

## *Research Brief: Implementing Wit & Wisdom with Fidelity*

### June 2021

*This brief examines the challenges of implementing a new curriculum and offers promising evidence of the short-term impacts of implementing Wit & Wisdom with fidelity on students' early literacy outcomes in 1<sup>st</sup> through 3<sup>rd</sup> grade.*

### **Challenges with Fidelity and Time**

Research suggests that a well-implemented curriculum can ensure more coherent instruction and better allocation of instructional time, and also can encourage teacher collaboration.<sup>1</sup> However, implementing a new curriculum is challenging for teachers, as well as for state, district, and school leaders. For example, districts might purchase new curriculum, but not fully support its use (e.g., through encouraging adequate building-level buy-in and supporting teachers with aligned professional development).<sup>2</sup> On the other hand, teachers may not embrace the new curriculum, or use the new materials unevenly. In addition, school leaders may not fully understand that the full impact of a new curriculum may take time to realize,<sup>3</sup> nor anticipate the challenges and possible, often temporary, dips in student outcomes.<sup>4</sup> Therefore, schools may panic and halt implementation before the new curriculum's full potential has been reached. In short, merely purchasing high-quality materials does not ensure that teachers will use the materials with fidelity, or at all.<sup>5</sup>

Most American teachers do not currently use a published curriculum for the great majority of their teaching needs. For example, in a nationally representative sample of American teachers,<sup>vii</sup> 99% of elementary ELA teachers stated that they used ELA instructional materials “developed and/or selected myself,” even as 96% also reported mixing them with those developed or selected by their district. The same sample found that 90% of elementary ELA teachers reported using self-selected or self-created materials at least weekly, compared with 74% of these teachers who used their district materials at least weekly. Many teachers frequently use a combination of materials from different sources to plan their lessons, instead of following a single published curriculum endorsed by their school or district.

A recent study found no relationship between a district's selection of high-quality math materials and increased student achievement. A closer look, however, revealed that most teachers did not actually *use* the materials selected by the district: less than 7% of teachers exclusively used their assigned textbooks, and only 46% of teachers used the endorsed materials in more than half of their lessons.<sup>6</sup> In short, adopting a high-quality curriculum does not translate into student success when teachers either do not consistently or appropriately use it.

Research also shows that successfully implementing a new curriculum takes time.<sup>7</sup> A 2019 report from Leading Innovation for Tennessee Education,<sup>8</sup> which details the districts'

process of implementing high-quality instructional materials, shows that at the beginning of implementation, teachers aligned only 4% of their classroom practices to Tennessee’s ELA standards. After three years of implementation, roughly half of the classrooms showed partial or full alignment to state standards. These findings highlight the implementation challenge that inducing, training, and supporting teachers’ effective use of standards-aligned curriculum takes time.

### **The impact of high-fidelity implementation using Wit & Wisdom**

The following section highlights results from a recent study of the first-year impact of implementing the Wit & Wisdom (W&W) curriculum. This study uses a statistical technique called “propensity score matching,” which compares students who were taught using W&W to similar students who were not taught using W&W. The purpose of this part of the study is to compare the differences in reading outcomes between these two groups.

The findings below use two different sets of W&W impact estimates. The first set is the average impact of W&W on students across all implementing schools. The second set of estimates are the average impact of W&W on students among teachers who report using W&W with *high fidelity*.

This second set of estimates is necessary, because prior research shows that the impact of a curriculum depends – unsurprisingly - on the extent to which a teacher uses it.<sup>9</sup> In our teacher surveys for this study, we find that while almost all teachers in the adopting schools used W&W (i.e., 95% of teachers in schools that adopted W&W used W&W materials to plan their ELA lessons, and 86% of teachers in W&W schools used W&W as their *primary* curriculum), many teachers struggled to teach W&W with fidelity during the first year of implementation. For example, only 40% of teachers in W&W-implementing schools reported that they use every component of W&W “always,” and classroom observers saw virtually no evidence of some W&W components, including concluding lesson elements (i.e., land, wrap, or deep dives) and writing instruction.

In Figure 1 below, the fidelity measure is based on teacher survey responses to the following question:

Figure 1

In a typical month, how often do you use Wit & Wisdom for each of the following components of instruction?

	Never	Sometimes	About half the time	Most of the time	Always
Learning Goals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Content Framing Questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Essential Questions and/or Focusing Questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Craft Questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Text Dependent Questions (TDQs) or task	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Informal and Formal Assessments (including daily Check for Understanding)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Writing Instructions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

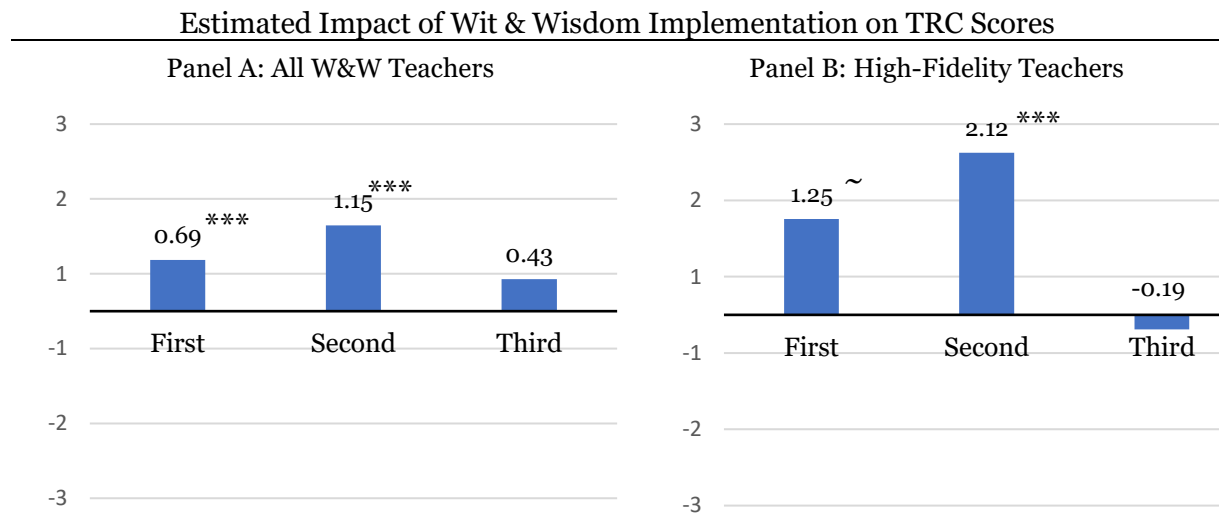
Each W&W component in the question is scored from 1, if a teacher never uses the component, to 5, if the teacher always uses the component. Teachers in grade bands (i.e., K-1 and 2-3) with an average score of 30 or higher are defined as “high-fidelity” users of W&W. That is, teachers must report using at least five of the main W&W components “Most of the time” or “Always” to be a high-fidelity user. High-fidelity teachers comprised roughly 67% of teachers in W&W-implementing schools, on average.

Student outcomes are measured using the TRC, a leveled reading assessment administered three times a year. The TRC is used to determine a student’s reading level in kindergarten through 5<sup>th</sup> grade, and combines both decoding and reading comprehension. Reading levels<sup>1</sup> are reported as a rating from A-Z and a proficiency designation of “instructional” or “independent” within that alphabetical level. The proficiency levels indicate that students need instructional support to fully comprehend and accurately read a book at an alphabetical reading level (i.e., “instructional”), or that students can accurately read and fully comprehend the book independently (i.e., “independent”). Therefore, a student reading at an instructional level (e.g., instructional G) is not as strong a reader as one reading at the same level independently (e.g., independent G). As such, a 1-point score increase in this TRC measure indicates a move from instructional to independent within the same alphabetical reading level, or a move from independent in one alphabetical reading level to instructional in the next level.

Figure 2 (below) shows the estimated treatment effects of W&W on students’ early literacy skills, as measured by TRC scores, in comparison to matched students. The estimated effects of W&W across all implementation schools are presented in Panel A. Panel B presents the estimated effects of a high-fidelity W&W implementation. Each panel presents separate estimate for 1<sup>st</sup>-, 2<sup>nd</sup>-, and 3<sup>rd</sup>-grade students.

<sup>1</sup> After two pre-reading levels: Print Concepts and Reading Behaviors.

Figure 2



NOTES—1. The statistical significance refers to the difference in the average student achievement between comparison and Wit & Wisdom students. ~ $p < .10$ , \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

Panel A shows that the W&W implementation across all classrooms is associated with positive or null effects on students' early literacy, as measured by TRC scores. Specifically, the implementation is associated with a significant increase in 1<sup>st</sup>- and 2<sup>nd</sup>-grade TRC scores, and no significant change in 3<sup>rd</sup>-grade scores.

For example, in 1<sup>st</sup> grade, the implementation of W&W is associated with almost a 1-point (i.e., 0.69) increase in a students' TRC scores. This is equivalent to a change from either instructional to independent within an alphabetical reading level (e.g., instructional G to independent G) or between alphabetical reading levels (e.g., independent G to instructional H). Panel A suggests that 1<sup>st</sup>- and 2<sup>nd</sup>-grade students in classrooms where their teachers used W&W had TRC composite scores that were roughly 1-point higher than they would have been if their teachers had not used W&W, and 3<sup>rd</sup>-grade students had approximately the same TRC scores whether or not their teachers used W&W.

However, the findings presented in Panel B suggest W&W has a generally more positive impact on students when their teachers use the curriculum with high fidelity. Panel B shows that in high-fidelity W&W classrooms, students experienced significant gains in 1<sup>st</sup> and 2<sup>nd</sup> grades, although no significant change in 3<sup>rd</sup> grade. For example, Panel B shows that 1<sup>st</sup>-grade students experienced more than a 1-point (1.25) increase in their TRC score when their teachers used W&W with high fidelity over what they would have experienced had their teachers not used W&W with fidelity or at all. Also, note that while 3<sup>rd</sup>-grade scores in Panel B appear negative, this estimated change is not statistically different from zero.

The results in 3<sup>rd</sup> grade are puzzling. Further investigation showed that teacher experience has no relationship to 3<sup>rd</sup>-grade teachers' use of W&W with fidelity. In addition, our classroom observations do not illuminate systematic differences in 3<sup>rd</sup>-grade classrooms during this first year of implementation. We do wonder whether the results in 3<sup>rd</sup> grade are related to the fact that students begin taking the state test during that grade, and how that impacts teachers' instruction. However, further investigation is needed in order to better understand these 3<sup>rd</sup>-grade results.

## Conclusion and further discussion

Ultimately, switching to high-quality curricula and using them with fidelity is a cost-effective reform that benefits both teachers and students. However, teachers must use the curriculum with fidelity to realize maximum results. While the impacts presented above make the case for fidelity of use, it is also worth recalling that these are first-year, and thus, short-term impacts. Given the literature cited above, we would expect these results to increase with time, as more teachers move towards greater fidelity. Schools and districts must support teachers in their use of the curriculum and be patient, as new curriculum implementation takes time to reach its full potential.

---

<sup>1</sup> Chyrs Dougherty, C. (2016). Keeping track of improvements in educational practices. Iowa City, IA: ACT.

<sup>2</sup> Ruth Chung Wei et al., “Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad.” (Dallas, TX: National Staff Development Council, 2009).

<sup>3</sup> Ira Nichols-Barrer and Joshua Haimson, “Impacts of Five Expeditionary Learning Middle Schools on Academic Achievement” (Washington DC: Mathematica Policy Research, June 8, 2013), <https://www.mathematica.org/our-publications-and-findings/publications/impacts-of-five-expeditionary-learning-middle-schools-on-academic-achievement>.

<sup>4</sup> Sharon Roberts et al., “Instructional Materials Implementation Guidebook, 2019,” 2019, <https://lifteducationtn.com/wp-content/uploads/2019/11/LIFT-Guidebook-2019-FINAL.pdf>.

<sup>5</sup> David Blazar et al., “Learning by the Book: Comparing Math Achievement Growth by Textbook in Six Common Core States,” March 2019, [https://cepr.harvard.edu/files/cepr/files/cepr-curriculum-report\\_learning-by-the-book.pdf](https://cepr.harvard.edu/files/cepr/files/cepr-curriculum-report_learning-by-the-book.pdf); V. Darleen Opfer, Julia H. Kaufman, and Lindsey E. Thompson, “Implementation of K–12 State Standards for Mathematics and English Language Arts and Literacy,” *Santa Monica, CA: RAND*, 2016; Thomas Kane et al., “Teaching Higher: Educators’ Perspectives on Common Core Implementation,” 2016, <https://cepr.harvard.edu/files/cepr/files/teaching-higher-report.pdf>.

<sup>6</sup> David Blazar et al., “Learning by the Book: Comparing Math Achievement Growth by Textbook in Six Common Core States.”

<sup>7</sup> David Blazar et al.

<sup>8</sup> Sharon Roberts et al., “Instructional Materials Implementation Guidebook, 2019.”

<sup>9</sup> David Blazar et al., “Learning by the Book” (Cambridge, MA: Harvard University, March 2019), [https://cepr.harvard.edu/files/cepr/files/cepr-curriculum-report\\_learning-by-the-book.pdf](https://cepr.harvard.edu/files/cepr/files/cepr-curriculum-report_learning-by-the-book.pdf).