

The following is a list of program databases with various criteria for determining program effectiveness. The ratings provided for each program are listed by database and a definition of these labels are provided. While programs often appear in more than one database their ratings will vary dependent on the definitions used.

The [PEW Results First Clearinghouse Database](#) is a compilation of evidence based programs for a series of other databases including CEBC, SPTW, WWC, RTIP, TPP Evidence Review, NIJ/MPG, NREPP, and WWH. Programs are searchable by name, category, setting, rating, and clearinghouse database. To address the disparate ranking schemes by the individual clearinghouses, PEW uses a colored dot system. The evidence considered by the various clearinghouses is ranked Highest rated, Second-highest rated, Mixed effects, No effects, and Negative effects. Programs that appear in multiple clearinghouses compiled by PEW will have a dot for each listing. The dots may be different colors for the same programs listed by different clearinghouses.

The [National Institute of Justice Crime Solutions/Model Programs Guide](#) (NIJ MPG) bases its ratings on experimental or quasi-experimental evaluations and whether significant effects were found in different domains. Their ratings are Effective, Promising, and No Effects, and each rating has two levels to indicate how much empirical support it has, either one study or more than one study. Replicability is a priority. Programs that are rated Effective in more than one study tend to be the strongest evidence based practices while programs rated No Effects in more than one study should be avoided.

NIJ/MPG [Scoring](#):

Effective – Programs that have strong evidence indicating they achieve their intended outcomes when implemented with fidelity.

Promising – Programs have some evidence indicating they achieve their intended outcomes. Additional research is recommended.

No effect – Programs have strong evidence indicating that they did not achieve their intended outcomes when implemented with fidelity.

The [Children Exposed to Violence Evidence Based Guide](#) (CEV EBG) report lists programs that serve children exposed to violence as the treatment population. In compiling this guide, the authors searched the NREPP, NIJ MPG, the SAMHSA National Child Traumatic Stress Network (NCTSN), and the OJJDP CEV EBG. They report the program ratings provided by these sources, organized as two blocks: effective and promising. This resource presents programs that contribute to at least one of the three goal areas: increase resilience, reduce trauma/symptoms, and reduce incidence.

[What Works in Reentry Clearinghouse](#) (WWRC) is yet another tool that organizes the NIJ MPG, and uses the same rating system with a specific emphasis on reentry programs. The majority of the programs listed here are either promising or show no effects. It is important to avoid implementing programs that sound good, but are ineffective.

The [Institute of Educational Sciences](#) (IES) What Works Clearinghouse (WWC) lists education related programs and policies that are evidence based. They rate the programs on a continuum from positive effects to negative effects, their criteria is described in a [one page summary](#), linked below. They also have a useful glossary of [research terms](#).

Positive Effects: Strong evidence that an intervention had a positive effect on outcomes.

Negative Effects: Strong evidence that an intervention had a negative effect on outcomes.

Potentially Positive/Negative: Evidence that an intervention had positive/negative effect on outcomes with no overriding contrary evidence.

Mixed Effects: Evidence that an intervention's effect on outcomes is inconsistent. This means the effects may differ by site, implementation strategy, and any number of additional variables.

No Discernable Effects: No evidence that an intervention had an effect on outcomes.

[Top Tier Evidence Initiative/Social Programs that Work](#) (SPTW)

Top Tier- “Programs shown in well-conducted RCTs [Randomized Control Trial], carried out in typical community settings, to produce sizable, sustained effects on important outcomes. Top Tier evidence includes a requirement for replication – specifically, the demonstration of such effects in two or more RCTs conducted in different implementation sites, or, alternatively, in one large multi-site RCT. Such evidence provides confidence that the program would produce important effects if implemented faithfully in settings and populations similar to those in the original studies.”

Near Top Tier- “Programs shown to meet almost all elements of the Top Tier standard, and which only need one additional step to qualify. This category primarily includes programs that meet all elements of the Top Tier standard in a single study site, but need a replication RCT to confirm the initial findings and establish that they generalize to other sites. This is best viewed as tentative evidence that the program would produce important effects if implemented faithfully in settings and populations similar to those in the original study.”

Suggestive Tier- “Programs that have been evaluated in one or more well-conducted RCTs (or studies that closely approximate random assignment) and found to produce sizable positive effects, but whose evidence is limited by only short-term follow-up, effects that fall short of statistical significance, or other factors. Such evidence suggests the program may be an especially strong candidate for further research, but does not yet provide confidence that the program would produce important effects if implemented in new settings.”

[Pennsylvania's Evidence-Based Prevention and Intervention Support Center](#) (EPISCenter) displays rigorously evaluated programs that have an effect on preventing and/or reducing delinquency and promoting youth development. They do not have a rating system.

The [Promising Programs Network](#) (PPN) was last edited in 2014 but still has valuable information about evaluated programs. The criteria are summarized below, for greater detail see the second link.

PPN [Criteria](#):

Proven, meets all of the following criteria: Direct impact on relevant indicator, effect size of +20% or 0.25 SD, minimum of one effect size is statistically significant at the 0.05 level, used a convincing comparison group (i.e. a randomized control trial), had a sample size greater than 30 in treatment and control groups, and the program documentation is publically available.

Promising, meets all of the following criteria: Impact on an intermediary outcome related to relevant indicator, small effect size of +1%, at least one small effect size is marginally statistically significant at the 0.10 level, has a comparison groups that may be somewhat flawed or insufficiently controlled, had a sample size greater than 10 in treatment and control groups, and the program documentation is publically available.

Other reviewed programs, meets any of the following criteria: Impact on irrelevant outcome and no evidence of PPN indicators, no effect size is greater than 1%, no outcome change is marginally statistically significant at the 0.10 level, does not have a convincing comparison group, sample size in one group is less than 10, or the program documentation is restricted.

The [Blueprints programs](#) listing is searchable by desired outcome, target population, program specifics, and risk/protective factors. Multiple programs can be selected and compared side by side. There are three levels of ratings: Model plus, model, and promising.

Model Plus: “Evaluation Quality: A minimum of (a) two high quality randomized control trials or (b) one high quality randomized control trial plus one high quality quasi-experimental evaluation.

Positive intervention impact is sustained for a minimum of 12 months after the program intervention ends.

Independent Replication: In at least one high quality study demonstrating desired outcomes, authorship, data collection, and analysis has been conducted by a researcher who is neither a current or past member of the program developer's research team and who has no financial interest in the program.

Model: Evaluation Quality: A minimum of (a) two high quality randomized control trials or (b) one high quality randomized control trial plus one high quality quasi-experimental evaluation.

Positive intervention impact is sustained for a minimum of 12 months after the program intervention ends.

Promising: Intervention specificity: The program description clearly identifies the outcome the program is designed to change, the specific risk and/or protective factors targeted to produce this change in outcome, the population for which it is intended, and how the components of the intervention work to produce this change.

Evaluation quality: The evaluation trials produce valid and reliable findings. This requires a minimum of (a) one high quality randomized control trial or (b) two high quality quasi-experimental evaluations.

Intervention impact: The preponderance of evidence from the high quality evaluations indicates significant positive change in intended outcomes that can be attributed to the program and there is no evidence of harmful effects.

Dissemination readiness: The program is currently available for dissemination and has the necessary organizational capability, manuals, training, technical assistance and other support required for implementation with fidelity in communities and public service systems. European programs have not undergone the Blueprints certification process to determine dissemination readiness.”

The [California Evidence Based Clearinghouse](#) (CEBC) rates programs along a 5-point scientific scale and a 3-point measurement scale. Programs are organized either alphabetically or by topic (challenge) area.