

Welcome/Do-Now Activity

Grab a blue dot sticker and indicate your position of comfort and interest on the three graphs on the side wall

- ESSA Evidence Tiers
- What Works Clearinghouse
- Research Methods



ESSA Tiers of Evidence Workshop With the Michigan Data Hub

Lyzz Davis | Matt Linick

05/29/2019



Objectives for today's meeting

1. Discuss the need for and use of evidence under the Elementary and Secondary Education Act (ESSA).
2. Gain a better understanding of the ESSA evidence tiers, and how they align to What Works Clearinghouse (WWC) standards.
3. Dig into specific examples of studies to determine evidence.
4. Practice applying knowledge through a small-group activity.

Today's presenters



Matt Linick

*Senior
Researcher*
mlicnick@air.org



Lyzz Davis

*Senior
Researcher*
edavis@air.org

Tell us about you!

- Name.
- Organizational role.
- If you could go anywhere in the world for one day, where would you go and why?



Agenda

1. The importance of evidence
2. ESSA evidence tiers: Overview
3. Aligning ESSA with existing resources
4. Test your knowledge
5. Small group activity & share out

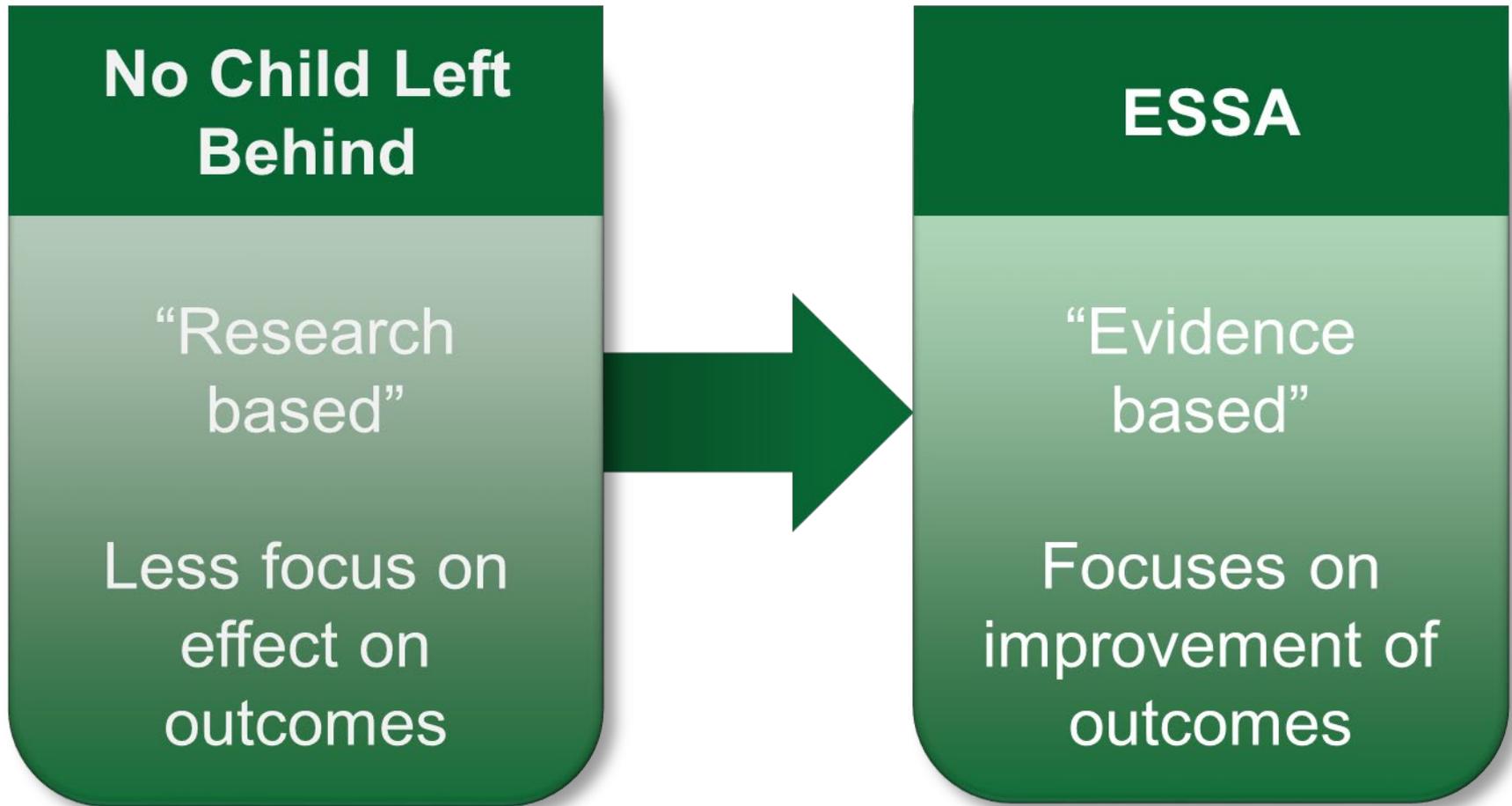
Why do we care about ESSA tiers of evidence?

- Schools identified for targeted supports must implement at least one intervention that meets **promising** evidence.
- Some federal grant programs (such as Striving Readers and Promise Neighborhoods) require interventions that meet **strong** or **moderate** evidence.
- Other activities require interventions that at least **demonstrate a rationale**.

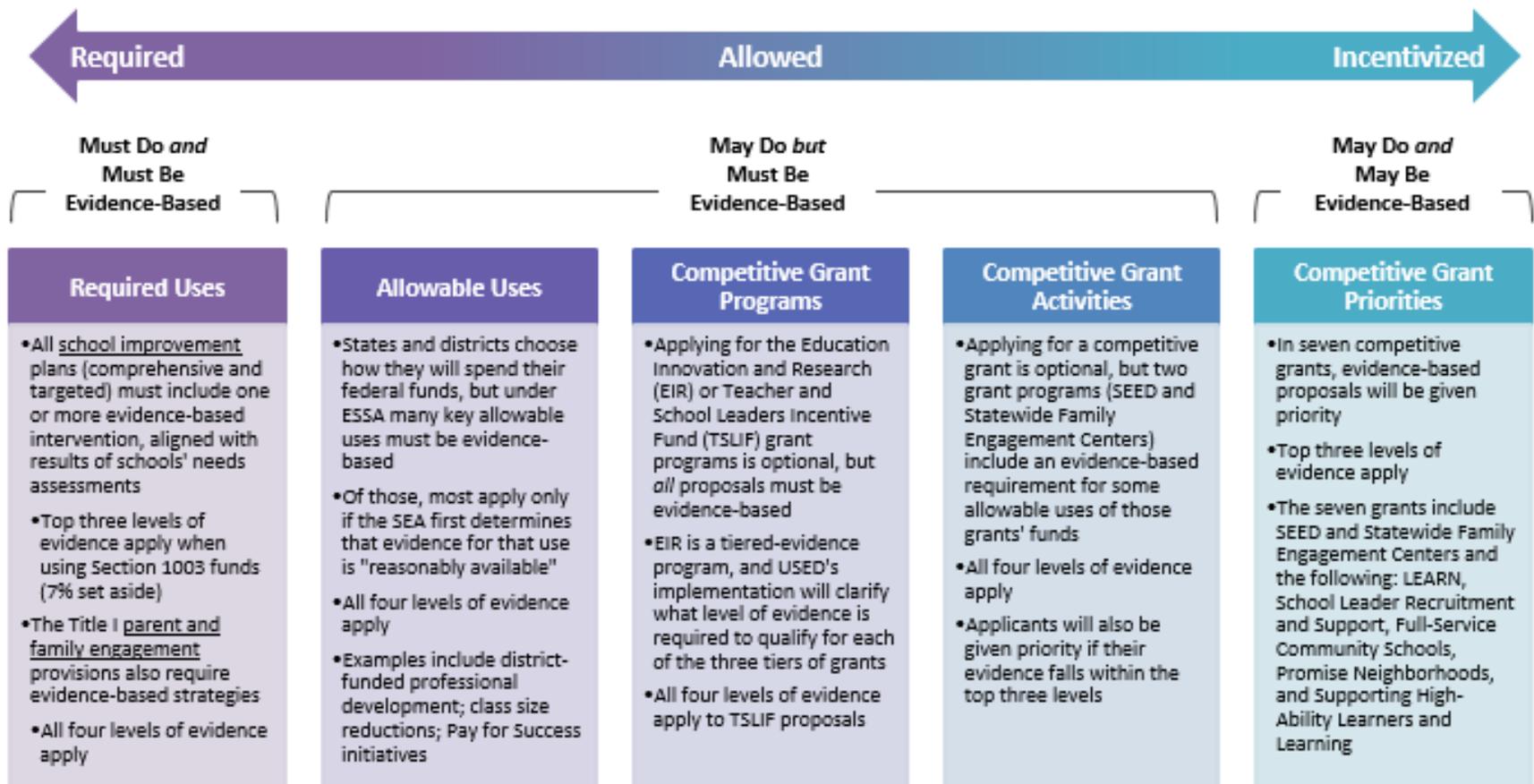
Evidence requirements across federal programs

ESSA program (unless noted)	Evidence requirement(s)
Title I, Section 1003: School Improvement	Minimum of one intervention meets Tier I, II, or III.
Title I, Part A: Schoolwide/ Targeted Assistance	External providers must have expertise in using evidence-based practices (EBPs) (Tier I, II, III, or IV).
Title II, Part A: Effective Instruction	Some requirements for Tier I, II, III, or IV, where evidence is reasonably available (for example, professional development, induction, mentoring).
Title IV, Part A: Student Support Grant	Some requirements for Tier I, II, III, or IV, where evidence is reasonably available.
Title IV, Part B: 21st CCLCs	Tier I, II, III, or IV evidence, when deemed appropriate.
Title IV, Part D: Magnet School Assistance	Competitive preference is given for proposals with evidence-based activities (Tier I, II, III, or IV).
Title IV, Part F: Education Innovation	Includes program-specific evidence requirements.
Title IV, Part F: National Community Support	<ul style="list-style-type: none"> • Promise Neighborhoods: Some requirements and competitive preference for Tier I, II, III, or IV. • Full-Service Community Schools: Competitive preference for Tiers I to IV.
Perkins V (Perkins ACT)	<ul style="list-style-type: none"> • Professional development for career and technology education (CTE) must be EBPs. • Title I Innovation for CTE proposal must include EBPs.

Why is it important to focus on evidence?



Using the ESSA tiers of evidence



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

What's the difference?

ESSA tier	Group formation	Group equivalence
1		
2		
3		

ESSA Tier 1

Strong Evidence

Key terms

Treatment group

Receives the intervention, practice, strategy, or program. Also known as intervention group.

Control group

Does not receive the intervention, practice, strategy, or program.

Key terms

Random assignment

- Method of assigning people (or schools) to the treatment and control groups.
- Helps ensure the two groups are as similar as possible before intervention.
- Must take place before groups are formed and before intervention begins.

Key terms

Attrition

Total percentage of participants who left the study after random assignment.

Differential attrition

The percentage point difference between attrition in the treatment group and attrition in the control group.

Key terms

Statistically significant effect

To understand this, first we should ask:

What is a p-value?

The “p” stands for “probability”—that is, the probability that there is no difference between groups.

So.....

Key terms

Statistically significant effect

A 95 percent (or higher) chance that there is a difference between the two groups, OR

A 5 percent (or lower) chance that there is no difference.

Example: Grade 3 students who participated in a new mathematics program had significantly higher standardized test scores (M = 361) than students who did not participate (M = 352; $p < 0.05$).

Key terms

Confounding factor

A factor other than the intervention that is unique to either the treatment group or the control group.

*To determine whether an intervention causes an outcome, we need to be sure that the **intervention is the only difference** between the groups.*

Example: All the intervention students are taught by one teacher, and there is no way to distinguish between the effect of the intervention and the effect of the teacher.

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?

NOTE: This criteria aligns with WWC's *Meets Standards Without Reservations*.

Tier 1: Strong Evidence

Statistically significant favorable effect on a relevant outcome:

- Studies often examine impact on more than one outcome.

No overriding negative effects from experimental or quasi-experimental studies

- Look to WWC to find this information.

Tier 1: Strong Evidence

Large sample

- At least 350 participants in the sample.

Multisite sample

- Study was conducted in more than one school.

NOTE: Samples and settings can be combined across studies to meet these criteria.

Both population and setting in the study are similar to your population and setting.

What's the difference?

ESSA tier	Group formation	Group equivalence
1	Random (equal chance of assignment)	Assumed
2		
3		

ESSA Tier 2

Moderate Evidence

Key terms

Nonequivalent groups

Treatment and control groups created using assignment that is nonrandom.

Matching

Using statistical methods to create treatment and comparison groups (rather than random assignment).

Key terms

Before and after intervention groups

Using time to create treatment and control groups.

- Control group: Before intervention is implemented.
- Treatment group: After intervention is implemented.

Baseline equivalence

Establishing that the treatment and control groups are **similar on key measures** before the intervention began.

Tier 2: Moderate Evidence

Well-executed quasi-experimental design

- **Group formation:** Can be through matching, nonequivalent groups, or before and after.
- **Baseline equivalence:** Treatment and control are similar on key measures before the intervention was implemented.

NOTE: This criteria aligns with WWC's *Meets Standards With Reservations*.

Tier 2: Moderate Evidence

Statistically significant favorable effect on a relevant outcome

- Studies often examine impact on more than one outcome.

No overriding negative effects from experimental or quasi-experimental studies

- Look to WWC to find this information.

Tier 2: Moderate Evidence

Large sample

- At least 350 participants in the sample.

Multisite sample

- Study was conducted in more than one school.

NOTE: Samples and settings can be combined across studies to meet these criteria.

Either population or setting in the study are similar to 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.

What's the difference?

ESSA tier	Group formation	Group equivalence
1	Random (equal chance of assignment)	Assumed
2	Nonrandom but purposeful	Establish the two groups are statistically similar on key characteristics before the intervention (baseline equivalence)
3		

ESSA Tier 3

Promising Evidence

Key terms

Selection bias

When people “self-select” into an intervention, they may have systematically different characteristics than those who don’t self-select.

Example: Students with higher grade-point averages (GPAs) may be more likely to self-select into a dual-enrollment course than students with lower GPAs, or be more likely to be encouraged by faculty to take the course.

Key terms

Statistical controls

Accounting for factors that could influence the outcome other than the intervention.

Example: Accounting for GPA, race/ethnicity, ACT/SAT scores, gender, and parent and teacher expectancy when examining the association between enrolling in dual-credit courses in high school and college outcomes.

Tier 3: Promising Evidence

Well-designed, well-implemented correlational study

- Uses statistical controls to account for differences between treatment and control groups.

OR

A study that otherwise would meet Tier 1 or Tier 2, but does not meet the large/multisite sample requirement or the population/setting overlap requirement.

Tier 3: Promising Evidence

Statistically significant favorable effect on a relevant outcome

- Studies often examine impact on more than one outcome.

No overriding negative effects from experimental or quasi-experimental studies

- Look to WWC to find this information.

What's the difference?

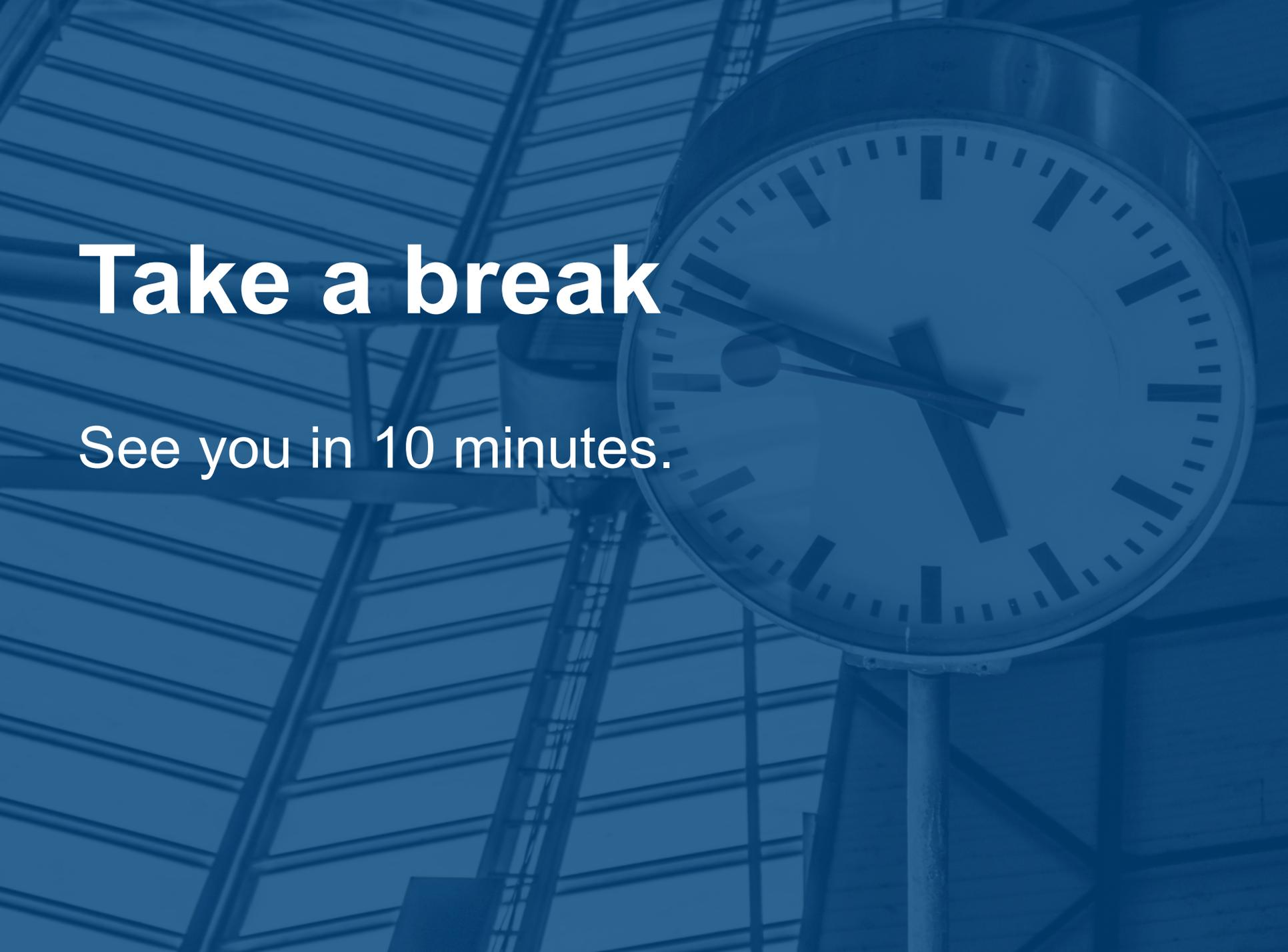
ESSA tier	Group formation	Group equivalence
1	Random (equal chance of assignment)	Assumed
2	Nonrandom, but purposeful	Establish the two groups are statistically similar on key characteristics before the intervention (baseline equivalence)
3	Nonrandom, not purposeful	No baseline equivalence, but statistically control for selection bias

ESSA Tier 4

Demonstrates a Rationale

Tier 4: Demonstrates a Rationale

- A well-specified logic model that explains how intervention is likely to improve outcomes.
- Supported by rigorous research in the field.
- An effort to study the effects is **currently or will be under way**.



Take a break

See you in 10 minutes.

Aligning ESSA evidence tiers with WWC design standards



New WWC Video!

<https://www.youtube.com/watch?v=hu4XnpyiKxw>

What do we need to know?

WWC rating

Outcomes

Sample size

Setting(s)

Context



What do we need to know?

WWC rating

- Meets Design Standards **Without** Reservations
Remember? Aligns with “well-designed, well-implemented experimental study.”
- Meets Design Standards **With** Reservations
Remember? Aligns with “well-designed, well-implemented quasi-experimental study.”

What do we need to know?

Outcomes

For each outcome of interest:

1. Is there a statistically significant favorable effect on a relevant outcome?
2. Are there countervailing negative effects from causal studies?

Note: WWC reports on all relevant outcomes in studies, and each one can have its own ESSA tier of evidence.

What do we need to know?

Sample size

- Were at least 350 students included in the sample(s)?

Location

- Was more than one district included in the study (or studies)?

Setting and population

- Is the study's setting and/or population similar to the district considering implementation?

Remember: You can pool across different studies that examine the same intervention on the same outcome.

ESSA Tiers 1 and 2 on the WWC

- WWC lists ESSA Tiers 1 and 2 for qualifying findings:
 - Studies reviewed under WWC standards 2.1 and higher (i.e., not for 1.0 or 2.0)
 - Findings meet standards with or without reservations
 - Significant favorable effect with no significant negative effects from other studies
 - Large/multisite samples

Reviews of Individual Studies – Review Details



Review Details | Findings | Sample Characteristics | Study Details | Additional Sources

Reviewed: May 2017

For:

- Saxon Math Intervention Report - Primary Mathematics

Using:

- Primary Mathematics Review Protocol 3.1
- Review Standards 3.0

Rating:

Meets WWC standards with reservations because it is a compromised randomized controlled trial, but the analytic intervention and comparison groups satisfy the baseline equivalence requirement.

MEETS WWC STANDARDS WITH RESERVATIONS

AT LEAST ONE STATISTICALLY SIGNIFICANT POSITIVE FINDING

ESSA TIER 2

AT LEAST ONE FINDING SHOWS MODERATE EVIDENCE OF EFFECTIVENESS

Reviews of Individual Studies – Findings



Review Details **Findings** Sample Characteristics Study Details Additional Sources

General Mathematics Achievement outcomes—Indeterminate effects found ⓘ

Outcome measure ⓘ	Comparison ⓘ	Period ⓘ	Sample ⓘ	Intervention mean ⓘ	Comparison mean ⓘ	Significant? ⓘ	Improvement index ⓘ	ESSA rating ⓘ
Early Childhood Longitudinal Study-Kindergarten (ECLS-K): 2nd grade adaptation	Saxon Math vs. Investigations in Number, Data, and Space®	2 Years	Grade: 2; 882 students	71.72	67.31	Yes		

+ More Outcomes

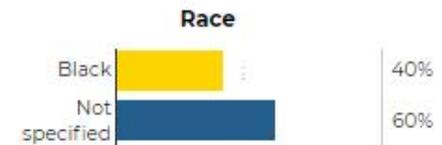
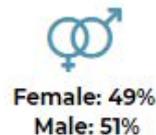
Source: <https://ies.ed.gov/ncee/wwc/Study/81830>

Reviews of Individual Studies – Sample Characteristics



Review Details | Findings | **Sample Characteristics** | Study Details | Additional Sources

Characteristics of study sample as reported by study author.

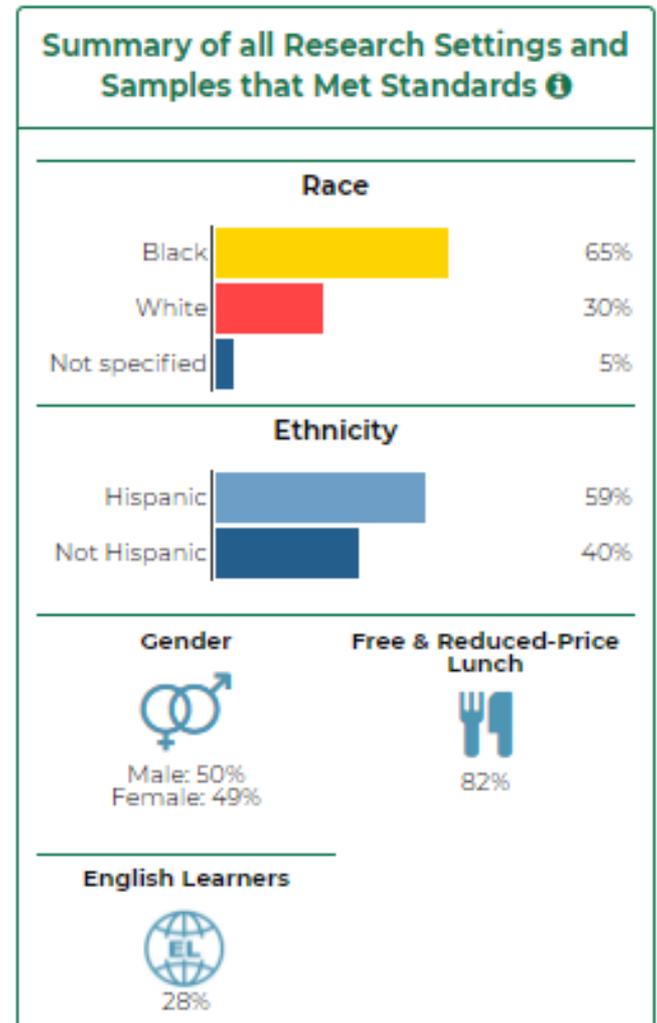


Source: <https://ies.ed.gov/ncee/wwc/Study/81830>

WWC: Contextual information provided

Evidence snapshots:
Summary of research
settings and samples can
include:

- Race/ethnicity.
- Gender.
- Free/reduced-price lunch.
- Delivery method.
- Locale.



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.

What Works Clearinghouse™ U.S. DEPARTMENT OF EDUCATION

WWC Intervention Report
A summary of findings from a systematic review of the evidence

ies INSTITUTE of EDUCATION SCIENCES

Adolescent Literacy Updated November 2016

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 alphabets. (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 alphabets for adolescent readers.

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 practice guides

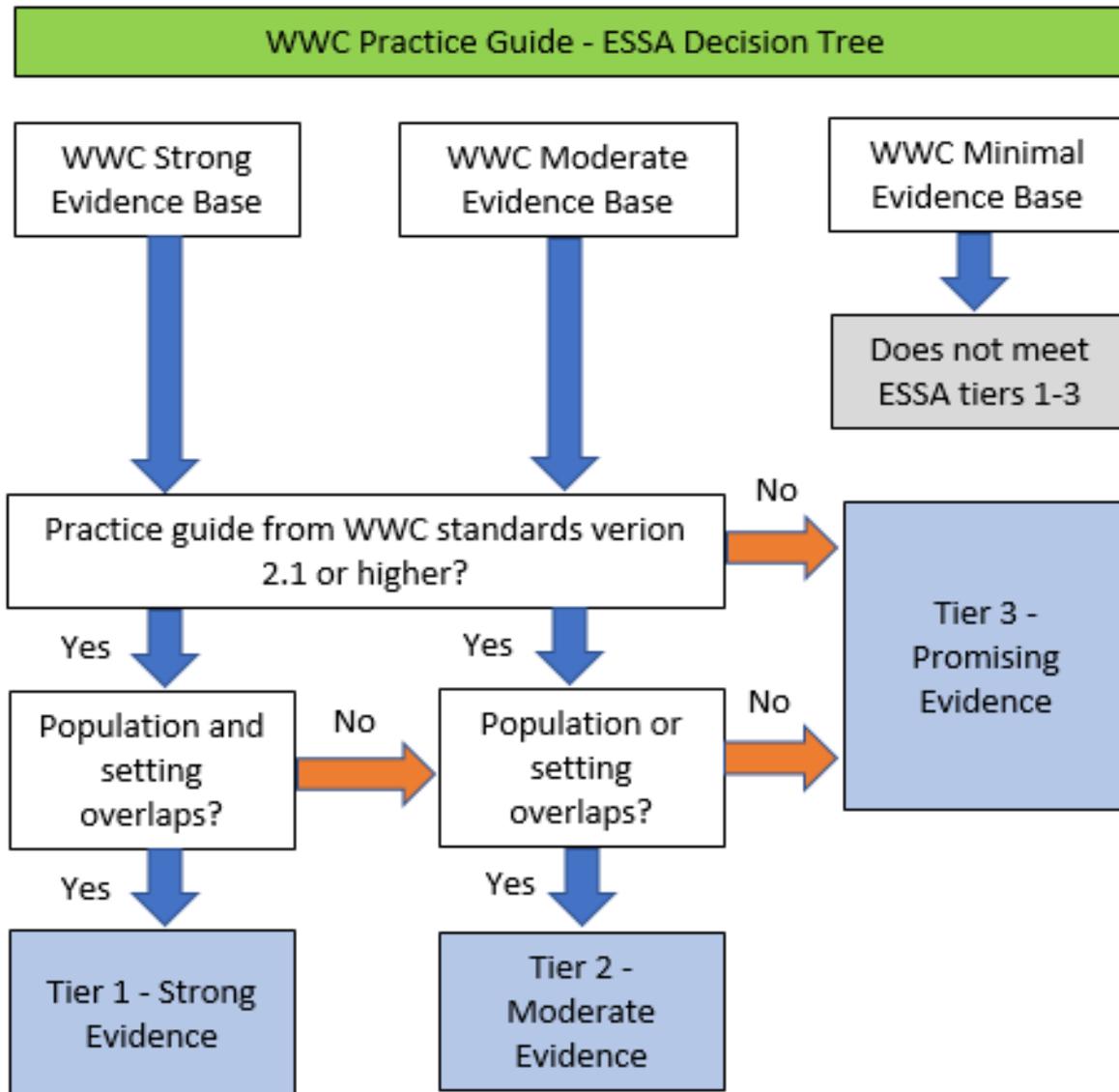
WWC practice guides

Three standards of evidence in the practice guides:

1. Strong evidence base
2. Moderate evidence base
3. Minimal evidence base

Overlap of terms can be confusing!

WWC practice guides



Activity #1

What is the ESSA evidence tier?

Activity: What is the ESSA tier?

An **experimental study** that tested the effectiveness of a new math program on state standardized test scores in mathematics **meets WWC standards without reservations**. The researchers found that the math program **significantly increased mathematics test scores**, and a search of the intervention on the WWC shows **other studies of this intervention have also found significant positive increases**. There were **562 students** from **10 high schools** included in the analysis.

What evidence tier would you assign this outcome?

Tier Rating: Strong (Tier 1)

Activity: What is the ESSA tier?

A **quasiexperimental study** that tested the effectiveness of a science curriculum on science achievement **meets WWC standards with reservations**. The researchers found that the science curriculum **significantly improved science achievement**, and a search of the intervention on the WWC shows **no other studies of this curriculum have been reviewed**. There were **200 3rd graders across 3 elementary schools** included in the analysis.

What evidence tier would you assign this outcome?

Tier Rating: Promising (Tier 3)

Activity: What is the ESSA tier?

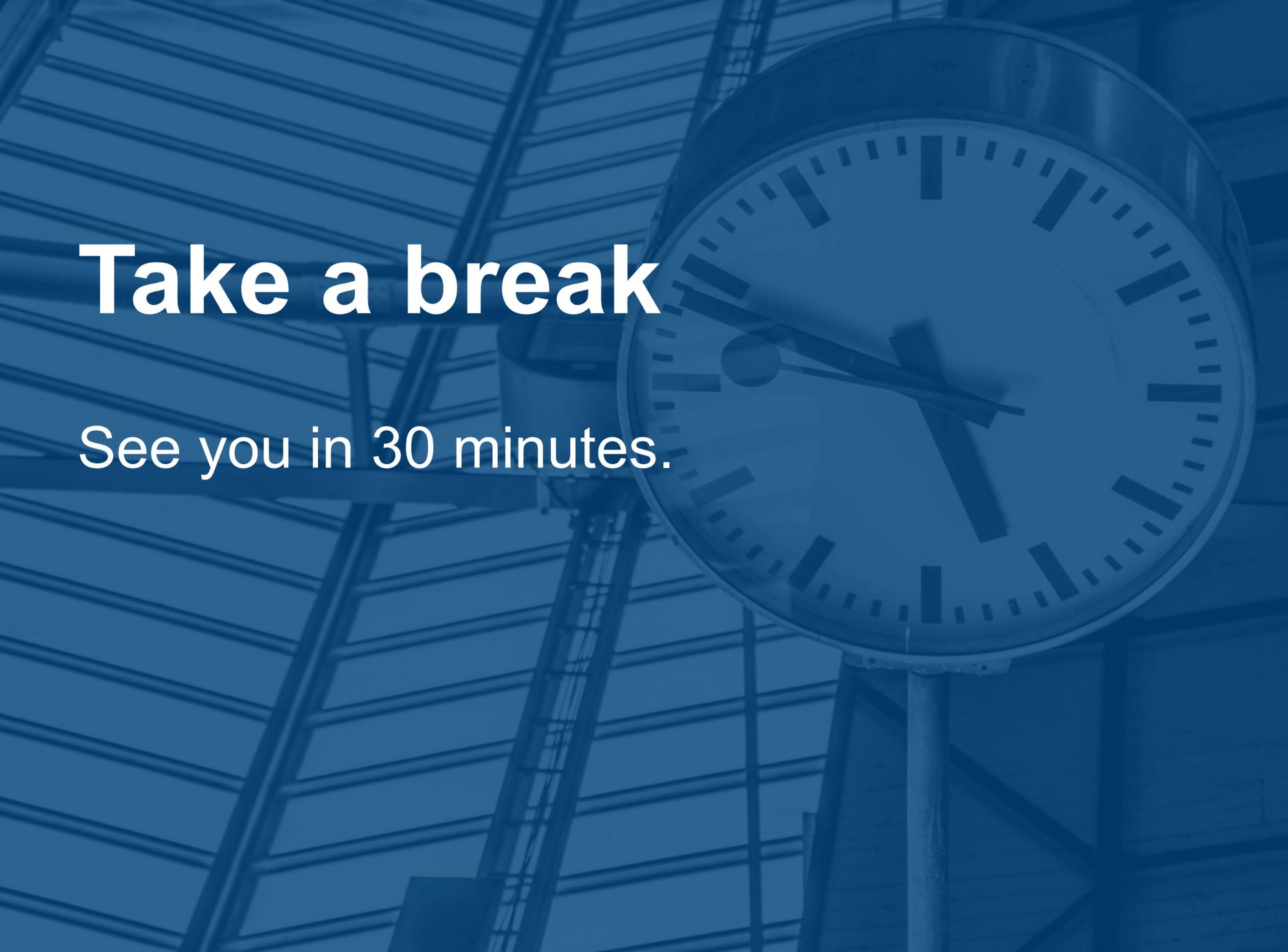
A **quasiexperimental** study looks at the effect of a principal professional development program on student achievement in 40 schools with an average of 300 student per school. The study took into account school size and locale, and student race and socioeconomic status. The treatment and comparison groups **were statistically similar** on math scores before the intervention started, but not on reading. The study found **significant positive effects** for both outcomes. No other studies in the WWC have looked at this intervention.

What evidence tier would you assign the math outcome?

Tier rating: Moderate (Tier 2)

What evidence tier would you assign the reading outcome?

Tier rating: Promising (Tier 3)



Take a break

See you in 30 minutes.

Activity #2

Using the WWC to select evidence-based practices

Small group activity (1 hour)

- Split up in small groups.
- 30 minutes: Use the activity worksheet to complete the scenarios. (Make sure you have a laptop!)
- 30 minutes: Report out findings.



#4

Practice guide:

Foundational skills to support reading for understanding in Kindergarten through 3rd grade

Practice:

Ensure that each student reads connected text every day to support reading accuracy, fluency, and comprehension.

Context:

Universal free lunch, large urban district

Thank you!!



Matt Linick

*Senior
Researcher*
m linick@air.org



Lyzz Davis

*Senior
Researcher*
edavis@air.org