



Selecting Evidence-Based Practices

Ohio Department of Education Literacy Academy

Lyzz Davis, PhD

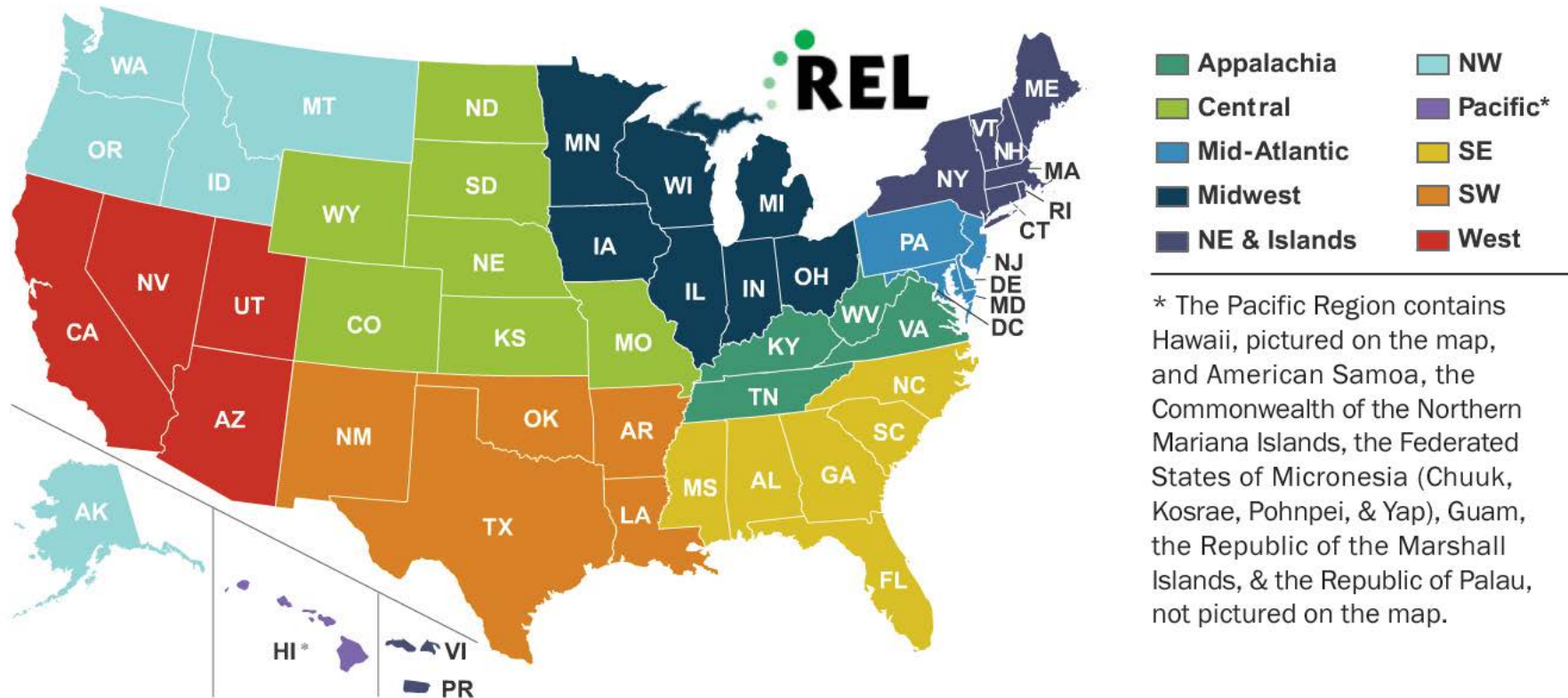
January 17, 2018



Agenda

1. Every Student Succeeds Act (ESSA) evidence tiers
2. Existing resources for identifying interventions
3. Aligning standards
4. Questions

Regional Educational Laboratories



ESSA Tiers of Evidence

Four tiers of evidence under ESSA

Tier 1: Strong Evidence

Tier 2: Moderate Evidence

Tier 3: Promising Evidence

Tier 4: Demonstrates a Rationale

Tier 1: Strong Evidence

“Well-executed” experimental study:

- Uncompromised random assignment:
 - Equal chances of being in treatment or control
 - No adding, switching, or dropping
- Low attrition:
 - How many people left the study after randomization and before the analysis?

Tier 1: Strong Evidence

“Well-executed” experimental study:

- Uncompromised random assignment:
 - Equal chances of being in treatment or control
 - No adding, switching, or dropping
- Low attrition:
 - How many people left the study after randomization and before the analysis?

Statistically significant favorable effect on a **relevant outcome:**

- Studies often examine impact on multiple outcomes

Tier 1: Strong Evidence

- No overriding negative effects from experimental or quasi-experimental studies:
 - Look to resources like What Works Clearinghouse (WWC)
- Large, multisite sample
- Population and setting in the study overlap with your population and setting

Tier 1: Strong Evidence

- No overriding negative effects from experimental or quasi-experimental studies:
 - Look to resources like What Works Clearinghouse
- **Large, multisite sample**
- Population and setting in the study overlap with your population and setting

Tier 1: Strong Evidence

- No overriding negative effects from experimental or quasi-experimental studies:
 - Look to resources like What Works Clearinghouse
- Large, multisite sample
- Population and setting in the study overlap with your population and setting

Tier 2: Moderate Evidence

Well-executed quasi-experimental design:

- **Some type of matching:** Can be through statistical method like propensity scores or not
- **Baseline equivalence:** Treatment and control have similar attributes or scores before the intervention

Tier 2: Moderate Evidence

- Statistically significant favorable effect on a **relevant outcome**
- No overriding negative effects from experimental or quasi-experimental studies:
 - Look to WWC
- Large, multisite sample
- Population and setting in the study overlap with your population and setting

A quick note about ESSA Tiers 1 and 2



Deciding whether a study is “**well designed and well implemented**” for Tiers 1 and 2 requires a review against WWC standards.

Tier 3: Promising Evidence

Well-designed, well-implemented correlational study:

- Uses methods to account for differences between treatment and control groups:
 - Statistical controls for selection bias

Tier 3: Promising Evidence

- Statistically significant favorable effect on a **relevant outcome**
- No overriding negative effects from experimental or quasi-experimental studies:
 - Look to WWC

Tier 4: Demonstrates a Rationale

- Well-specified logic model:
 - How is the intervention **likely** to improve outcomes?
 - Based on previous research
- An effort to study the effects is **currently or will** be underway

Strong Evidence	Moderate Evidence	Promising Evidence
<p>Well-designed and implemented experimental study</p> <p>Significant positive effect on relevant outcome</p> <p>No overriding negative effects from causal studies</p> <p>Large, multisite sample</p> <p>Overlaps with population of interest</p>	<p>Well-designed and implemented QED or RCT with high attrition</p> <p>Significant positive effect on relevant outcome</p> <p>No overriding negative effects from causal studies</p> <p>Large, multisite sample</p> <p>Overlaps with population of interest</p>	<p>Well-designed and implemented correlational study or well-designed and implemented RCT or QED without a large/multisite sample</p> <p>Statistical controls for selection bias</p> <p>Significant positive effect on relevant outcome</p> <p>No overriding negative effects from causal studies</p>

QED is quasi-experimental design, RCT is randomized controlled trial.

Existing resources

Evidence-based clearinghouses focusing on literacy

- What Works Clearinghouse
- Top Tier Evidence
- Blueprints Programs (more on this tomorrow)

To apply ESSA standards, we need to know:

1. **Types** of studies reviewed
2. **Criteria** that factor into ratings
3. How **conflicting outcomes** are handled
4. What **contextual** information is provided



What Works Clearinghouse

WWC: Types of studies reviewed

Experimental designs:

- Randomized controlled trials (RCT)
- Regression discontinuity designs (RDD)*
- Single case designs (SCD)*

Quasi-experimental designs (QED):

- Matched groups

*Criteria are slightly different for these designs.

WWC: Criteria factored into ratings

Study design:

- How were treatment and control groups formed?
- Free of confounding factors?
- Examine an eligible outcome?

Sample attrition:

- How many participants left or dropped out during the study?

Baseline equivalence:

- Were the groups equal on key characteristics before the intervention?

WWC: How conflicting outcomes are handled



Evidence of effectiveness icons (by intervention)

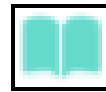
223 Results filtered by:

Literacy x

Evidence of effectiveness ⓘ

Intervention ⓘ

	Literacy Express
	Phonological Awareness



Positive/potentially positive effects



Mixed/no discernable evidence



No evidence



WWC: How conflicting outcomes are handled

Six effectiveness ratings (by outcome):

--	-	0	+ -	+	++
----	---	----------	-----	---	----

WWC: How conflicting outcomes are handled

Six effectiveness ratings (by outcome):

--	-	0	+ -	+	++
----	---	---	-----	---	----

Negative

Potentially negative

No discernable

Mixed

WWC: How conflicting outcomes are handled

Six effectiveness ratings (by outcome):

--	-	0	+-	+	++
----	---	---	----	---	----

Negative

Potentially negative

No discernable

Mixed



ESSA Tiers 1-3

WWC: How conflicting outcomes are handled

Six effectiveness ratings (by outcome):

--	-	0	+ -	+	++
----	---	---	-----	---	----

Potentially positive

Positive

WWC: How conflicting outcomes are handled

Six effectiveness ratings (by outcome):

--	-	0	+ -	+	++
----	---	---	-----	---	----

Potentially positive
Positive

 **ESSA Tiers 1-3**

WWC: Contextual information provided

Summary of evidence page for each outcome:

- Effectiveness rating
- Number of studies meeting standards
- Grades examined
- Number of students
- Improvement index

 INTERVENTION REPORT (587 KB)

 REVIEW

	Grades examined ⓘ	Students ⓘ	Improvement index
	PK	722	--
	PK	185	--
	PK	1,002	-50
	PK	1,004	-50
	PK	999	-50

WWC: Contextual information provided

Evidence snapshots:

- Summary of all research settings and samples from studies meeting standards can include:
 - Race/ethnicity
 - Gender
 - English learners
 - Free or reduced-price lunch
 - Delivery method
 - Locale

Summary of all Research Settings and Samples that Met Standards

Race



Ethnicity



Gender



Male: 50%
Female: 49%

Free & Reduced-Price Lunch



82%

English Learners



28%

WWC: Contextual information provided

Intervention reports go into a lot more detail:

- Program information, including implementation and cost
- All studies reviewed and summary of their findings
- Sample characteristics

READ 180®

Program Description¹

READ 180® is a reading program designed for struggling readers who are reading 2 or more years below grade level. It provides blended learning instruction (i.e., combining digital media with traditional classroom instruction), student assessment, and teacher professional development. READ 180® is delivered in 45- to 90-minute sessions that include whole-group instruction, three small-group rotations, and whole-class wrap-up. Small-group rotations include individualized instruction using an adaptive computer application, small-group instruction with a teacher, and independent reading. READ 180® is designed for students in elementary through high school. This review of READ 180® focuses on students in grades 4–12.

Research²

The What Works Clearinghouse (WWC) identified nine studies of READ 180® that both fall within the scope of the Adolescent Literacy topic area and meet WWC group design standards. Three studies meet WWC group design standards without reservations, and six studies meet WWC group design standards with reservations. Together, these studies included 8,755 adolescent readers in more than 66 schools in 15 school districts and 10 states.

The WWC considers the extent of evidence for READ 180® on the reading achievement of adolescent readers to be medium to large for four outcomes—comprehension, general literacy achievement, reading fluency, and alphabetics. (See the Effectiveness Summary on p. 7 for more details of effectiveness by domain.)

Effectiveness

READ 180® was found to have positive effects on comprehension and general literacy achievement, potentially positive effects on reading fluency, and no discernible effects on alphabetics for adolescent readers.

Table 1. Summary of findings³

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of students	Extent of evidence
		Average	Range			
Comprehension	Positive effects	+6	-4 to +16	6	3,882	Medium to large
General literacy achievement	Positive effects	+4	0 to +7	6	6,235	Medium to large
Reading fluency	Potentially positive effects	+4	+4 to +4	2	561	Medium to large
Alphabetics	No discernible effects	0	-1 to +2	2	746	Medium to large

Report Contents	
Overview	p. 1
Program Information	p. 2
Research Summary	p. 4
Effectiveness Summary	p. 7
References	p. 11
Research Details for Each Study	p. 22
Outcome Measures for Each Domain	p. 39
Findings Included in the Rating for Each Outcome Domain	p. 41
Supplemental Findings for Each Outcome Domain	p. 47
Endnotes	p. 54
Rating Criteria	p. 56
Glossary of Terms	p. 57

This intervention report presents findings from a systematic review of READ 180® conducted using the WWC Procedures and Standards Handbook, version 3.0, and the Adolescent Literacy review protocol, version 3.0.

WWC—Ratings

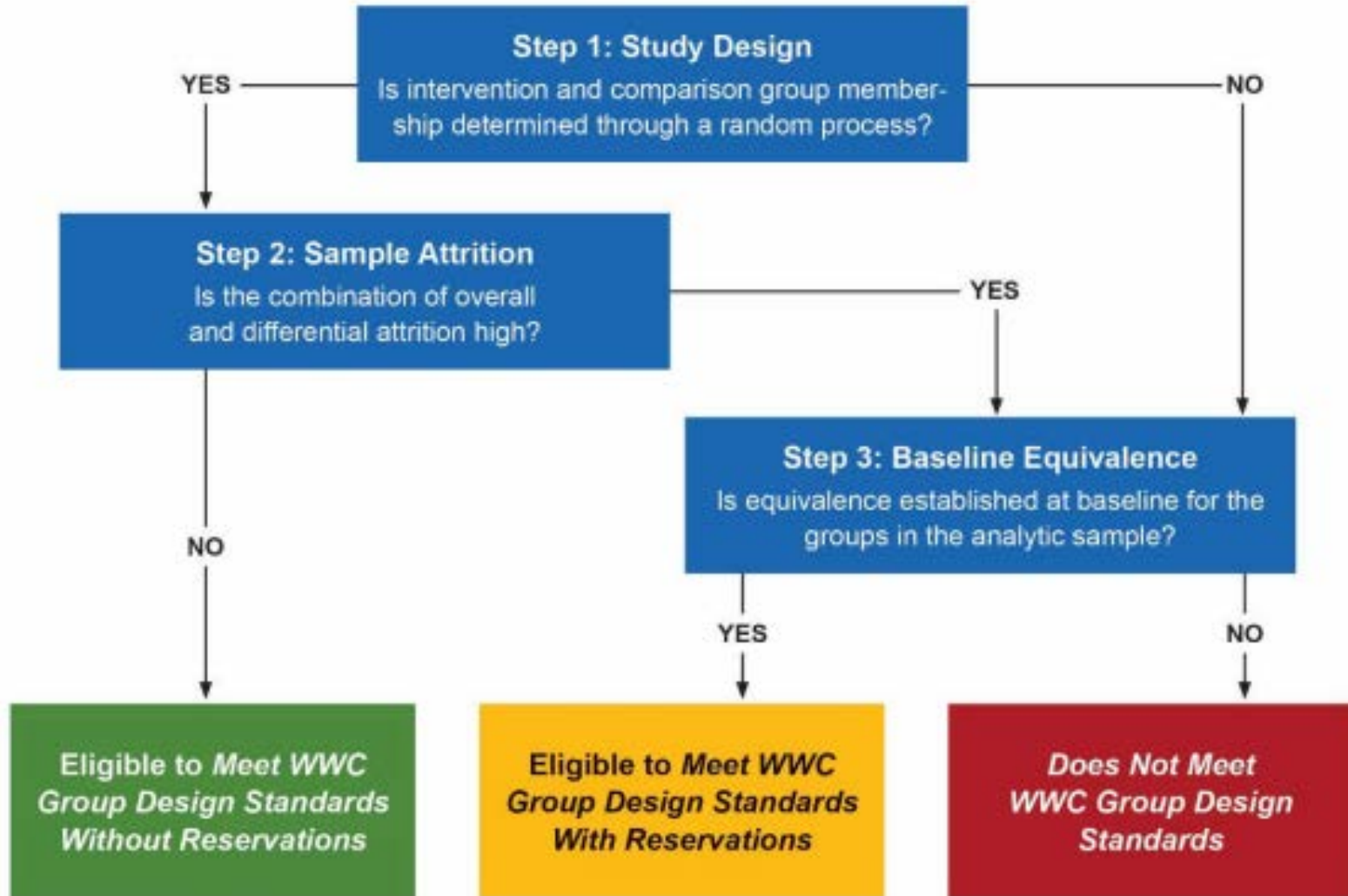
Three ratings for RCTs and QEDs:

- Meets Group Design Standards **Without** Reservations
- Meets Group Design Standards **With** Reservations
- Does **Not Meet** Group Design Standards

Ratings for single case designs and RDD:













- Meets **Without** Reservations
- Meets **With** Reservations
- Does **Not Meet** Standards

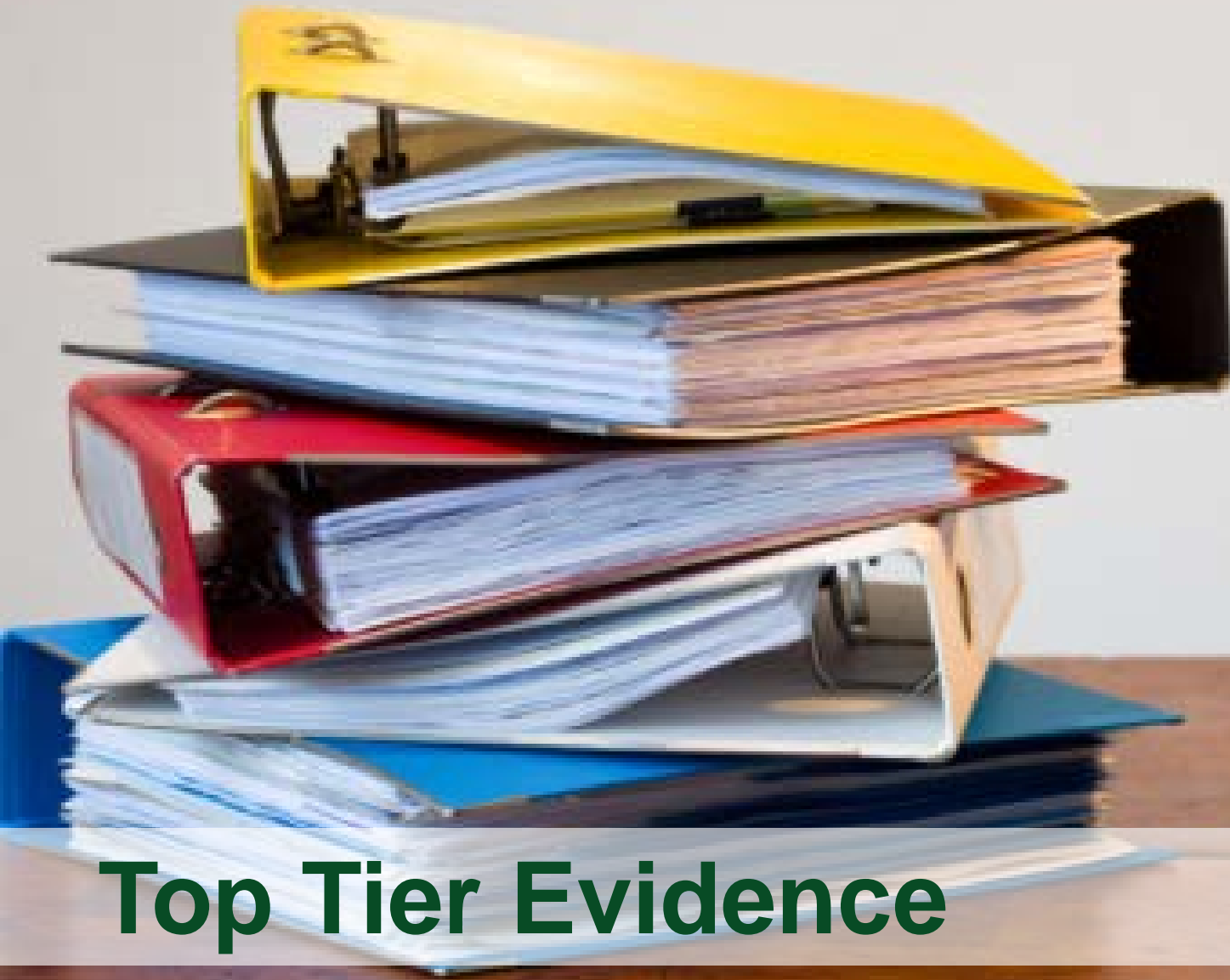
WWC: Criteria that factor into ratings



Source: What Works Clearinghouse. (2014). Procedures and standards handbook (Version 3.0). Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf.

WWC-ESSA alignment

WWC Standard	Positive/ potentially positive	Large, multisite sample	ESSA standard
Meets standards without reservations			Strong Evidence (Tier 1)
			Promising Evidence (Tier 3)
			Does not meet ESSA Tiers 1-3
Meets standards with reservations			Moderate Evidence (Tier 2)
			Promising Evidence (Tier 3)
			Does not meet ESSA Tiers 1-3



Top Tier Evidence

TTE: Types of studies reviewed

Randomized Controlled Trials

TTE: Criteria factoring into ratings

Overall study **design**

Outcome measures

Reporting of intervention **effects**

TTE: How conflicting outcomes are handled

“No strong countervailing evidence”

Applies to **no effects** in addition to **negative effects**

TTE: Contextual information provided

Evidence summaries give information on things like:

- Sample size
- Sample characteristics
- Region of country
- Cost

Reader needs to “dig” for it.



Top Tier Evidence Initiative:

Evidence Summary for Success For All, in Grades K-2

HIGHLIGHTS:

- **Intervention:** A school-wide reform program, primarily for high-poverty elementary schools, with a strong focus on reading instruction.
- **Evaluation Methods:** A large, multi-site randomized controlled trial.
- **Key Findings:** The program increased second-grade reading achievement in Success for All schools by 25-30% of a grade-level, three years after random assignment.
- **Other:** Strong evidence of effectiveness applies to the program as implemented in grades K-2 (as opposed to later elementary school). Per-student program cost is low. Longer-term study follow-up would be desirable to see if effects continue beyond second grade.

I. Finding of the Top Tier Evidence Initiative's Expert Advisory Panel:

Success for All meets the Top Tier Evidence Standard, defined by recent Congressional legislation to include: *Interventions shown in well-designed and implemented randomized controlled trials, preferably conducted in typical community settings, to produce sizeable, sustained benefits to participants and/or society.*

II. Description of the Intervention:

Success for All is a comprehensive school-wide reform program, primarily for high-poverty elementary schools, with a strong emphasis on early detection and prevention of reading problems before they become serious. Key program elements include: (i) daily 90-minute reading classes, each of which is formed by grouping together students of various ages who read at the same performance level; (ii) a K-1 reading curriculum that focuses on language development (e.g., reading stories to students and having them re-tell), teaching students the distinct sounds that make up words (i.e. phonemic awareness), blending sounds to form words, and developing reading fluency; (iii) daily one-on-one tutoring (in addition to the classes) for students needing extra help with reading; and (iv) cooperative learning activities (in which students work together in teams or pairs) starting in the grade 2 reading classes.

The program costs approximately \$120,000 per elementary school (for implementation in grades K-5) in the first year, \$55,000 in the second year, and \$45,000 in the third year, in 2008 dollars. These costs include materials and training; schools may incur additional costs of reallocating staff from other functions (e.g., to provide a higher ratio of tutors).

[Click here to go to the program's website.](#)

III. Evidence of Effectiveness

This summary of the evidence is based on a systematic search of the literature, and correspondence with leading researchers, to identify all well-designed and implemented randomized controlled trials of the Success for All school-wide reform program. Our search identified one such trial, summarized below. Importantly, this trial evaluated the program as implemented in grades K-2 but not grades 3-5; thus, its findings apply only to the K-2 elements of the program.

TTE—Ratings

Top tier standard:




- Well-designed, well-implemented RCTs in replicable settings. Large, sustained effects. Must be large sample* and multisite.

Near top tier standard:

- Meets most top tier standards; only needs replication to qualify.

*Not necessarily large enough to meet ESSA requirements.

TTE-ESSA—Alignment

TTE Standard	Large multisite sample	ESSA Standard
Top Tier		Strong Evidence (Tier 1)
		Promising Evidence (Tier 3)
Near Top Tier		Promising Evidence (Tier 3)

Other resources and clearinghouses

- Blueprints Programs
 - Blueprints for Healthy Youth Development
- Crime Solutions
 - National Institute of Justice
- California Evidence-Based Clearinghouse for Child Welfare
 - California Department of Social Services Office of Child Abuse Prevention
- National Registry of Evidence-Based Programs and Practices
 - Substance Abuse and Mental Health Service Administration
- Home Visiting Evidence of Effectiveness
 - US Department of Health and Human Services

Questions?





Elisabeth (Lyzz) Davis

*AIR Senior
Researcher*

edavis@air.org