

ESSA tiers of evidence webinar series

Session 1: An overview of the ESSA criteria

06/17/2020

Lyz Davis | Matt Linick

Objectives for today's webinar

1. Discuss the need for and use of evidence under the Every Student Succeeds Act (ESSA).
2. Gain a better understanding of the ESSA evidence tiers.
3. Practice applying knowledge through knowledge checks.

Today's facilitators



Lyzz Davis

Principal Researcher
edavis@air.org



Matt Linick

*Senior
Researcher*
mnick@air.org

Tell us about you!

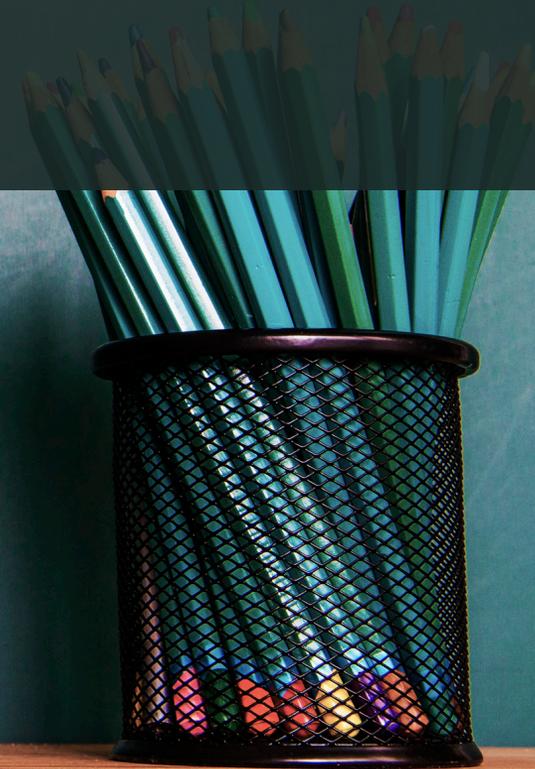
- Name.
- Organizational role.



Agenda

- The importance of evidence
- ESSA and the Michigan context
- ESSA tiers of evidence: Overview

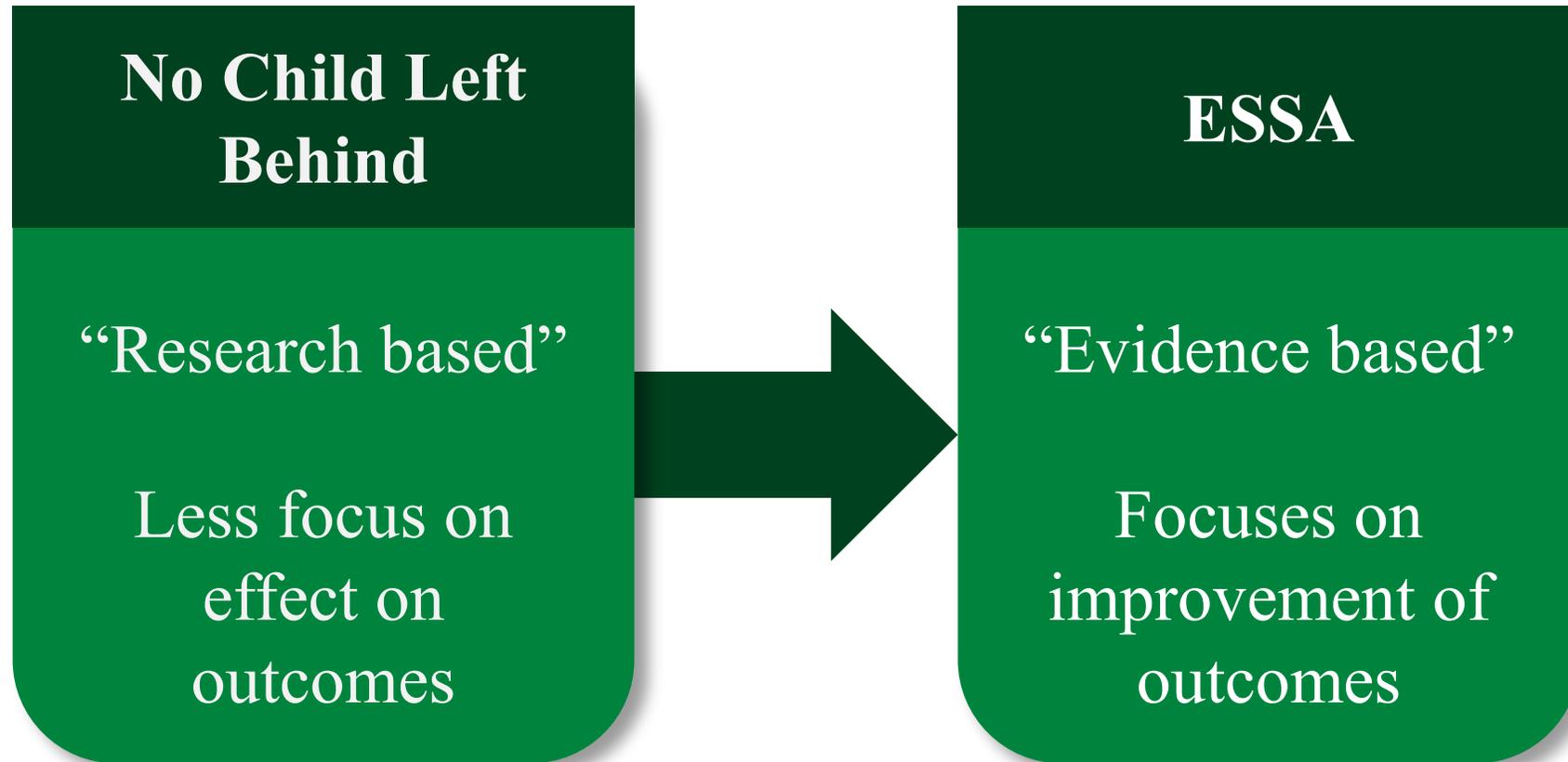
The importance of evidence



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**.

Why is it important to focus on evidence?



In Michigan, ESSA tiers are incorporated into MiStrategyBank

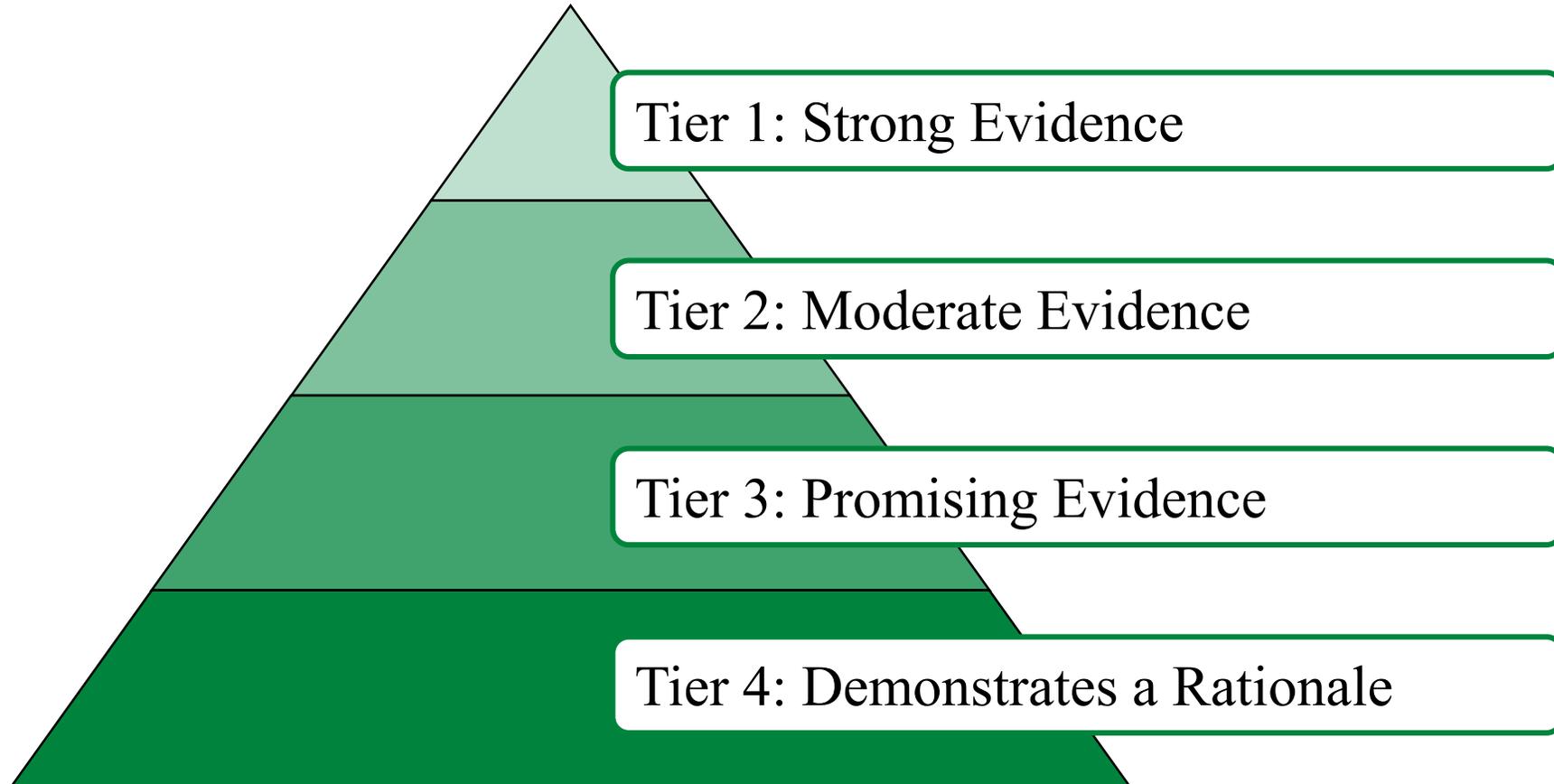
- ESSA tiers a key attribute in MiStrategyBank.
- MiStrategyBank informs, promotes, and encourages the use of increasingly rigorously researched strategies.
- ESSA tags help streamline processes for districts.



ESSA tiers of evidence



Four tiers of evidence under ESSA



ESSA tier 4: Demonstrates a rationale



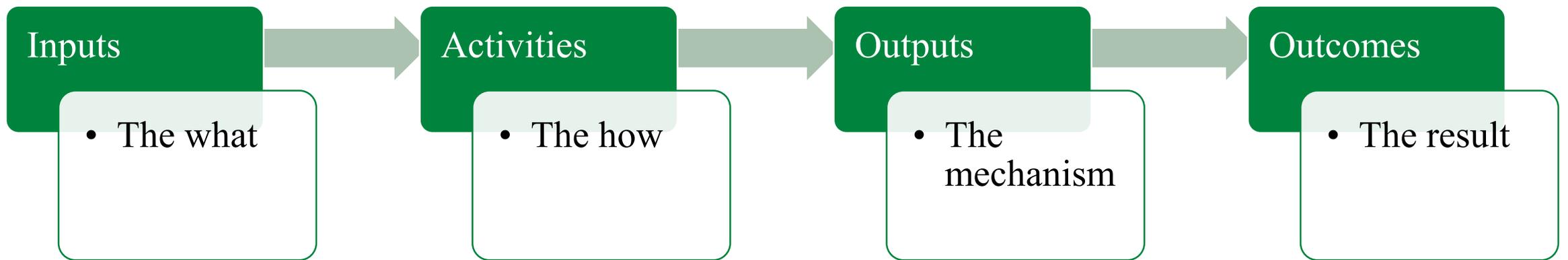
Tier 4: Demonstrates a rationale

- 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**.

What is a well-specified logic model?

A map of *how* a strategy will work

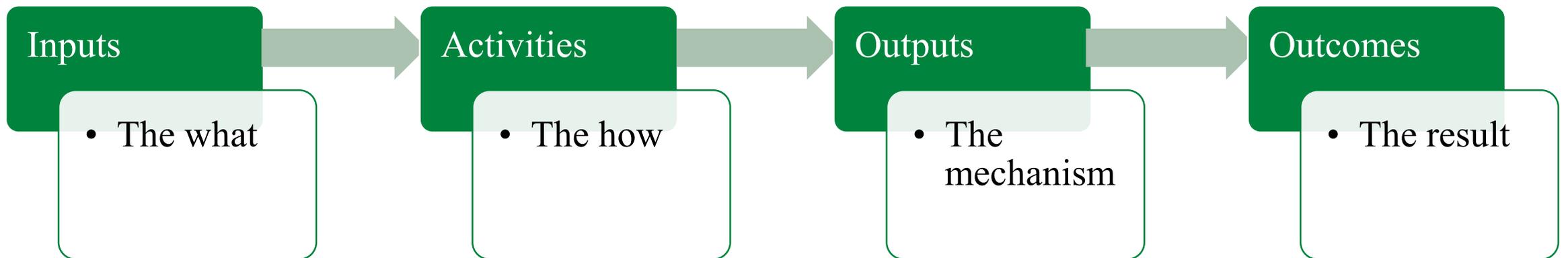
Example:



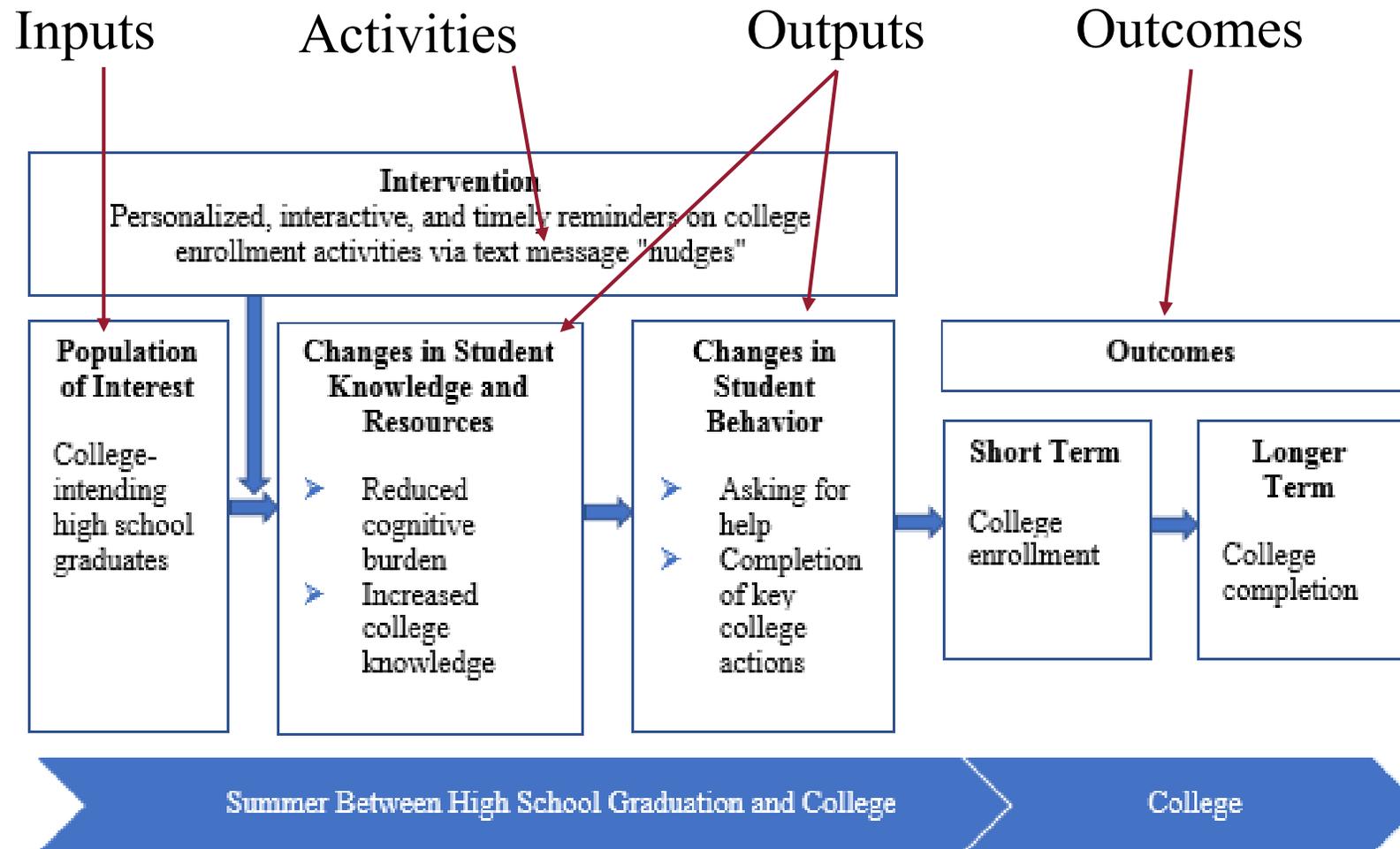
What is a well-specified logic model?

“Well-specified” means all elements are

- Specified.
- Measurable.
- Grounded in rigorous research.



An example



What about the evaluation component of tier 4?

Research needs to be planned, under way, or completed.

Essentially, this means testing your logic model.

What is NOT considered tier 4?

These types of publications might give insight into potential strategies, but rarely review or describe methods used to evaluate a strategy.



Promotional materials



Blog posts



Magazine articles



Books



Websites

Knowledge check 1

Which of the following is not a requirement for a strategy to qualify for tier 4?

- A. The strategy is supported by a well-articulated logic model or theory of action.
- B. The strategy has demonstrated statistically significant, positive findings in a peer-reviewed, published study.
- C. There is a study under way to examine the effects of the strategy on the outcomes of interest.

Which of the following is not a requirement for a strategy to qualify for tier 4?

- A. The strategy is supported by a well-articulated logic model or theory of action.
- B. The strategy has demonstrated statistically significant, positive findings in a peer-reviewed, published study.**
- C. There is a study under way to examine the effects of the strategy on the outcomes of interest.

ESSA tier 3: Promising evidence



What's the difference?

ESSA tier	Group formation	Group equivalence
1		
2		
3		

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		
2		
3	Nonrandom, not purposeful	No baseline equivalence, but statistically control for selection bias

Knowledge check 2

What is one possible reason a finding from an evaluation would *not* qualify for tier 3 status?

- A. The finding is from a study that does not meet WWC standards.
- B. The sample size for the finding was $n = 200$.
- C. A different study of the strategy that meets WWC standards found a significant negative effect for the same finding.

What is one possible reason a finding from an evaluation would *not* qualify for tier 3 status?

- A. The finding is from a study that does not meet WWC standards.
- B. The sample size for the finding was $n = 200$.
- C. A different study of the strategy that meets WWC standards found a significant negative effect for the same finding.**

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.

What's the difference?

ESSA tier	Group formation	Group equivalence
1		
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

Knowledge check 3

If a study has been reviewed by the WWC, what rating(s) could be applied to a finding and have the finding still be eligible for tier 2 status?

- A. Meets Standards Without Reservations
- B. Meets Standards With Reservations
- C. Does Not Meet Standards
- D. All of the above
- E. A and B

If a study has been reviewed by the WWC, what rating(s) could be applied to a finding and have the finding still be eligible for tier 2 status?

A. Meets Standards Without Reservations

B. Meets Standards With Reservations

C. Does Not Meet Standards

D. All of the above

E. A and B

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.

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	Nonrandom, not purposeful	No baseline equivalence, but statistically control for selection bias

Knowledge check 4

True or false: All studies that Meet WWC Standards Without Reservations are eligible for tier 1.

A. True

B. False

True or false: All studies that Meet WWC Standards Without Reservations are eligible for tier 1.

A. True

B. False

Wrapping up



Next week

A deep dive into navigating ESSA and the WWC!



Thank you!



Lyzz Davis

*Principal
Researcher*
edavis@air.org



Matt Linick

*Senior
Researcher*
mnick@air.org