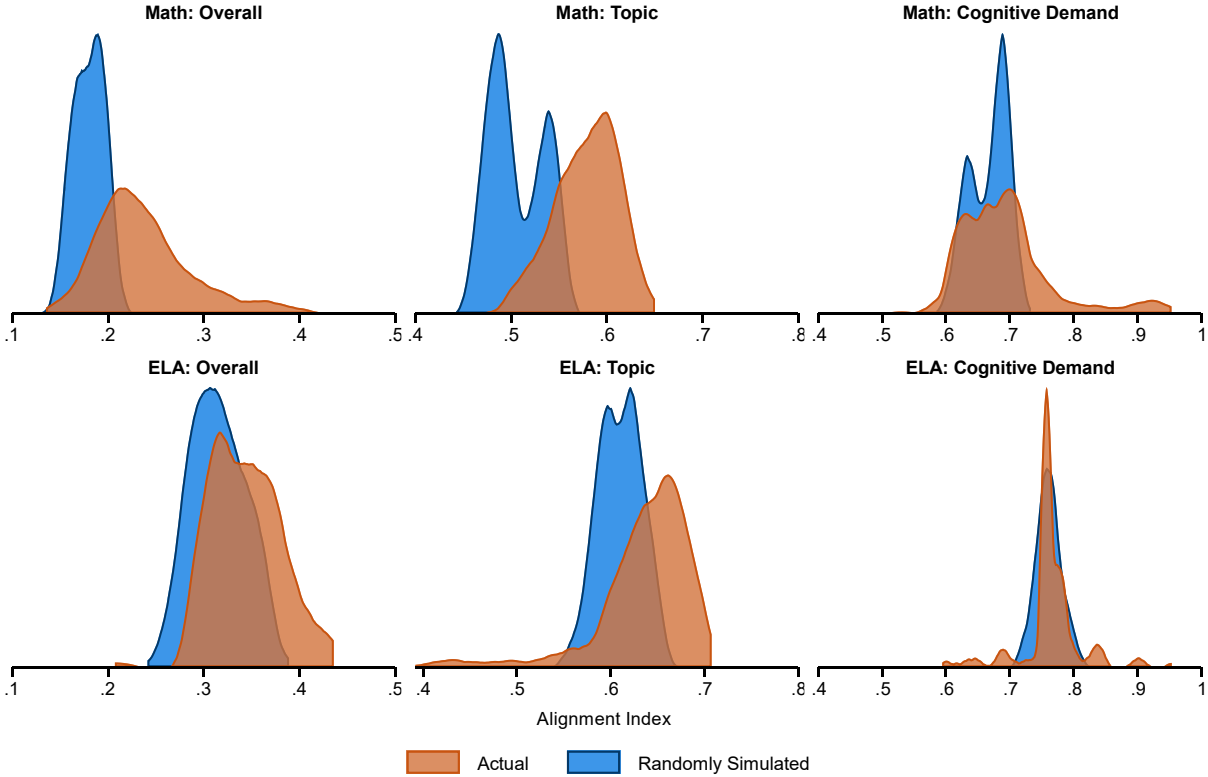
Notes: Sample size for math analysis = 192 teacher-year observation, for 135 unique teachers, in 51 schools. Sample size for ELA analysis = 128 teacher-year observations, for 93 unique teachers, in 46 schools.

As Exhibit 3 shows, for a large portion of teachers in both math and ELA, their actual cognitive demand alignment indices fell within the range of where they would be expected to score had they selected their emphasis levels at random. In ELA in particular, the average of teachers' actual cognitive demand

alignment index values was not significantly different from that based on simulated data ($p > .05$ based on a t test).In contrast, the distribution of teachers' topic alignment for both subjects clearly is shifted to the right of the topic alignment distribution based on randomly simulated data. This suggests that most teachers' alignment by topic was higher than what it would have been had they randomly chosen levels of emphases when responding to the FAST survey. The same is true for overall alignment. This analysis contributes additional evidence that teachers did not properly differentiate cognitive demands in their reporting.

Implications

This study highlights some of the challenges associated with assessing the extent to which teachers align their instruction with state standards. In particular, the overall alignment index, constructed based on teacher-reported emphasis on both topics and cognitive demands, was less predictive of teacher value-added than an alignment index constructed based on instructional emphasis by topic alone. Further, several analyses cast doubt on whether teachers in this study were able to differentiate different types of cognitive demands in a self-report survey of content coverage of their instruction.

Future work on the use of SEC-based instruments to measure alignment could explore several questions such as the following:

- What are the limitations and affordances of different ways to ask about cognitive demands in SEC-based instruments? The FAST Instructional Survey used one approach, but there may be others. For example, instead of reporting on the level of emphasis given to topic and cognitive demand pairs, teachers may first report the number of days spent on each topic, then indicate the level of emphasis given to each cognitive demand for each topic. Teachers may be better able to differentiate their reporting on emphasis given to cognitive demands if they are decoupled from the topics. Teachers may also be able to more accurately recall their instructional emphasis if they are asked to report on their instruction over shorter time periods (e.g., a semester or a month rather than a whole school year).

- Could instructional alignment be measured better by using teacher self-reported data on topic coverage combined with classroom observations of the extent to which teachers emphasize different cognitive demands? This approach would not enable analyses at the intersection of topics and cognitive demands, but it may offer a more valid measure of the extent to which teachers emphasize different cognitive demands in their instruction.

- To what extent do teachers need support to understand the distinctions among the cognitive demands prior to completing SEC-based instruments? Some of the teachers in our study participated in the FAST program and worked closely with a coach to use the topics and cognitive

demands to reflect on the alignment of their instruction with state standards. Others did not. Those teachers may have needed more support to understand the differences between different cognitive demands so that they could better differentiate among them in their reporting.

Conclusion

Our analyses raise questions regarding approaches to measuring the alignment of teachers' instruction with state standards. We have suggested a few potential directions for future work aimed at addressing the challenges that surfaced through this work with the aim of measuring instructional alignment with improved validity.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95–135.

Hannaway, J. (2003). Accountability, assessment, and performance issues: We've come a long way … or have we? In W. L. Boyd & D. Miretzky (Eds.), *American educational governance on trial: Change and challenges* (pp. 20–36). National Society for the Study of Education.

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review, 47*, 180–195.

Polikoff, M. (2021). *Beyond standards: The fragmentation of education governance and the promise of curricular reform.* Harvard Education Press.

Polikoff, M. S. (2012). Instructional alignment under No Child Left Behind. *American Journal of Education*, *118*(3), 341–368. http://www.jstor.org/stable/10.1086/664773

Porter, A. (2000). Doing high-stakes assessment right. *School Administrator, 57*(11), 28–31. https://eric.ed.gov/?id=EJ619580

Porter, A. C., & Smithson, J. L. (2001). *Defining, developing, and using curriculum indicators.* Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education.

AIR®