Aligning Evidence-based Clearinghouses with the ESSA Tiers of Evidence

Clearinghouse characteristics

Clearinghouse	Focus areas	Types of studies included	Criteria factored into ratings	How conflicting outcomes are handled	Contextual information provided
What Works Clearinghouse (Institute of Education Sciences)	Academic, behavioral, student subgroups, teacher excellence, dropout prevention, postsecondary success.	Experimental (RCTs) and quasi-experimental designs, including Regression Discontinuity Designs (RDD) and Single Case Designs (SCD).	Strength of the study methodology (design, outcome requirements, confounding factors).	They are reported in intervention reports, but do not factor into ratings.	Intervention report and evidence snapshot pages provide sample size, setting, and student demographic breakdowns for each intervention.
Top Tier Evidence (Coalition for Evidence-based Policy)	All areas of social policy: early childhood, K–12, and postsecondary education, employment/training, health, community, international development.	RCTs only.	Design, implementation fidelity, setting, magnitude of effect with sustained benefits, replicability.	Interventions require "no strong countervailing evidence"—evidence of negative effect or absence of effect from a well conducted RCT.	Evidence summaries (~ 4–10 pages) provide sample sizes, settings, benefit/cost information, and demographics.
Blueprints Programs (Blueprints for Healthy Youth Development)	Programs designed to reduce antisocial behavior, and promote healthy youth development and adult maturity. Specific outcomes in the areas of problem behavior, education, emotional well-being, physical health, & positive relationships.	Experimental (RCTs) and quasi-experimental designs.	Design, implementation, dissemination, magnitude of effect with sustained benefits, replicability.	Interventions require "an absence of iatrogenic effects for intervention participants"; in other words, no harmful effects on specified Blueprints outcomes, either as a whole or for any subgroups.	Sample sizes provided in detailed evaluation abstracts, provide searchable dashboard by outcomes, target population, program specifics, risk/protective factors. Some programs have cost/benefit data.

Clearinghouse	Focus areas	Types of studies included	Criteria factored into ratings	How conflicting outcomes are handled	Contextual information provided
National Registry of Evidence-based Programs and Practices (NREPP)	Programs designed to impact mental health, substance abuse, and wellness. NREPP provides a taxonomy of outcomes.	RCTs and QEDs with "inactive" control group (total absence of intervention, meaning the control group doesn't get any amount of the treatment. No "dosing" studies).	Rigor of design, sample size, magnitude of effects, implementation, articulation of program goals and components. If two outcomes are reported in two studies, a proprietary rating system is used to come up with a rating.	Effect size and design are combined to come up with these categories: Favorable, Possibly Favorable, Trivial, Possibly Harmful, and Harmful.	Users can search by population parameters, outcomes, program type. Program overviews give study specifics, ratings for each outcome, implementation support, cost.
Crime Solutions (separate process for studies and meta-analyses)	Programs and practices designed to improve criminal justice, juvenile justice, and crime victim services outcomes. Outcomes must relate to crime, delinquency, victimization prevention, intervention, response. Single studies are reviewed as **Programs**, and meta-analyses are reviewed as **Practices.**	**Programs**: RCTs and QEDs. A program's evidence rating is based on up to three studies representing the most rigorous study designs and methodologies from all available evaluations of the program.	**Programs**: Research design and analysis methods, implementation, sample size, evaluator independence, publication year.	**Programs**: Studies are classified into five classes based on study rigor and the direction and statistical significance of the effect. Studies with significant negative effect and studies with nonsignificant or null effect are not eligible for effective or promising ratings.	**Programs**: Users can search by evidence rating, population parameters, extent of evidence, setting, topic. Sample sizes for individual studies are provided in Evaluation Methodology. Cost also is included. Links to the studies are provided.

Clearinghouse	Focus areas	Types of studies included	Criteria factored into ratings	How conflicting outcomes are handled	Contextual information provided
		Practices: Meta-analyses. Studies in the meta-analysis need to have "an appropriate control, comparison, or counterfactual condition."	**Practices**: Overall quality of the studies, methods used in meta-analysis, procedures, and proportion of RCTs in the meta-analysis.	**Practices**: Outcomes within a meta-analysis are classified into five classed based on the overall quality rating for the meta-analysis and the direction and statistical significance of the average effect size. The average effect for an outcome has to be statistically significant and positive to get effective or promising rating.	**Practices**: Users can search by evidence rating, population parameters, extent of evidence, setting, topic. Overall sample size is provided. Links to the studies are provided.

ESSA tiers of evidence			
Strong evidence	**Moderate evidence**	**Promising evidence**	**Demonstrates a rationale**
- Well designed and well implemented experimental study. - Significant favorable effect on relevant outcome. - No overriding negative effects from causal studies. - Large, multisite sample.[a] - Overlaps with population.	- Well designed and well implemented QED or RCT with high attrition. - Significant favorable effect on relevant outcome. - No overriding negative effects from causal studies. - Large, multisite sample. - Overlaps with population.	- Well designed and well implemented correlational study or well-designed and implemented RCT or QED without a large/multisite sample. - Statistical controls for selection bias. - Significant favorable effect on relevant outcome. - No overriding negative effects from causal studies.	- Well specified logic model. - An effort to study the effects is or soon to be under way.

[a] A large sample is 350 or more students, or 50 or more groups with 10 or more students. A multisite sample is more than one site, and a site is a local education agency, locality, or state. Multiple studies can be combined to meet the large and multisite sample requirement as long as all studies meet the other requirements and examine the effects of an intervention on the same outcome domain.

ESSA alignment with existing evidence-based clearinghouses

Clearinghouse	Study/program ratings	Criteria	Alignment with Every Student Succeeds Act evidence tiers
What Works Clearinghouse	Meets standards without reservations	• Well designed, well implemented experimental study with low attrition. • Well designed, well implemented Regression Discontinuity Design (RDD).	If positive/potentially positive effectiveness rating with large multisite sample = **Strong Evidence.** If positive/potentially positive effectiveness rating without large multisite sample = **Promising Evidence.**
	Meets standards with reservations	Well designed, well implemented quasi-experimental design with baseline equivalence (or an RCT with high attrition that can be reviewed as a quasi-experimental design).	If positive/potentially positive effectiveness rating with large multisite sample = **Moderate Evidence.** If positive/potentially positive effectiveness rating without large multisite sample = **Promising Evidence.**
Top Tier Evidence	Top Tier	• Well-designed, well-implemented RCTs in replicable setting. • Large, sustained effects. • Must be multisite.	If sample size is large = **Strong Evidence.** If sample size is not large = **Promising Evidence.**
	Near Top Tier	• Meet most Top Tier standards; only need one additional step to qualify (such as replication).	**Promising Evidence.**
Blueprints for Healthy Youth Development	Model+ Programs	• At least two high-quality RCTs or one RCT and one QED. • Significant sustained positive impact on intended outcomes. • No evidence of negative effects. • Intervention specificity, outcomes, risk/protective factors, and logic model all specifically described. • Results have been independently replicated.	If large/multisite sample = **Strong Evidence.** If no sample size information is available or sample is not large/multisite = **Promising Evidence.**

Clearinghouse	Study/program ratings	Criteria	Alignment with Every Student Succeeds Act evidence tiers
	Model Programs	• At least two high-quality RCTs or one RCT and one QED. • Significant sustained positive impact on intended outcomes. • No evidence of negative effects. Intervention specificity, outcomes, risk/protective factors, and logic model all specifically described.	If large/multisite sample = **Strong Evidence.** If no sample size information is available or sample is not large/multisite = **Promising Evidence.**
	Promising Programs	• One high-quality RCT or two high-quality QEDs. • Significant positive impact on intended outcomes. • No evidence of negative effects. • Intervention specificity, outcomes, risk/protective factors, and logic model all specifically described.	If large/multisite sample and RCT = **Strong Evidence.** If large/multisite sample and 2 QEDs = **Moderate Evidence.** If no sample size information is available or sample is not large/multisite = **Promising Evidence.**
	Effective Outcomes	• Strong methodological rigor. • Short-term favorable outcome with a substantial effect favoring the treatment group.	If RCT and large/multisite sample = **Strong Evidence.** If only QEDs and large/multisite sample = **Moderate Evidence.** If no large/multisite sample = **Promising Evidence.**
National Registry of Evidence-based Programs and Practices	Promising Outcomes	• Sufficient methodological rigor. • Short-term effect likely to be favorable with the effect likely to be substantial.	Not aligned.
	Effective	• Strong evidence to indicate intended outcomes achieved. • Implemented with fidelity.	If includes RCT and large/multisite sample = **Strong Evidence.** If includes only QEDs and large/multisite sample = **Moderate Evidence.** If no large/multisite sample = **Promising Evidence.**

Clearinghouse	Study/program ratings	Criteria	Alignment with Every Student Succeeds Act evidence tiers
Crime Solutions: Programs	Promising	• Some evidence to indicate intended outcomes were achieved.	If includes RCT and large/multisite sample = **Strong Evidence.** If includes only QEDs and large/multisite sample = **Moderate Evidence.** If no large/multisite sample = **Promising Evidence.**