APPENDIX 3
CHECKLIST FOR BLUEPRINT PROGRAM EVALUATION

The team used the checklist for the Blueprint Program Evaluation to guide the quality assessment of the evidence identified in the program reports reviewed. Studies received a score of 1 point for each criterion they met. The Blueprint checklist used assigns the same weight to all the following elements:

1. Study used an experimental design (randomized controlled trial—RCT).
2. The sample was clearly described. This variable summarizes a range of conditions, including clarity describing the size and the representativeness of the sample as well as other characteristics of the sample.
3. The reliability or validity of tests and measures was described.
4. The study used intention to treat analysis.
5. The analysis was done at the proper level. This refers to a range of steps in the analysis. It may refer to whether data were analyzed using the correct instrument—for example, using the correct type of regression analysis when the dependent variables are nominal, ordinal, or intervalar. It could also refer to whether the interpretations of analysis were based on the correct unit of analysis or level of such analysis. For example, if a sample was taken in a rural area alone, the statements in the analysis should not suggest the study took place in rural and urban areas.
6. The analysis controlled for baseline outcome measures.
7. The analysis demonstrated baseline equivalence between conditions. This refers to the analysis conducted to determine if two groups are similar enough that one can be used as an “intervention” group and another as a “comparison” group in experimental studies. This analysis typically involves comparing the two groups during baseline.
8. The study demonstrated that attrition is below 5% or unrelated to group assignment, sociodemographic characteristics, and baseline measures of the outcomes.

Studies with evaluation results were divided into the following four categories, according to study design and quality score: high-quality experimental studies, high-quality quasi-experimental studies, lower-quality experimental studies and lower-quality nonexperimental studies.

Program Name

1. Does the study have a high-quality design? A randomized trial is ideal, but two or more studies with quasi-experimental designs may be sufficient. Report on the use of randomization or the nature of the quasi-experimental design.
2. Does the study clearly describe the sample size at each stage of data gathering? Report the number of subjects at each stage, including the N at baseline and the Ns and percentages of the baseline sample remaining at post-test and each follow-up.
3. Are the measures reliable and valid? Report the information provided by the study (e.g., interrater reliability, Cronbach’s alpha).
4. Does the study use an intent-to-treat analysis? The study should attempt to follow and analyze all subjects as assigned to their original condition.
5. Is the analysis done at the proper level? Report on whether the analysis matches the level of the intervention. (For example, if schools are randomized, the analysis should compare schools, not persons, or use multi-level statistical methods that adjust for clustering).
6. Does the analysis control for baseline outcome measures? Report on the use of change scores, baseline outcomes as covariates, or group-by-time interactions.
7. Does the analysis demonstrate baseline equivalence between conditions? Report on whether a test was performed, non-equivalent findings, and potential adjustments.

8. Does the study demonstrate that attrition is below 5% or unrelated to group assignment, sociodemographic characteristics, and baseline measures of the outcomes? Report on whether a test was performed, evidence of significant differential attrition, and potential adjustments.