

Guidelines for Blueprints Checklist – September 7, 2016

1. High Quality Design: *Did the study have a high-quality design?*
 - Y: Randomized (individual or clustered)
 - ?: High-quality QED or randomized with flaw (e.g., consent after assignment)
 - N: Not randomized, crude matching, or criteria for assignment unclear

2. Sample Ns Tracked – *Did the study clearly describe the sample size at each stage of data gathering?*
 - Y: Attrition rates clear for all assessments (usually shown in CONSORT chart)
 - ?: Some details missing or inconsistencies
 - N: No information on attrition, only analysis N

3. Measures Independent – *Was measurement of the outcomes done independently from the delivery of the intervention?*
 - Y: Self reports, raters blind, or archives/records
 - ?: Some but not all measures are independent (note that both are used)
 - N: Program deliverer (including parents and teachers) provided all measures

4. Measures Valid/Reliable – *Are the outcome measures reliable and valid?*
 - Y: Reported alphas and validation by others, or used well-known measures
 - ?: Reported some alphas, had low alphas, or changed well-known measures
 - N: New measures, no information

5. Behavioral Outcome Measure – *Did the measures include a Blueprints behavioral outcome (see list)?*
 - Y: At least one BP behavioral outcome
 - ?: Unclear
 - N: No, tied closely to program content or knowledge, attitudes, intentions only

6. Intent-to-Treat – *Did the study perform an intent-to-treat analysis?*
 - Y: Excluded only those refusing to provide data or lost to follow-up, or used multiple imputation or FIML to include all randomized subjects
 - ?: Didn't try to follow program dropouts, inadequate imputation, or unclear
 - N: Dropped subjects because they did not complete program

7. Proper Level – *Was the analysis done at the level of randomization?*
 - Y: Randomized individuals or adjusted for clustering (e.g., multilevel)
 - ?: Assigned groups but multilevel said to be unnecessary (e.g., low ICC)
 - N: Assigned groups but analyzed at individual level

8. Baseline Outcome Controls – *Did the analysis control for baseline outcome measures?*
 - Y: Used always (change scores, covariates, group-by-time interaction)
 - ?: Used sometimes or not possible (as in survival analysis)
 - N: Not used

9. Baseline Equivalence – *Did the analysis demonstrate baseline equivalence between conditions for randomized subjects?*

Y: Tests complete and showed no significant differences or chance differences

?: Tested for only some variables or some significant differences appeared

N: No tests or many differences

10. Differential Attrition Minimal – *Did the study demonstrate that attrition was below 5% or unrelated to socio-demographics and baseline outcomes?*

Y: Attrition < 5% or tests showed no significant relationships with attrition

?: Tests incomplete or showed some significant relationships with attrition

N: No tests

11. Tested Baseline-by-Condition Attrition – *Did the study check to see if baseline equivalence held for the analysis sample, after loss of subjects from attrition?*

Y: Yes, and no differences or attrition < 5%

?: Yes, but some differences

N: No

12. Posttest Effect on Behavioral Outcome – *Did the results find a significant ($p < .05$ two-tailed) program benefit on an independently measured behavioral outcome, beyond what would be expected by chance across multiple tests?*

Y: Significant effect on an independently measured behavioral outcome

?: Weak, inconsistent effect, or significant effect only on a non-independent outcome measure

N: No significant effect on a behavioral outcome

13. Iatrogenic Free – *Were the results free of any significant iatrogenic effects?*

Y: No significant harmful effects

?: Possible but inconsistent harmful effect

N: Found significant harmful effect (list what it is)

14. Long-Term Effect on Behavioral Outcome – *Did the study find a significant program benefit on an independently measured behavioral outcome that lasted at least one year beyond the end of the intervention?*

Y: Significant effect on an independently measured behavioral outcome 1 year after program end

?: Weak, inconsistent effect or significant effect only on a non-independent outcome measure

N: No significant effect on a behavioral outcome or not tested

15. Effects on R&P Factors – *Did the study find any significant effects on risk and protective factors or measures that are not BP behavioral outcomes?*

Y: Consistent significant effect

?: Weak, inconsistent effect

N: No significant effect or not measured

16. Sample General – *Was the sample broad and general rather than narrow?*
Y: Diverse subjects from multiple states, regions, cities, schools
?: Limited sites, other countries, or many exclusion factors
N: One place, small sample, or highly selective, unusual sample
17. Fidelity of Implementation – *Did the study provide quantitative measures of fidelity of implementation?*
Y: Quantitative evidence of high fidelity and acceptable attendance
?: Some weak or incomplete quantitative evidence
N: No quantitative information
18. Effect Sizes – *Were effect sizes presented?*
Y: Reported d, r, beta, eta², OR
?: Reported but very small effects sizes (d < .20)
N: Not reported
19. Mediation – *Did the study examine mediators that help account for the relationship between the intervention and outcome?*
Y: Full analysis with indirect effects and path diagram
?: Analysis but weak evidence of mediation
N: Not examined
20. Recommended for BP Board – *Was the study strong enough to be reviewed by the BP Board?*
Y: Definitely, no major limitations
?: Maybe, some limitations but not major and can request more results
N: Major limitations