SOLUTION SPOTLIGHT

How Effect Size Is Raising the Bar for Edtech Evaluation

Lexia® All for Literacy

Thank you for taking the time to read this Solution Spotlight.

I'm Dr. Liz Brooke, chief learning officer for Lexia® Learning. In this role, I spend a lot of time in conversations with educators about how they evaluate the efficacy of literacy programs. And during these unprecedented times of interrupted learning, it is critical we understand the nuances of what the results of research studies really mean. Inside this Spotlight, you'll learn why one state made the decision to update its CARES/ESSER grant criteria for K–3 reading programs to require strong, moderate, or promising ESSA evidence of effectiveness, AND an effect size of .20 at a minimum. I believe the pairing of these criteria is a potential game changer for how educators evaluate programs.

In this Spotlight, we explain why effect size is an important addition to levels of evidence for evaluating rigor of efficacy, what effect size is, its value in comparing the efficacy of edtech literacy programs head to head, and how to identify effect size studies that don't make the grade. We hope and expect other states will soon adopt these paired criteria as well.

I am following this topic with interest and will continue to bring news and insight to you as effect size and other rigorous aspects of research are adopted as an evaluation criterion for edtech.

Dr. Liz Brooke, CCC-SLP
Chief Learning Officer
Lexia Learning

One State's Shift

Education had already been "disrupted" multiple times before the pandemic, but never as quickly, widely, or radically. Now, education leaders are seizing the moment by building on the disruption caused by the coronavirus and the influx of federal emergency funds to advance educational initiatives already in progress.

In December 2020, the Florida DOE released its High Quality Curriculum for Reading and Civics guidance for CARES funds grants. The guidance required one of the Key Assurances of the grant be that the Local Education Agency (LEA) would:

> **Select a K–3 program and/or an instructional practice aligned to…the science of reading (explicit and systematic instruction in phonological awareness, phonics, fluency, vocabulary and comprehension, as applicable to need) with strong, moderate or promising levels of evidence as defined by ESSA, [that] has an effect size of .20, at a minimum, and meets the needs of the target population.**
>
> **—Florida Department of Education, 2020**

The Florida DOE's expectation that LEAs select literacy programs based on the strength of the evidence as defined by the Every Student Succeeds Act (ESSA) and effect size is a first. Moreover, it represents a potentially seismic shift in how LEAs evaluate literacy edtech programs, especially head to head.

Effectiveness as Defined by ESSA

ESSA's predecessor, the No Child Left Behind (NCLB) Act, defined a program of merit as being built upon "scientifically based research." Although this definition was intended to implement a data-driven system, it ended up being too vague to be much use to educators.

Passed in 2015, ESSA promotes evidence-based programs by ensuring their capacity to produce results and improve outcomes. ESSA levels of evidence reflect the quality, rigor, and statistical significance of research study designs and findings. The kind of evidence described in ESSA has generally been produced through formal studies and research. Under ESSA, there are four tiers, or levels, of evidence:

Tier 1 — Strong Evidence

supported by one or more well-designed and well-implemented randomized control experimental studies.

Tier 2 — Moderate Evidence

supported by one or more well-designed and well-implemented quasi-experimental studies.

Tier 3 — Promising Evidence

supported by one or more well-designed and well-implemented correlational studies (with statistical controls for selection bias).

Tier 4 — Demonstrates a Rationale

practices that have a well-defined logic model or theory of action, are supported by research, and have some effort underway by an SEA, LEA, or outside research organization to determine their effectiveness.

These evidence requirements—despite being an improvement over NCLB standards—still leave educators to determine whether a literacy program is grounded in up-to-date research, proven to be an effective teaching tool, and capable of addressing their students' unique needs.

For educators working to evaluate literacy programs' effectiveness, a requirement of selecting programs that meet an effect size threshold and that meet these levels of evidence under ESSA is a significant breakthrough.

The Value of Effect Size

By placing the emphasis on the size of the effect as opposed to answering the question of effectiveness with the opaque binary of yes/no, educators can understand the strength of the intervention and the potential impact on student outcomes. Effect size is especially helpful when gauging the potential value of two educational tools, as it allows educators to avoid an "apples-to-oranges" comparison that does not provide insight into the expected impact on student learning.

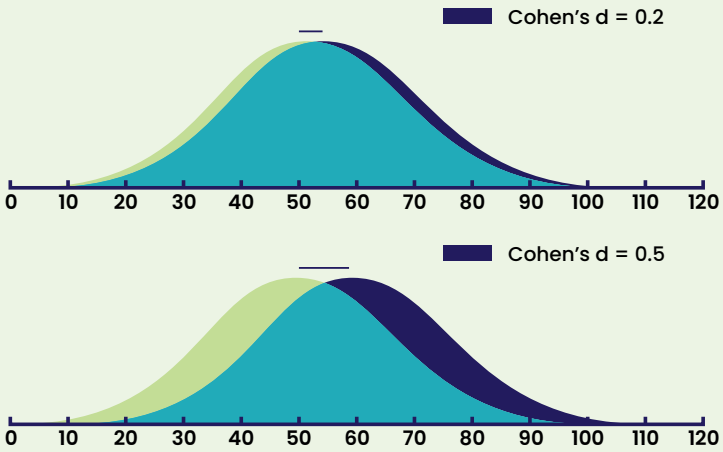
As the University of Michigan explains, "To know if an observed difference is not only statistically significant but also important or meaningful, you will need to calculate its effect size. Rather than reporting the difference in terms of, for example, the number of points earned on a test…effect size is standardized. In other words, all effect sizes are calculated on a common scale—which allows you to compare the effectiveness of different programs on the same outcome."

How Effect Size is Calculated

The Centre for Evaluation and Monitoring defined "effect size" as "simply a way of quantifying the size of the difference between two groups. It is easy to calculate, readily understood and can be applied to any measured outcome in Education…" In more technical words, effect size is the standardized mean difference between two groups:

$$\text{EFFECT SIZE} = \frac{[\text{Mean of experimental group}] - [\text{Mean of control group}]}{\text{Standard deviation}}$$

Here's an example: An effect size of 0.6 means an average student in the intervention group scores 0.6 standard deviations higher than an average student in the control group (that is, the scores of students in the intervention group exceed 73% of the scores of students who did not receive the intervention).



Although the equation shown on the previous page is straightforward, EdWeek has noted that effect-size studies "vary in quality, and many features of studies give hugely inflated estimates of effect sizes." With this in mind, the publication recommended that educators watch for:

- Use of measures made by the researchers
- Very brief studies
- Studies with small sample sizes

Flawed effect-size studies can easily produce effect sizes of +1.00 or more. Such studies should be disregarded by readers who are serious about knowing what does and does not work in real classrooms. Florida's innovation of combining ESSA standards with effect size reduces the risk of relying on effect size alone.

Educators may also need some context for understanding the meaningfulness of the effect size. In 1988, Jacob Cohen, an American psychologist and statistician best known for his work on statistical power and effect size, initially described an effect size of 0.20 as "small," 0.50 as "medium," and 0.80 as "large." The Institute of Education Sciences subsequently issued a report by Mark Lipsey and colleagues that challenged these characterizations, pointing out that in real-life educational experiments with broad measures of achievement and random assignment to treatments, effect sizes as large as +0.50—let alone +0.80—are hardly ever seen, except on occasion in studies of one-to-one tutoring. **This suggests that effect sizes up to 0.50 are a more appropriate range for evaluating educational tools.**

Lexia also supports the work of Evidence for ESSA, an independent organization that evaluates programs using specific criteria derived from the ESSA guidelines. The website was created by the nonprofit Center for Research and Reform in Education at Johns Hopkins University with the goal of providing clear and authoritative information about programs that meet the ESSA evidence standards, and enabling educators and communities to select effective educational tools to improve student success.

Evidence for ESSA is an impartial source for up-to-date, reliable information about programs that meet ESSA evidence standards. By using effect size as an "apples-to-apples" measurement of the size of the intervention impact, educators can make more informed investments with greater confidence.

Research and Effect Size at Lexia

Research is the bedrock of Lexia's educational mission. Founded in 1984 with a grant from the National Institute of Child Health and Human Development (NICHD), Lexia has an ongoing commitment to rigorous efficacy, and learning outcomes research is at the center of our pedagogical approach.

Lexia has the largest impact on student reading outcomes.

Very few edtech providers have programs that have earned "Strong" ratings at BOTH the elementary and secondary levels. Of those, the estimates suggest that Lexia® Core5® Reading and Lexia® PowerUp Literacy® have the largest impact on student reading outcomes, as measured by the average effect sizes.

Lexia CORE5 READING

Lexia's Core5 Reading is one of the most rigorously researched, independently evaluated, and respected reading programs in the world. In an independent review of Lexia Core5 Reading research, Evidence for ESSA awarded it a "Strong" rating, concluding the following:

> **The impact of Core5 was examined in a cluster-randomized study of five schools in the greater Chicago metropolitan area. The study focused on 116 students in grades K–5 receiving special education support for reading difficulties. Students received 'push-in' and/or 'pull-out' support from a special education teacher. After one year, students who used Core5 had significantly higher MAP scores compared to a control group (ES = +0.23), qualifying it for an ESSA 'Strong' rating.**

Lexia POWERUP LITERACY

At the secondary level, Lexia's PowerUp has also received a "Strong" rating from Evidence for ESSA. This study contributes to Lexia's growing body of evidence on PowerUp's efficacy, and shows the program can be up to five times more effective than the average middle school reading intervention. Evidence for ESSA concluded:

> **One cluster-randomized study of PowerUp was conducted with 155 students in grades 6-8 attending one of two Title 1 middle schools. Ten classes of students with reading levels below the 35th percentile were randomly assigned to use PowerUp for 40 minutes of daily supplemental reading instruction or business-as-usual. After one semester, PowerUp students performed significantly higher on the STAR Reading Assessment (effect size = +0.36), qualifying PowerUp for the ESSA "Strong" category.**

	Rating	Effect Size
Core5	Strong and Promising	.23 and .28
PowerUp	Strong	.36

In comparison, the Department of Education estimates that the average middle school intervention has an effect size of .11 and the average elementary intervention has an effect size of .07.

The Future of Effect Size

At Lexia Learning, we welcome the introduction of effect size into states' literacy tool evaluation criteria. The fact that all research is not created equal has long been an issue for educators—**one study does not constitute an evidence base, and research that has not been externally reviewed may be labeled by vendors as more rigorous than it actually is.**

Lexia Learning is hopeful that effect size will continue to be adopted by states and districts as a clear and reliable metric to evaluate literacy programs. We are confident our programs will continue to meet the high standards established by independent organizations such as the National Center on Intensive Intervention (NCII) and Evidence for ESSA.

All students deserve the same opportunity to become successful readers and confident learners. Let us show you how Lexia's research-proven, personalized literacy programs can help.

Check out the research

Lexia®

Lexia is the Structured Literacy expert. For more than 35 years, the company has focused solely on literacy, and today provides a full spectrum of solutions for both students and teachers. With robust offerings for differentiated instruction, personalized learning, assessment, and professional learning, Lexia helps more learners read, write, and speak with confidence.

lexialearning.com

a **cambium** company

v04 03.31.22