

Integrating Implementation Research into an Impact Evaluation:  
Understanding Variation in Early Grade Reading Program Effects to Drive Improvement in Haiti

Authors

TJ D'Agostino, Danice Brown Guzman, Paul Perrin, Anasthasie Liberiste-Osirus, and Kate Schuenke-Lucien

This paper contributes to understanding the effectiveness of early grade reading interventions in low-income-country contexts and how and why such interventions vary in their effectiveness across schools. It presents the results from an impact evaluation of an early grade reading intervention in Haiti, which integrates implementation research using a mixed-methods study of factors that explain variation in program effects. Such utilization of implementation research approaches is necessary to understand how interventions work, for whom, and under what conditions, which is vital if programs are to be able to scale effectively, but are exceedingly rare in international education research in foundational skills in low-and middle-income countries.

Some promising solutions have emerged to address learning poverty (World Bank, 2021). For example, recent reviews of evidence point to positive average effects of structured programs in improving students' reading skills (Evans and Popova 2016; Kim et al. 2020). Kim et al.'s (2020) recent meta-analysis reported an average effect size of .30 from 67 studies across 32 countries. Scholars also note, however, wide variation in treatment effects across contexts, which impedes the efficacy and scalability of programs (Bold et al. 2018). Kim et al. (2020) found negative results in approximately 20 percent of studies in their analysis. Even in early grade reading (EGR) programs that demonstrate large average effects, high proportions of students often fail to make progress.

We study program impact of a structured early grade reading program using a randomized controlled trial experimental research design and find significant treatment impact on students' literacy skills, including on students' reading comprehension in Creole, an outcome that is rarely seen in the literature based upon recent meta-analyses (Kim et al., 2020).

We use a mixed-methods design to explore the mechanisms that explain variation in program effects across school sites. We utilize monitoring data from monthly coaching visits to explore the factors that predict higher growth in the treated schools on a range of literacy outcomes using multi-level modeling, considering factors related to how the intervention worked and in what ways as well as site level factors that may moderate program effects.

We then identified the schools with the highest and lowest academic growth on an index of variables for key early grade reading outcomes, and conducted qualitative research with key informants most familiar with the schools to understand what site-level and context related factors were common in the highest and lowest growth schools and to help identify sources of site-level variation in program effectiveness.

Key findings related to implementation fidelity, take-up, and site/context based sources of variation included the following:

- Pacing: We used a unique measure of treatment on the treated related to percentage of target lessons taught by a given date. Students in classrooms that were the furthest behind (due to teacher absenteeism, school closures, or using a combination of the intervention materials and older materials) consistently scored significantly lower than students that were not as far behind;
- Teaching quality: Students in classrooms where the teachers were comfortable with and adhered to the curriculum (on a measure of instructional quality and fidelity of implementation) scored higher on a range of literacy outcomes, including in the first grade cohort;
- Leadership: School leadership was important for fostering teacher buy-in to the program, fostering high student and teacher attendance, and facilitating remedial classes and extra student supports. Surprisingly (and contrary to the intervention's theory of change), we found weak and mixed evidence of school leaders' role in fostering accountability and instructional leadership, as measured by an indicator of their frequency of conducting classroom observation and coaching visits. We theorize about a possible hierarchy of impactful roles and behaviors of school leaders in low-income contexts, namely, in supporting the more foundational elements of opportunity to learn as more vital than focusing on the finer elements of instructional leadership.

We believe that this study is also a model for the field of how implementation research with a mixed methods design can be integrated into an impact evaluation of a large-scale foundational skills program to yield rich insights, a design that is far too rare in the scholarly literature on foundational learning. This allowed the study to explore not only what works, but how and why it works, for whom, and under what conditions. This helped to reveal which schools were performing well and which schools were failing to do so, in response to the treatment, and why, as well as which elements of the intervention's theory of change appeared to be working and which failed to foster learning gains. This approach provides rich - and sometimes surprising - learning that informs how to make structured early-grade reading interventions better as they seek to maintain and enhance their impact as they go to scale.

This study also demonstrates the impact of a promising early-grade reading intervention in Haiti using a randomized control trial evaluation, which yielded reading comprehension gains for students in the participating schools. This outcome is a rare achievement in similar programs globally, and the program achieved these promising results in a challenging context.