



***CrimeSolutions* Intervention Assessment and
Content Development Services**

**Senior Researcher Procedures Manual
for *CrimeSolutions* and *Model Programs Guide***

Updated March 2023

**Submitted to
National Institute of Justice
Office of Juvenile Justice and Delinquency Prevention
Office of Justice Programs
810 Seventh Street NW
Washington, DC 20531**

**Submitted by
Development Services Group, Inc.
7315 Wisconsin Avenue, Suite 800E
Bethesda, MD 20814
301.951.0056
www.dsgonline.com**

**Prepared under Contract 47QRAA20D002V
Order No. 15PNJD20F0000003**

Changelog for Senior Researcher Procedures Manual

1. Overall
 - a. Updated formatting as needed.
2. Sending the Program Studies Forward for Review
 - a. Added information on the Program and Study Goals Tables that were added to the Program Screening Form.
 - b. Added guidance on how to complete the Program Screening Form with new program and study goals denoted.
3. Appendix D
 - a. Replaced previous version of Program Screening Form with new version.

Contents

Prioritizing Programs and Practices	2
Screening Programs	2
Selecting Studies for Review	4
Selecting the Outcomes and Analyses to Score	6
Sending the Program Studies Forward for Review	11
Screening Practices	11
Identifying an Eligible Meta-Analysis	16
Selecting Effect Sizes for the Evidence Summary	19
Sending the Meta-Analyses Forward for Review	24
Reviewing Final Ratings of Programs and Practices	24
Dispute Resolution Process for Programs and Practices	24
Reviewing Profiles of Accepted Programs and Practices	25
Contact Information	28
Appendix A	29
Appendix B	31
Appendix C	32
Appendix D	33
Appendix E	33

THIS MANUAL DESCRIBES THE GUIDELINES that Senior Researchers should follow when prioritizing and screening programs and practices, handling dispute resolutions, and reviewing program and practice profiles for inclusion in *CrimeSolutions*, the *Model Programs Guide* (the *MPG*), and the National Mentoring Resource Center (NMRC). This document also describes how to prepare programs and practices to be sent out for review (e.g., selecting primary and secondary outcomes).

Prioritizing Programs and Practices

Senior Researchers will receive lists of eligible programs and practices to prioritize for review. All eligible programs and practices have been prescreened by DSG Research Assistants to ensure they meet the minimum eligibility criteria. For the programs, the list will include only those that have been categorized as “high priority.”* For example, those studies that are randomized controlled trials or quasi-experimental designs published within the past 5 years are categorized as “high priority,” while time-series analyses published more than 10 years ago are categorized as “low priority.” This categorization does not happen for practices, as there are not as many meta-analyses available for review.

When Senior Researchers receive the list for their topic areas, they should prioritize the programs and practices in the order in which they think they should be reviewed. The prioritization will depend on the number of programs and practices eligible for review. For instance, if there are 10 or fewer programs in a list, Senior Researchers will be asked to place all programs in the order in which they wish to see them reviewed. If there are more than 10 programs on a list, Senior Researchers may be asked to select their top 10 choices for review (in order) and will not need to prioritize any other studies on the list. The Project Director or Deputy Project Director who sends out the list will instruct Senior Researchers on how to prioritize their lists.

The lists include 1) the name of the program or practice/meta-analysis, 2) the reference/references of eligible studies or meta-analyses, and 3) a brief abstract of the study or meta-analysis. If Senior Researchers want more information about a specific program or practice (including a copy of the study or meta-analysis), they can ask the Project Director or Deputy Project Director. The abstract will generally include a brief description of the program or practice, the sample size, and the specific study design. In addition, within each abstract, DSG will boldface the text describing the specific study design, and boldface and underscore the text describing the outcomes available for selection from the study. This information should be considered while Senior Researchers are prioritizing programs for review.

It may help for Senior Researchers to provide notes for the programs and practices that they select for review; however, this is not required. Senior Researchers may also wish to screen out or deprioritize studies based on the information provided on the list (see below for more information on deprioritizing programs).

Screening Programs

After prioritized programs are approved for review by National Institute of Justice (NIJ) staff, Senior Researchers will receive each program’s evidence base to screen from the Project Director or Deputy Project Director. Although the studies have already been initially screened by the Research Assistants, Senior Researchers must confirm that the studies meet the following criteria for inclusion on *CrimeSolutions* and *MPG*:

1. The primary **goal** of the intervention falls within the **scope** of *CrimeSolutions* and *MPG*.

*The three categories of priority are “high,” “medium,” and “low” and are based on two factors: study design and year of publication.

- It aims to prevent or reduce crime, delinquency, and related problem behaviors (such as aggression, gang involvement, and school attachment).
 - It aims to prevent, intervene, or respond to victimization.
 - It aims to improve justice systems or processes.
 - It targets individuals who have committed offenses or an at-risk population (that is, individuals who have the potential to become involved in the justice system).
2. The **study design** meets the criteria required by *CrimeSolutions* and the *MPG*.
 - Experimental (randomized controlled trial, including a control group)
 - Quasi-experimental (including a comparison group)
 - Time-series design (with comparison group or single group)
 3. The **outcomes** fall within the **scope** of *CrimeSolutions* and the *MPG*.
 - They aim to prevent or reduce crime, delinquency, or related problem behaviors (such as aggression, gang involvement, or school attachment), which may be presented as individual behaviors, community-level behaviors, crime rates, and the like.
 - They aim to prevent, intervene, or respond to victimization.
 - They aim to improve justice systems or processes.
 - They aim to reduce risk factors for crime and delinquency, including school failure, psychological problems or mental illness, and so forth.
 4. There must be at least one **behavioral outcome** of interest (that is, studies that include only Tier 3 outcomes measuring attitudes, beliefs, or knowledge are not eligible; additional information on the tiers of the outcomes is provided below).
 5. The study must be published in a peer-reviewed publication or documented in a comprehensive evaluation report.
 6. The study must be published in or after 2000.
 7. There is some evidence that the study authors measured **implementation fidelity**. Two aspects of fidelity are scored.
 - The authors documented information about fidelity to the program.
 - The authors reported the degree to which the implementation adhered to the program.
 - Although a good fidelity score is required for a Class 1, 3, or 4 study rating, it is not required for a Class 2 study rating. Thus, if this information is not provided in study or supplemental materials, it does not automatically require the study to be screened out, but it may affect the Senior Researchers' decision.
 - If there is no information on implementation fidelity provided in the evidence base, Senior Researchers may conduct an additional Internet search to look for additional relevant publications.

If any of the studies do not meet the above minimum eligibility criteria, Senior Researchers should inform the Project Director or Deputy Project Director of their reason for screening out the study. This will then be documented within the internal *CrimeSolutions/MPG* program database, and the program will be included on the list of [Screened-Out Program Evaluations](#).

Deprioritizing studies. Senior Researchers may find that a study meets the minimum eligibility criteria stated above but should still *not* move forward for review (i.e., there may be limitations to the study design such that the Senior Researcher believes the study would be rated Class 5, Inconclusive Evidence). In this instance, the Senior Researcher should instead **deprioritize** the study. Deprioritized studies will remain on the list of programs eligible for review, but these studies will likely not be reviewed because of the finite resources

available every year to complete program reviews.* Deprioritized studies will *not* be included on the list of Screened-Out Program Evaluations. The deprioritization process allows the study to continue to be available for review, should it need reconsideration in the future (e.g., if the topic area becomes of particular interest to NIJ, or there is another compelling reason to review the program). NIJ has made the list of deprioritized studies available to the public on the *CrimeSolutions* website, [Programs Held for Future Consideration](#).

Supplemental documents. In addition to outcome evaluation studies, a program’s evidence base may include supplemental studies, which should be screened. Supplemental materials may include longer reports of published articles that include detailed information on program elements or logic models; reports on process evaluations or implementation fidelity; or any other information that could help the Study Reviewers score a particular evaluation.

Senior Researchers may elect to send forward any of the supplemental materials identified by the Research Assistants. Senior Researchers may also decide to send a study forward as supplemental material if they do not think it should be scored but think it still should be considered for other information it provides (e.g., more detailed program description, implementation fidelity).

Selecting Studies for Review

Once Senior Researchers have decided to send a program forward for review, they will need to select the study or studies they want to have scored and which (if any) they want to include as supplemental studies.

NUMBER OF STUDIES

No *more than three* studies should be scored for a program review. If a program has more than three studies to select for the evidence base, Senior Researchers should use the following criteria to determine the three most rigorous studies for inclusion:

- Strength of research design[†]
- Breadth of documentation
- Type of analytic procedures used
- Year of publication

If a Study Reviewer believes there is a compelling reason to review more than three studies, he or she may contact the Senior Researcher to request additional studies for review. Similarly, a Study Reviewer may disagree with the inclusion of a study in the program’s evidence base. The Senior Researcher will then make the final determination of which studies will be included in a program’s evidence base.

POSSIBLE EXCLUSIONARY CRITERIA

The following are additional criteria that may exclude a study from the review.

Statistical comparisons. For each study that meets the criteria for inclusion, statistical comparisons should be made *between* the treatment and comparison groups (that is, between-group differences), and not for within-group differences (for example, pre–post differences within the treatment group only). If the study authors do not conduct analyses examining the between-group differences, the study should be screened out

*Almost all the programs sent forward for review are categorized as “high” priority. Deprioritizing a study will recategorize the program as “medium” or “low” priority in the *CrimeSolutions.gov/MPG* program database.

[†]Prioritize randomized controlled trials when possible.

(the exception is single-group time series designs; see guidance on page 9 on Time Series Designs for more information).

Comparative effectiveness research. At this time, it has been decided to screen out comparative effectiveness research—that is, those studies for which the only comparisons are between two treatment groups that do not have a true no-treatment or treatment-as-usual comparison group. Please make a specific note of this when screening out the study, as these instances will be tracked, and the topic revisited in the future.

To ensure that comparative effectiveness research is not screened forward for review, Senior Researchers should clearly identify the condition of the comparison or control group, to confirm it is appropriate. If Senior Researchers are uncertain about the appropriateness of the comparison condition/control group, they should reach out to the Project Director for additional guidance.

Hazard ratios/interrupted time series analysis. When a study uses hazard ratios or an interrupted-time-series analysis, the statistics they produce are largely interpretable only within the context of the study parameters and the precise statistical model used. In these instances, the Senior Researcher should instruct the Study Reviewers to score the two-by-two frequency table of the outcomes and to use the effect size calculator to establish an effect size. If the study does not provide this incidence rate, the study should be screened out.

Pooled results of multiple programs within a single study. If *different* programs are evaluated within a single study, then the effect sizes of each program must be reported separately. However, if the effect sizes of these different programs are pooled together into an overall effect size (without reporting the separate effect size per program), then the study should be screened out. For example, a study evaluating six different mentoring programs (implemented in different jurisdictions, by different providers, and through independent organizations) provided an overall effect size that pooled the effects across the six programs. Because the study did not provide an effect size per program, the study was screened out.

However, if the *same* program is implemented across different sites, and the results are pooled across the sites into an overall effect size, the study should *not* be screened out. For example, a study evaluating the same mentoring program (implemented in different jurisdictions or by different providers) provided an overall effect size that pooled the effects across the sites. Because it is the same program implemented across the sites, the study is not screened out.

If there are any questions about whether a study is evaluating the same program or different programs, Senior Researchers can ask for guidance from the Project Director.

Regression discontinuity design (RDD). RDD is a quasi-experimental design that examines the causal effects of interventions by assigning a cutoff above or below which an intervention is assigned, and then comparing the observations that are close to either side of the cutoff. As Lee and Lemieux (2010, p. 281) explain, RDD is “a way of estimating treatment effects in a nonexperimental setting where treatment is determined by whether an observed ‘assignment’ variable (also referred to in the literature as the ‘forcing’ variable or the ‘running’ variable) exceeds a known cutoff point. In their initial application of regression discontinuity designs, Thistlethwaite and Campbell (1960) analyzed the impact of merit awards on future academic outcomes, using the fact that the allocation of these awards was based on an observed test score.

The main idea behind the research design was that individuals with scores just below the cutoff (who did not receive the award) were good comparisons to those just above the cutoff (who did receive the award).”*

Currently, the Program Scoring Instrument does not allow for appropriate scoring of this type of study design. Therefore, any studies/evaluations that use RDD should be screened out.

Selecting the Outcomes and Analyses to Score

OUTCOMES

Senior Researchers will select outcomes using the Tiered Outcomes List for guidance. The Tiered Outcomes List was developed to provide important guidance for Senior Researchers during the outcome selection process. An updated version of the Tiered Outcomes List (which is in Excel spreadsheet format) will be sent to Senior Researchers with every new program they need to screen.

The Tiered Outcomes List divides potential outcomes for review into three tiers:

- **Tier 1 Outcomes.** These are outcomes that Senior Researchers should categorize as “primary,” unless they can provide justification for categorizing an outcome as “secondary.” For all topic areas, crime/delinquency outcomes are categorized as Tier 1. Additional Tier 1 outcomes depend on the topic area.
- **Tier 2 Outcomes.** These are outcomes that Senior Researchers can categorize as “primary” or “secondary,” depending on the theory of change and the context, intent, and goal of the program. Tier 2 outcomes should never replace any Tier 1 outcomes as primary unless the Senior Researcher can justify the decision.
- **Tier 3 Outcomes.** These are outcomes that Senior Researchers should categorize as “secondary,” unless they can provide justification for categorizing an outcome as “primary.” For all topic areas, attitudes/beliefs/knowledge outcomes (i.e., nonbehavioral outcomes) are categorized as Tier 3. Additional Tier 3 outcomes depend on the topic area. Please note that current *CrimeSolutions* policy states that studies that include only secondary outcomes should be screened out. Thus, if a study has outcomes categorized only as Tier 3, the study screens out.

Guidance to using the Tiered Outcomes List. The Tiered Outcomes List has different prioritizations for each topic area. However, there may be situations in which a program could have been reviewed in more than one topic area (for example, a drug court could be scored within the courts topic area or the drugs topic area). If a program falls into multiple topic areas, the Senior Researcher should rely on the highest possible tier for selecting outcomes. For example, if the drug court program were sent to the Senior Researcher of the court’s topic area, the drugs/substance abuse outcomes would be Tier 2. If the drug court program were sent to the Senior Researcher of the drugs topic area, the drugs/substance abuse outcomes would be Tier 1. Therefore, the drugs/substance abuse outcomes should be considered Tier 1 regardless of under which topic area the program is reviewed. Senior Researchers will be alerted when they are sent a program that could have been reviewed in another topic area.

*Lee, David S., and Thomas Lemieux. 2010. “Regression Discontinuity Designs in Economics.” *Journal of Economic Literature* 48:281–355. The *What Works Clearinghouse* provides additional information on [RDD](#).

The Tiered Outcomes List is meant to be comprehensive, meaning if Senior Researchers come across an outcome that is *not* on the list or an outcome marked as “N/A” (meaning not applicable) on the list in their topic area, they should reach out to the DSG Project Director and the NIJ staff to ensure the outcome should be scored. New outcomes that are approved for review will be added to the list and categorized as Tier 1, 2, or 3. Outcomes on the list marked as “N/A” will also be categorized as Tier 1, 2, or 3. The Tiered Outcomes List is dynamic and will be updated and enhanced as needed.

More than five primary outcomes to select. If a study has more than five Tier 1 outcomes to review, Senior Researchers should consider the intent/goal of the program to select the five primary outcomes to score. Any additional Tier 1 outcomes should be scored as secondary outcomes. For example, if a Senior Researcher were reviewing a corrections program that had more than five measures of recidivism, they could focus on the overall measures of recidivism as primary outcomes, and specific measures of recidivism (e.g., misdemeanor recidivism, felony recidivism) could be secondary outcomes. Senior Researchers should provide justification for why Tier 1 outcomes are scored as primary and/or secondary, when more than five are available to review.

Please note: Senior Researchers should not feel obligated to choose the maximum number of outcomes (10), but only those that are appropriate for inclusion in *CrimeSolutions* and *MPG* and according to the Senior Researcher’s knowledge of the subject area and the goals/purpose and theory of change of the program.

Considering theory of change when selecting outcomes. In addition to using the Tiered Outcomes List to help select and categorize outcomes, Senior Researchers should consider the intent and theory of the change of the program when screening a study. It is not sufficient to select an outcome because the study authors measured it. Rather, there should be justification in the study to explain why outcomes were measured and examined in the evaluation of the program. There should be a reasonable assumption for why an outcome was measured or an explanation of why a program would be expected to make an impact on the measured outcome.

Selecting outcomes related to intermediary actors. Intermediary actors are those who may be targeted by a program or intervention (i.e., the service delivery target population), with the intent of affecting the primary population of interest for *CrimeSolutions*. The program should show an effect or impact on the intermediary actors before the program can achieve an effect or impact on the primary population of interest for *CrimeSolutions*. For example, in parent training programs, the intermediary actor is the parent participating in the training program. The primary population of interest is the child of the parent participating in the training program. Outcomes related to intermediary actors are considered Tier 3 and should always be selected as secondary.

Behavior Item. The purpose of the “Behavior” item on the Program Scoring Instrument is to assess the degree to which the evaluation demonstrates that receiving the program is related to changes in behavior. *CrimeSolutions* considers a program to be more robust if evaluation results indicate changes in behavior, compared with how program participants think about the behavior (such as attitudes toward aggression, knowledge of characteristics of a healthy relationship, and beliefs that one would act in a certain way in a certain situation). For *CrimeSolutions*, *behavior* is used to define the ways (i.e., whether at the individual level or other units of analysis such as classroom level or neighborhood level) that people interact with the world around them. This could include acts such as committing crimes, being arrested, using drugs, engaging in violence/aggression, seeking help or treatment, standing up for victims of bullying, and making calls for

service. For purposes of scoring this item, mental health diagnoses and symptoms (such as depression, trauma symptoms, and anxiety) are also considered behavior, in part because symptoms typically include behaviors (for example, social withdrawal for depression, avoidance of certain stimuli or situations for anxiety), and because they affect the way people think about themselves, relate to others, and interact with the world around them. In addition, if it is unclear whether an outcome is a distinct measure of behavior or attitudinal/intention/beliefs, it is recommended that the outcome be considered a measure of behavior.

Multiple comparison groups. If the study includes more than one control/comparison group or version of a program, the study should probably be reviewed separately by control/comparison group (for example, in a law enforcement study that looks at multiple policing strategies, with each treatment group compared with a treatment-as-usual comparison group, each policing strategy should be reviewed separately). However, such issues should be discussed with the Project Director to determine the best approach.

Excluding outcomes. Outcomes, in general, should *not* be excluded based on methods considerations (for example, for marginal reliability), as the Study Reviewers will be assessing the instrumentation (that is, reliability and validity of outcome measures) in the design quality section of the instrument. However, if the measure is judged to be fundamentally flawed (e.g., it has little or no basis or support in the literature as a viable measure of the outcome), it may be (but does not have to be) excluded.

Certain types of programs are excluded for review from *CrimeSolutions* and *MPG*, as follows:

From *MPG*: Pregnancy-prevention programs, where only pregnancy outcomes are measured, are excluded. However, if a pregnancy-prevention program examines other related outcomes such as risk factors or risky sexual behavior, the program may be reviewed.

From *CrimeSolutions*: Police- or correctional-officer wellness programs are excluded. Training programs for police officers, correctional officers, or other justice staff are also excluded, unless the study examines the impact of training on officers' behaviors. For more information on the inclusion criteria of training programs on *CrimeSolutions*, please see **Appendix A**. Suicide prevention programs also are excluded, unless they target a justice-involved population (for example, youths in detention or adults in jail).

SUBGROUPS

Only outcomes related to the full sample of study participants should be selected for scoring; that is, do **not** select outcomes related to subgroups.

- The study must report analyses that compare the full study sample to the comparison group (for example, the study may provide separate results for high-dosage participants versus the control group, and low-dosage participants versus the control group, but the study can screen in only if the results for the combined high- and low-dosage participants versus the control group are reported). If the study does not report analyses that compare the full study sample with the comparison group, and there are no other eligible outcomes, the study must **screen out**.
- This is true even when it appears that the most appropriate group to examine is a subgroup. For example, a program may have provided an anti-teen dating violence (TDV) curriculum to a whole school, but TDV was measured only among those in a dating relationship during the reporting period (a subgroup of the sample). In this case, the TDV outcomes could **not** be scored, and the study could be eligible for review only if there were relevant outcomes reported for the full sample.

INTENT-TO-TREAT ANALYSES

To the extent they are available, intent-to-treat (ITT) analyses are preferable to those that analyze data only on those for whom there are complete data. However, treatment-on-treated (TOT) analysis is acceptable when ITT analysis is not available.

BIVARIATE ANALYSIS VERSUS MULTIVARIATE ANALYSIS

To the extent they are available, results examined using multivariate analyses are preferable to results that were examined using only bivariate analyses.

TIME SERIES DESIGNS

Under the **Research Design** item, time-series design studies can receive a score of '2' if the study is a comparison group design or a score of '1' if the study is a single-group design. A study can receive credit as a time series *comparison group* design (i.e., receive a score of '2') only if the study authors conducted significance tests to directly compare the results between the treatment group and the comparison group. If the study authors do not conduct significance tests to directly compare the results between the groups, then the study design should be scored as a time series *single group* design (i.e., receive a score of '1'). For example, if the study authors conducted tests of significance only within the treatment group and within the comparison group, and then compared those results descriptively, then this design should be scored as a time-series single-group design.

DIFFERENCE-IN-DIFFERENCE ANALYSIS

Difference-in-difference (DID) analysis is a quasi-experimental design that makes use of longitudinal data from treatment and control groups to obtain an appropriate counterfactual to estimate a causal effect. It is typically used to estimate the effect of a specific intervention or treatment (such as a passage of law or enactment of policy) by comparing the changes in outcomes over time between a population that is enrolled in a program (the treatment group) and a population that is not (the control group). The structure of the experiment implies that the treatment group and control group have similar characteristics and are trending in the same way over time. This means that the counterfactual (unobserved scenario) is that, had the treated group not received treatment, its mean value would be the same distance from the control group in the second period.

Specifically, DID estimation uses four data points to deduce the impact of a policy change or some other shock (i.e., treatment) on the treated population: the effect of the treatment on the treated. The four data points are the observed mean (average) of each group and calculated as *the difference in average outcome in the treatment group before and after treatment minus the difference in average outcome in the control group before and after treatment*.

$$\begin{array}{rcl} \text{Treatment:} & 40 \text{ (AVG before)} - 0 \text{ (AVG after)} = & 40 \text{ (Difference)} \\ \text{Control:} & 20 \text{ (AVG before)} - 20 \text{ (AVG after)} = & 0 \text{ (Difference)} \\ \hline \text{DID} & & 40 \end{array}$$

In terms of *CrimeSolutions*, an assessment of the DID approach would fall under the **Statistical Adjustment** item. It *assesses the use of statistical controls to account for the initial measured differences between the groups. Any outcome-relevant variable on which the groups may differ should be identified and included in the statistical adjustment*. It is suggested that a study using a DID approach is eligible for CS and would likely receive a '2' or '1' on the **Statistical Adjustment** item, depending on other characteristics of the study.

FOLLOW-UP PERIOD

As a general rule, Senior Researchers should choose the most distal outcomes. However, there may be instances in which Senior Researchers may want to make an exception and select proximal outcomes. In these instances, Senior Researchers should use their discretion and expert knowledge of their topic area to decide whether it is more appropriate to select a proximal outcome, given the context and goals of the program. Senior Researchers will need to provide written justification for not selecting the most distal outcomes to the Project Director or Deputy Project Director. This justification will be saved, along with other documents from the program review, should questions about the outcome selection process arise in the future.

Additional information on the follow-up period. In some instances, the final time point of outcome assessment for an evaluation may occur before the scheduled end of the program. For example, a delinquency prevention program may be designed to provide support to students from first through sixth grade, but an evaluation may report on outcomes of the program through only fourth grade. These types of evaluations should be placed on hold and *not* sent forward for review, until an evaluation through the end of the program period is available.

In other instances, a program may include booster sessions or other similar components to help sustain effects after the scheduled end of the program. The timeframe of an evaluation does not need to encompass these types of additional activities to be eligible for review.

Additionally, some programs will have more open-ended timeframes. For example, a mentoring program may require a 1-year minimum commitment from mentors but also include ongoing support for mentoring relationships that extend beyond this timeframe. Senior Researchers are expected to use their subject-matter knowledge to determine what constitutes a minimally acceptable timeframe for evaluations of such programs to be reviewed.

Different outcomes with different follow-up periods across multiple articles or within the same report.

There may be occasional instances in which program reviews involve scoring multiple articles that examined the same study sample across different follow-up periods. For example, one article may look at outcomes at a 6-month follow-up, while another article looks at outcomes at a 3-year follow-up. In such instances, consistent with the guidelines above, Senior Researchers should select the article/publication with the more distal follow-up period and may sometimes include earlier studies as supplemental material (as consistent with the policy to select the most distal follow-up period). There may also be cases for which Senior Researchers should look at two articles of the same study sample, such as when outcomes examined in an article at an earlier follow-up period were not included in an article at a later follow-up period. For example, the article discussing the 6-month follow-up may include self-reported drug use and official records of recidivism, whereas the article discussing the 3-year follow-up may include only the results of the official records of recidivism, because obtaining the self-reported drug use results from the study sample was no longer possible.

In such instances, Senior Researchers should maintain the practice of scoring only the most distal outcomes for the same study sample, including cases in which multiple articles of the same study looked at the same outcome across different follow-up periods. In the example above, the self-reported drug use results would be scored in the 6-month study, and the official records of recidivism would be scored in the 3-year study (note that the recidivism outcome would *not* be scored in the 6-month study). However, if the Senior Researcher has not identified any primary outcomes in a study and the only eligible outcomes are secondary, that study should not be scored, even if it provides the most distal outcome data for the study sample.

Thus, if different outcomes (i.e., the self-reported drug use and official recidivism data) are reported within the same report, the longest follow-up period per outcome should still be scored (i.e., drug use at 6 months and recidivism at 3 years). For items in the Design Quality section that could be affected by the disparate time periods, such as the Follow-Up Period Item, Study Reviewers should give the benefit of the doubt to the study and score the longest follow-up period, even if not all of the outcomes are scored at that distal timepoint (as in the example above, the Study Reviewers could give the study credit for the 3-year follow-up period). This guidance could also be used when scoring other items, such as Sample Size, if the sample differs per outcome. The overall goal is to assess a program's effects on the study sample across the body of evidence of that program, whether that comprises one final report or multiple publications. However, the program will receive an icon for only a single study in the program's evidence base on *CrimeSolutions*.

Sending the Program Studies Forward for Review

Senior Researchers, when sending a program forward for review, should complete the *CrimeSolutions* Program Screening Form (see **Appendix D** for a copy).

Sections on the form include

- Name of the program, completed by DSG.
- Name of Senior Researcher, completed by DSG.
- Date of screening, completed by the Senior Researcher.
- Program Goals Table, completed by DSG and supplemented by the Senior Researcher as necessary.
- Study Goals Table, completed by DSG and supplemented by the Senior Researcher as necessary.
- Primary and Secondary Outcomes Tables, completed by the Senior Researcher.

The objective of the Program Goals Table is to ensure that Senior Researchers consider the goals of the program when selecting outcomes. The objective of the Study Goals Table is to capture the research goals of the evaluators.

When selecting outcomes, Senior Researchers should give greater consideration to outcomes that relate directly to program goals than to outcomes that relate directly only to study goals.

In some instances, program goals are not clearly stated (or even implied) in the available studies or supplemental materials. In such instances, Senior Researchers can relate outcomes to only the study goals. (Studies and outcomes are not automatically screened out if no program goals are identified.) If the study goals and related outcomes do not seem to be connected to what the Senior Researcher could reasonably assume is the intent of the program, they should relay such concerns to DSG.

To complete the Program and Study Goals Tables, DSG will use the study or studies sent forward for screening, along with any supplemental materials.

- **Program Goal.** DSG will include text from the studies or supplemental materials that describes the specific goal or goals of the program.
- **Source.** DSG will cite the studies or supplemental materials in which the program goal was identified.
- **Page or Table Number.** DSG will identify what page the text or language came from.

This approach should make it easier for Senior Researchers to find the goals in the studies or supplemental materials as they screen them. The same procedures will be used to complete the Study Goals Table. However, study goals will come *only* from the studies eligible for review, and not from the supplemental materials.

If a Senior Researcher believes DSG has not identified all of the program or study goals, they may add any additional goals to the tables and inform the Project Director of that change when they send back the screening form. There is no limit on the number of goals that can be added to the goals tables.

When the program under evaluation is an experiment, the program goals are assumed to be the same as the study goals. In this instance, in the Program Goals Table, it will read: “This program is an experiment; therefore, program goals and study goals are considered the same (see Study Goals Table below).”

In completing the primary and secondary outcome tables, the Senior Researcher can choose up to five primary outcomes and five secondary outcomes. Senior Researchers must complete five components for each outcome:

1. **Outcome name.** The outcome name listed in the Program Screening Form should be the exact name of the outcome used in the study, including (if relevant) the follow-up period. If the follow-up period is not specified by the Senior Researcher, DSG will assume that the longest follow-up given in the study is intended.
2. **Page and/or table number** in the study where the outcome can be found.
3. **Categorization from the Tiered Outcomes List** of the outcome, as specific as possible. This is the code for the outcome that can be found in the Tiered Outcomes List.
 - For example, robbery is listed under Crime and Delinquency | Violent Offenses | Robbery and therefore should be coded as 1Fviii on the screening form.
 - Violent offenses not otherwise specified is listed under Crime and Delinquency | Violent Offenses and should be coded as 1F.
 - If the outcome is not already included on the Tiered Outcomes List, the Senior Researcher should write “New” in this box. The newest version of the Tiered Outcomes List (which will be included along with the studies and supplemental documents to screen) should always be used.
4. **Tier Number** as listed in the Tiered Outcomes List (e.g., 1, 2, or 3).
5. **Goal Number.** For each outcome, Senior Researchers should use the Program and Study Goals Tables at the top of the form to identify which program and/or study goal relates to each of the selected outcomes. The Senior Researcher should denote the program and/or study goal for each outcome. For example, if program goal no. 2 was “to reduce recidivism” and the outcome selected is rearrest (a measure of recidivism), the Senior Researcher would write “2” in the Goal column. There can be multiple outcomes coded for each program and study goal. In the previous example, if the study also

included outcomes of reincarceration and reconvictions (also measures of recidivism), the Senior Researcher would again write “2” in the Goal column for each outcome.

Below are additional notes on relating outcomes to goals:

- Senior Researchers should relate each outcome to as many program and study goals as are appropriate.
- Every program goal may not have an accompanying outcome. There are many reasons why an evaluation might not explore a program’s impact on all of the stated program goals.
- If a Senior Researcher selects an outcome that does not relate to an identified program or study goal, the Senior Researcher should provide justification for this decision under the “Additional Notes” section. This information will be forwarded to NIJ for consideration, to ensure that it is suitable to score an outcome that does not relate directly to a program or study goal.
- If the Senior Researcher relates a Tier 2 outcome to a program goal, the Senior Researcher should lean toward making the outcome primary unless they can provide justification for making it secondary. If the Senior Researcher relates a Tier 3 outcome to a program goal, the Senior Researcher should lean toward making the outcome secondary unless they can provide justification for making it primary. Any departure from this guidance will be brought to NIJ for consideration.
- If the Senior Researcher relates a selected Tier 1 outcome to a study goal, the Senior Researcher should lean toward making the outcome secondary unless they can provide justification for making the outcome primary.
- There may be situations in which selected outcomes cannot be related to either program goals or study goals, but they do relate to program components. For example, a program with a stated goal of reducing delinquent behavior in juveniles may incorporate several lessons in the curriculum (such as improving family climate or emotional regulation) that are not directly tied to delinquency but target risk factors associated with it. In these instances, the Senior Researcher should write “Program Components” in the Goal column. In the Additional Notes section, they should specify which specific program component relates to the selected outcome. The Senior Researcher can rely on the Tiered Outcomes List to determine whether outcomes that relate to program components are primary or secondary.
- In instances where program goals are stated in the studies or supplemental materials, but the study authors appear to intentionally disregard the program goal in favor of unrelated study goals, the Senior Researcher should consider whether it is appropriate to score these studies. The Senior Researcher should reach out to the Project Director for discussion when considering screening out a study for this reason.

Senior Researchers need to complete the section “**Eligible Outcomes Not Selected for Scoring**” only if there are outcomes eligible for review that they do *not* select. (Ineligible outcomes, such as findings from subgroup analyses or results from shorter follow-up periods, do not need to be discussed because those outcomes would

never be eligible for review.) For example, there may be a study that includes more than 10 eligible outcomes for a Senior Researcher to select for review. In this instance, for any eligible outcomes not selected, the Senior Researcher would need to provide justification in this section.

Each study has a space available for “**Additional Notes**,” where the Senior Researcher can provide additional information, such as justification for outcomes that do not match any program goal. This space is also where Senior Researchers can explain why a new outcome not on the Tiered Outcomes List was selected, and the reason they want it included in the review. Senior Researchers must include justification when including outcomes not already on the Tiered Outcomes List. This information will be sent to NIJ for consideration.

If there are more than three studies to select from, Senior Researchers can list additional studies not sent forward for review in the “Additional Studies” box and briefly explain why the study or studies are not being sent forward. Citations of supplemental documents for the Study Reviewers can be listed in the “**Supplemental Documents**” box. At the bottom of the Screening Form, Senior Researchers should select two Study Reviewers and one backup Study Reviewer to review the program.

Screening Practices

On the *CrimeSolutions* website, a practice is distinguished from a program. Whereas the evidence base for a program is derived from one to three individual program evaluations, the evidence base for a practice is derived from one or more meta-analyses. **A practice, as defined by *CrimeSolutions*, is a general category of programs, strategies, or procedures that share similar characteristics with regard to the issues they address and how they address them.** *CrimeSolutions* uses the term “practice” to categorize causal evidence that comes from meta-analyses of multiple program evaluations. Using meta-analysis, it is possible to group program evaluation findings in different ways to provide information about effectiveness at different levels of analysis.

The existence of a meta-analysis does *not* mean it is an appropriate practice to include on *CrimeSolutions* (i.e., if a meta-analysis has been done on a particular topic, it does not mean it is automatically considered a practice by *CrimeSolutions*). For example, a specific program, such as Functional Family Therapy (FFT), may have been evaluated enough times for a meta-analysis to have been conducted on the available research, but within *CrimeSolutions* FFT would be considered a specific program and not a more general set of practices. In contrast, a meta-analysis of studies looking at multiple types of family therapy programs (such as FFT, Multidimensional Family Therapy [MDFT], and Multisystemic Therapy [MST]) could be considered a practice because it examines family therapy practices more generally. Thus, meta-analyses on more specific, individual programs should be distinguished from meta-analyses examining more general, heterogeneous practices. If Senior Researchers have any questions about the appropriateness of a specific meta-analysis as a practice for *CrimeSolutions*, they should contact the Project Director.

In addition, regarding **conflict of interest**, if a meta-analysis includes a primary study conducted by the Senior Researcher responsible for screening the meta-analysis or anyone else on the *CrimeSolutions* project, it has been decided that this is *not* a conflict of interest (COI). A copy of the COI form for *practice* reviews is available in **Appendix B**, and a copy of the COI form for *program* reviews is available in **Appendix C**. These forms must be completed by Reviewers before they begin the review process, to ensure they have no conflicts with the meta-analysis authors, the study authors, or the program developers. These forms should provide the Senior Researchers with clarification on what constitutes conflicts of interest in the review process.

DEFINING A PRACTICE

Senior Researchers should use the *CrimeSolutions* Practice Screening Form (see **Appendix E** for a copy) when defining all practices that are sent forward for review. All meta-analyses are screened to identify an eligible practice. An eligible practice is defined as a set of interventions that share one or more active components designed to address a justice-related problem. An **intervention** may refer to a specific program (e.g., Perry Preschool Project), a more general strategy (e.g., drug courts, community policing), or a government policy (e.g., drunk driving laws). For the purposes of *CrimeSolutions*, an **active component** is the element of an intervention that is expected to produce a therapeutic (or beneficial) effect. Conversely, an inactive component is the inert element of an intervention that by itself is not expected to have any effect.

In the first part of the Practice Screening Form, the Senior Researchers examine the eligibility requirements that must be met for a meta-analysis to be sent forward for review. They include

- The meta-analysis must focus on at least one:
 - Justice or eligible-justice related problem, *or*
 - Any problem involving a justice-involved population.
- The meta-analysis must target one or more active components through
 - The specified inclusion criteria, *or*
 - A defined moderator analysis
- The meta-analysis must calculate an effect size for the targeted component from at least two studies.

For the purposes of *CrimeSolutions*,

- A **justice problem** is any issue that involves the preventing, detecting, prosecuting, sentencing, or punishing of illegal behavior (including behavior related to status offenses) and any issue (i.e., victim assistance, officer training, system improvement, etc.) that involves at least one justice subsystem (i.e., pretrial services; child welfare agencies [juvenile justice system]; trial, family [juvenile justice], and appellate courts; victim service agencies; prosecution and public defender offices; probation and parole agencies; and custodial institutions [jails, prisons, reformatories, halfway houses, etc.]).
- A **justice-related problem** is limited to the following issues: violent and disruptive behavior, relational and social aggression, maltreatment, and legal substance use (e.g., alcohol and tobacco).
- A **justice-involved individual** includes victims of crime and any person who has had interactions with a justice subsystem, including pretrial services; child welfare agencies (juvenile justice system); trial, family (juvenile justice), and appellate courts; prosecution and public defender offices; probation and parole agencies; and custodial institutions (jails, prisons, reformatories, halfway houses, etc.).

In addition, all meta-analyses are screened to identify inclusion criteria that target a specific population, a unique setting, or a specified age range (this information is indicated on the Practice Screening Form).

In the second part of the Practice Screening Form, Senior Researchers screen to ensure the meta-analysis meets the minimum criteria (for more details on the criteria, see the next section on Identifying an Eligible Meta-Analysis).

Once both parts of the Practice Screening Form are completed, the Senior Researcher determines whether the meta-analysis is sent forward for review or not sent forward for review. If the meta-analysis is sent forward for review, the Senior Researcher must make decisions about the outcomes to score (for more details, see the section on Selecting Effect Sizes for the Evidence Summary, p. 18).

Identifying an Eligible Meta-Analysis

A meta-analysis is eligible only if the answer is “yes” to every question below. If even one question is answered as “no,” the meta-analysis is screened out and will be included on the list of [Screened-Out Meta-Analyses for Practices](#). These eligibility items are completed on the *CrimeSolutions* Practice Screening Form (see [Appendix C](#)).

INTERVENTION

Does the meta-analysis include at least two studies of the intervention of interest? The meta-analysis must include at least two program evaluations of the program protocol, program type, program infrastructure, or practice of interest. Meta-analyses typically locate a large number of potentially eligible studies in the literature search, which are then whittled down to the set of studies that meet the final eligibility criteria.

AGGREGATION

Does the meta-analysis aggregate the results from at least two studies? The meta-analysis must report an aggregate effect size. That is, the meta-analysis must aggregate or average the effects of at least two studies into a single estimate of the effectiveness of the treatment. If the meta-analysis presents results individually for at least two studies, but elects not to aggregate results across the studies, Senior Researchers should ***seek advice from an expert meta-analyst*** on whether it is possible and/or justifiable to combine findings from at least two of the included studies.

PRIMARY AIM OF THE INTERVENTION

Do the programs included in the meta-analysis aim to reduce crime, delinquency, overt problem behaviors, or victimization; improve justice system practices; or target an offender or at-risk population? The intervention that is the subject of the meta-analysis must focus on ***at least one*** of the following:

- Aims to prevent or reduce crime or delinquency, overt problem behaviors (such as aggression or gang involvement), or important risk factors for crime or delinquency (such as school failure)
- Aims to prevent, intervene, or respond to victimization
- Aims to improve justice systems or processes
- Targets an offender population or an at-risk population (that is, individuals who have the potential to become involved in the justice system)

LITERATURE SEARCH

Did the literature search include at least two sources and provide evidence that unpublished literature was sought in the search? The meta-analysis must be based on a comprehensive and systematic search of the literature. To meet this criterion at least minimally, the search must include ***both*** of the following components:

1. At least two sources for identifying potentially eligible research must be used. Sources include, for example, different electronic bibliographies, Google-type Internet searches, hand searches of journals, contacting relevant researchers in the field, and reviewing the references in pertinent studies and prior reviews.
 - a. Two clearly specified *different* electronic databases (e.g., *PsycINFO* and *Dissertation Abstracts International*) can be counted as two different sources.
 - b. Meta-search engines, such as EBSCO Host, ProQuest, or Dialog count as two (or more) different sources only if the meta-analysis clearly identifies two or more sources within the meta-search engine that was searched.

- c. Google Scholar as the sole search strategy is not sufficient to meet the two-source minimum criterion.
2. There is some indication that unpublished or “grey” literature was sought during the search, whether any unpublished research reports eventually met the eligibility criteria for inclusion in the meta-analysis. Unpublished or grey literature is defined here as literature that is not controlled by commercial publishers, including for example conference presentations, government and nongovernmental agency technical reports, and theses/dissertations. ***Meta-analyses that explicitly state that only published or peer-reviewed research was included are ineligible. Meta-analyses that explicitly or actively excluded unpublished or non-peer-reviewed research are ineligible.***
 - a. Mention of “fail-safe N” as a way to guard against publication bias is *not* sufficient to meet this criterion.
 - b. Careful review of the literature search documentation and the final bibliography may be necessary to assess the criterion above. Scanning the bibliography can illuminate whether unpublished research such as dissertations, technical reports, and conference presentations were included (or retrieved but deemed ineligible for inclusion in the final meta-analyses). Meta-analyses that mention including documents from the following sources can be assumed to have sought unpublished literature: *ERIC, SIGLE, NTIS, CrimDoc, NCJRS*, master’s theses or dissertations, conference proceedings, government websites, and research firm websites.

PRIMARY OUTCOMES

Does the meta-analysis report on at least one eligible outcome (defined below)? The meta-analysis must report results on at least one eligible outcome, as follows:

- Aims to prevent or reduce crime, delinquency, or related problem behaviors (such as aggression, gang involvement, or school attachment)
- Aims to prevent, intervene, or respond to victimization
- Aims to improve justice systems or processes
- Targets an offender population or an at-risk population (that is, individuals who have the potential to become involved in the justice system)

Eligible outcomes must fall completely within one of the single construct categories listed below and cannot be combinations of constructs across more than one category. That is, constructs that are combinations of education and mental health outcomes would not be eligible unless the education and mental health outcomes were also reported separately.

CONTROL GROUPS

Do all studies included in the meta-analysis include an appropriate control, comparison, or counterfactual condition?

- a. The studies included in the meta-analysis must have appropriate control groups. Specifically, one of the following two conditions must be met:
 - An appropriate counterfactual condition is required for all included studies in a meta-analysis, or
 - In meta-analyses that include studies without control groups, results for controlled studies are reported separately from those of uncontrolled studies.
- b. Senior Researchers should examine the inclusion criteria for the meta-analysis and any descriptive information about the studies that meet those criteria and are included in the meta-analysis. They should determine whether the selected studies have control or comparison groups, or other ways to characterize the counterfactual condition that represent what the outcome would be in the

absence of treatment. To meet these criteria, all studies included in the meta-analysis must estimate intervention effects relative to an empirical estimate of the counterfactual condition or separately report those that meet this condition from those that do not.

- c. In addition, simple before–after comparisons will *not* meet this condition, except in rare instances where evidence or very plausible assumptions indicate that no change would occur absent treatment.

REPORTING OF RESULTS

Does the meta-analysis report effect sizes that represent the magnitude of the treatment effect? The meta-analysis must report a common quantitative index of results that represents the magnitude of the treatment effect (i.e., effect sizes) across included studies. A quantitative synthesis of effect sizes is not required (that is, the meta-analysis does not have to calculate or report an average effect size across multiple studies), as long as individual study results are reported using a common index. In some unusual instances, all the studies in a meta-analysis may use exactly the same outcome measures so that conversion of effects to effect sizes to make them comparable across studies is not necessary. This is acceptable as long as the measures are identical—that is, they are operationalized the same way and produce values on the same exact scale in all the studies.

- a. Meta-analyses that report only the statistical significance or *p* values *of the included studies* as the index of treatment effects are *not* eligible. Meta-analyses may report on the statistical significance of the individual effect sizes they compute and/or on the aggregate mean effect sizes that summarize a number of included studies. What is *not* acceptable are meta-analyses that do not report effect sizes for the included studies, but rather only extract and report the statistical significance or *p* values from the original research reports.
- b. Meta-analyses that report only vote-counting results are not eligible. These types of meta-analyses might list included studies and indicate whether results were significant or whether the programs were effective. These kinds of meta-analyses are *not* eligible. To be eligible, the results of the included studies must be summarized with some form of effect size.

COMBINING DIFFERENT TYPES OF RELATIONSHIPS

If a quantitative synthesis of effect sizes is reported (i.e., a mean effect size is reported for multiple studies), do all effect sizes in the combination index the same type of relationship? If the meta-analysis reports an average mean effect size calculation (for all studies or for subgroups of studies), only effect sizes in the same metric and measuring the same type of relationship should be averaged together. This criterion has to do with the kinds of relationships and effect sizes that are combined, rather than with the types of programs that are combined. A meta-analysis that includes a diverse group of programs would not necessarily be problematic. However, a meta-analysis that combined standardized mean difference effect sizes for treatment-versus-control comparisons with pretest–posttest change effect sizes would be ineligible, unless the different effect size means were reported separately.

- a. Similarly, meta-analyses that combine treatment-versus-treatment effect sizes with treatment-versus-control effect sizes would also be ineligible, unless the different effect size means were reported separately. For example, a meta-analysis that combines effect sizes from studies comparing cognitive–behavioral therapy (CBT) with “no treatment” or “practice as usual” control groups with effect sizes from studies that compare CBT with family therapy, without presenting separate averages for the two groups, would *not* be eligible.

- b. It is acceptable under this criterion to combine effect sizes in the same metric from randomized controlled trials with effect sizes from quasi-experiments because both types of designs would lend themselves to the same kinds of relationships (i.e., treatment versus control differences).

PUBLICATION DATE

Were at least 50 percent of the studies included in the meta-analysis published/available in or after 2000?

At least 50 percent of the studies included in the meta-analysis had to have been published or first available in or after 2000.

AGE OF SAMPLES

Are the included samples in the meta-analysis restricted to either adults or juveniles? Or are mean effect sizes for adults and juveniles reported separately? All studies included in the meta-analysis must involve offender participants that are either all juvenile or all adult. If the meta-analysis includes studies that involve both juveniles and adults, mean effect sizes must be reported separately for the juvenile and adult studies.

- a. Meta-analyses that include studies that themselves combine juveniles and adults in the same programs without providing any options for splitting the juvenile and adult samples would **not** be eligible.
- b. The age criterion applies for meta-analyses of practices in the courts, corrections, drug/substance misuse, and juvenile justice topic areas. The age criterion does not apply for 1) meta-analyses of practices that focus on areas, hot spots, surveillance, and the like, which do not typically involve identifiable offenders, and 2) victim interventions.

Selecting Effect Sizes for the Evidence Summary

At this stage of coding, Senior Researchers should select the mean effect sizes that they will use for the evidence summary. There are several decision points to think about. In each case, they should make a decision about which mean effect sizes to carry forward and record only those effect sizes (and their accompanying ratings) in the evidence summary.

SELECTING SUBGROUPS

Some eligible subgroupings should be coded. For example, if reported in a meta-analysis, subsets of studies categorized by age (e.g., if a meta-analysis reports separate mean effect sizes for juvenile and adult participants) **must be coded separately**. If juvenile and adult subgroups are not reported separately in a meta-analysis, then the meta-analysis is **not** eligible for coding.

- a. Multiple outcomes may be reported in a meta-analysis in addition to any subgroupings of program types. Thus, multiple combinations of outcomes and subgroups may be possible.
- b. Subsets of studies categorized by program type may also be coded separately. For example, a meta-analysis of family-based therapies may include subgroup analyses that report separate effect sizes for each type of therapy (i.e., FFT, MDFT, MST, etc.). These subgroups may be coded separately; **however, they will not affect the overall outcome rating**. Instead, this information will be included in the description of moderator analyses results, in the Other Information section of the practice profile.

SELECTING FROM AMONG MULTIPLE OUTCOMES

Only two types of outcomes are eligible for coding: 1) general crime or delinquency, and 2) focal outcomes that are the intended targets of the interventions.

- a. Senior Researchers should identify the general crime/delinquency outcome in the meta-analysis. If more than one general offense outcome is reported (for example, more than one measure of recidivism is reported), Senior Researchers must select one option using the following criteria: 1) select the broadest or most general construct in the category or, if that cannot be determined, then 2) select the option with the largest number of studies. If neither of these two options results in a choice, Senior Researchers should consult an expert meta-analyst for assistance in selecting an effect size.
- b. For any program type, noncrime/nondelinquency outcomes may also be coded if the outcomes are the intended target (or targets) of the intervention, and the participant population is defined or identified by the targeted outcome (e.g., truancy outcomes from truancy interventions for chronic truants, achievement outcomes from tutoring programs for underachieving delinquents, behavior problem outcomes from violence prevention programs for children with conduct disorder).

For each effect size selected for coding, Senior Researchers should use the Outcome Construct Categories taxonomy on the following page and a half to categorize the constructs represented in each outcome effect size. If they cannot determine which Level 2 construct to use, then they should code the outcome under the “Other” Level 2 category, and work with DSG staff to make these final determinations.

Outcome Construct Categories	
Level 1 (Macro)	Level 2 (Micro)
Crime/Delinquency	<ul style="list-style-type: none"> • Multiple crime/offense types • Property offenses • Public order offenses • Sex-related offenses • Violent offenses • Drug and alcohol offenses • Status offenses • Terrorism and mass violence offenses • Technical violations • Other crime/offense types
Drugs and Substance Abuse	<ul style="list-style-type: none"> • Multiple substances (e.g., alcohol, drug use) • Alcohol • Cocaine/crack cocaine • Heroin/opioids • Marijuana • Methamphetamines • Other substances
Mental/Behavioral Health	<ul style="list-style-type: none"> • Multiple mental/behavioral health outcomes • Internalizing behavior (e.g., depression, anxiety) • Externalizing behavior (e.g., antisocial behavior, aggression, self-control) • Trauma/PTSD • Psychological functioning • Emotional well-being • Other mental/behavioral health
Education	<ul style="list-style-type: none"> • Multiple education outcomes • Academic achievement/school performance • Attendance/truancy

	<ul style="list-style-type: none"> • Graduation • Degree completion (e.g., GED) • Dropout • Expulsion/suspension • Other education
Victimization	<ul style="list-style-type: none"> • Multiple victimization outcomes • Domestic violence/intimate partner violence/family violence • Child abuse/neglect/maltreatment • Sexual abuse/exploitation • Resiliency • Hospitalizations/emergency room visits • Perceived social support • Self-concept/self-esteem • Other victimization
Family	<ul style="list-style-type: none"> • Multiple family functioning outcomes • Family functioning (e.g., adaptability, cohesion, communication) • Out-of-home placement/permanency • Parenting skills (e.g., behavior management skills) • Other family
Employment/Socioeconomic Status	<ul style="list-style-type: none"> • Multiple employment/socioeconomic status outcomes • Job placement • Job retention • Monthly/yearly earnings • Other employment/socioeconomic status
Juvenile Problem or At-Risk Behaviors	<ul style="list-style-type: none"> • Multiple juvenile problem/at-risk behaviors • Teen pregnancy • Running away • Ungovernability/incorrigibility • Association with antisocial peers • Aggression • Bullying • Other problem behavior
Attitudes/Beliefs	<ul style="list-style-type: none"> • Multiple attitudes/beliefs • Fear of crime • Antisocial beliefs/attitudes (e.g., favorable attitude toward drug use) • Legitimacy of police • Victim satisfaction (with restorative justice or justice system process) • Offender satisfaction (with restorative justice or justice system process) • Knowledge • Other attitudes/beliefs
Justice Systems/Processes	<ul style="list-style-type: none"> • Multiple justice systems or processes outcomes • Processing of offenders • Compliance with restitution/fines/payments • Other justice systems/processes

SELECTING EFFECT SIZES FROM DIFFERENT RESEARCH DESIGNS

If a meta-analysis reports mean effect sizes separately for randomized controlled trial (RCT) and non-RCT studies for the program (or programs) or populations (juvenile/adult) of interest, there will be circumstances in which only the mean effect size from the RCT studies should be carried forward. For example, if the mean effect size from the non-RCT studies has sufficient potential for bias, then it should be excluded from further consideration and only the mean from the RCTs should be carried forward.

- a. If any of the following conditions are met, then there is sufficient potential for bias and only the mean effect size for the RCT studies should be carried forward if it is separately reported or can be computed:
 - There are at least five RCTs, and a statistical test is reported that shows a statistically significant difference between the mean effect size for the RCT and non-RCT studies.
 - There are at least five RCTs, and the mean effect sizes for both the RCTs and non-RCTs do *not* fall within an approximate fixed effect 95 percent confidence interval around the mean effect size for both combined. If necessary, Senior Researchers should compute the combined mean as the weighted average of the RCT and non-RCT means, with each weighted by the number of studies contributing to the respective mean. The confidence intervals are defined as follows, based on the total number of studies in the combined mean:
 - 10 studies or fewer: $\pm .12$ around the combined mean
 - 11–20 studies: $\pm .09$ around the combined mean
 - 21–30 studies: $\pm .07$ around the combined mean
 - 31–50 studies: $\pm .06$ around the combined mean
 - 50 studies or more: $\pm .05$ around the combined mean
- b. If none of the above is the case, the mean effect size for the RCT and non-RCT studies combined should be the one carried forward to the validity and final Evidence Ratings. If there is no overall mean reported in the meta-analysis, the section below provides guidelines on effect-size combinations.
- c. In cases in which the RCTs and non-RCTs are not reported separately, the combined mean effect size may be carried forward, and evidence ratings should be based on the combined group of studies.
 - If it is necessary to compute the weighted average effect size for the RCT and non-RCT means to determine whether the separate or combined means are eligible for coding, Senior Researchers should use the following calculation:

$$\overline{ES}_{RCT \& non-RCT} = \frac{(ES_{RCT} * n_{RCT}) + (ES_{non-RCT} * n_{non-RCT})}{(n_{RCT} + n_{non-RCT})}$$

Where ES_{RCT} is the mean effect size for the RCT studies, $ES_{non-RCT}$ is the mean effect size for the non-RCT studies, n_{RCT} is the number of studies used to calculate the RCT mean effect size, and $n_{non-RCT}$ is the number of studies included in the non-RCT mean effect size.

SELECTING EFFECT SIZES FROM DIFFERENT FOLLOW-UP PERIODS

Some meta-analyses may report a specific effect size at different follow-up periods (for example, they report an effect size for recidivism at 6 months, 1 year, and 2 years separately) rather than provide an overall effect size across the different follow-up periods. If a meta-analysis reports mean effect sizes for the same outcome separately at different follow-up periods, Senior Researchers should choose the most distal effect size available per outcome.

For example, a meta-analysis (Zane et al. 2016) of studies on the effects of transferring juveniles to adult court looked at the length of follow-up periods as part of the moderator analyses.* Overall effect sizes for felony recidivism and violent felony recidivism were provided, and these outcomes were scored. In this instance, the effect sizes related to the follow-up periods would not be selected for review, although the findings could be included in the Other Information section of the practice profile. Conversely, in a meta-analysis by Armelius and Andreassen (2007) on cognitive behavioral programs for antisocial youth, which specifically reported recidivism effect sizes at 6 months', 12 months', and 24 months' posttreatment, no overall effect size for recidivism was provided. In this instance, the recidivism effect size at the 24-month follow-up period was scored.†

SELECTING A MEAN EFFECT SIZE WHEN BOTH RANDOM EFFECTS AND FIXED EFFECTS MODELS ARE REPORTED

Some meta-analyses will report both fixed effect and random effects means. *Always select the random effects mean*, even if the authors indicate that the fixed effect mean is preferred for some reason. The only time fixed effect means are eligible for coding is when a random effects mean is not available. In cases where an overall random effects mean is reported on a group of studies that should not be combined for purposes of *CrimeSolutions* coding (e.g., juveniles and adults are combined) and the subgroup means are reported as fixed effects, it is appropriate to select the fixed effect means in this case because the overall mean is not eligible for coding.

OUTLIERS

In addition to outcome type, design type, program category, program subgroups, and age subgroups, other situations may arise when it will be necessary to select one of several specific effect sizes to code. Some of the most common situations will be as follows:

Outlier analysis. If the meta-analysis reports effect size means with and without outliers, Senior Researchers should take the mean with the outlier (or outliers) removed over the mean that includes an outlier (or outliers), if possible.

Influence analysis. Some meta-analyses will conduct an influence analysis or “one-study removed analysis,” reporting a number of different mean effect sizes, each with a different study removed. These analyses are used to determine the influence of individual studies on the overall mean effect size; that is, they can be used to identify outliers if the mean effect size changes drastically when individual studies are removed. When multiple means are presented in such an analysis, Senior Researchers should not use them for coding, but take the overall mean that includes all the effect sizes. If, however, the influence/one-study removed analysis identifies an obvious outlier, which the author then removes to compute a mean, that mean should be selected for coding over the mean that includes the outlier.

*Zane, Steven N., Brandon C. Welsh, and Daniel P. Mears. 2016. “Juvenile Transfer and the Specific Deterrence Hypothesis: Systematic Review and Meta-Analysis.” *Criminology & Public Policy* 15(3):1–25.

†Armelius, Bengt-Åke, and Tore Henning Andreassen. 2007. “Cognitive–Behavioral Treatment for Antisocial Behavior in Youth in Residential Treatment.” *Campbell Systematic Reviews* 8.

Sending the Meta-Analyses Forward for Review

Senior Researchers should use the *CrimeSolutions* Practice Screening Form (see **Appendix E** for a copy) when screening all practices that are sent forward for review. As described above, the screening form requires Senior Researchers to define the practice and ensure the meta-analysis meets all minimum eligibility criteria, and to decide ultimately whether the meta-analysis should go forward for review. Once the form is sent back confirming the meta-analyses will go forward for review, the Project Director or Deputy Project Director will send an email with space to fill in information on the outcomes to score and the Practices Reviewers who should complete the review. ***Senior Researchers should be as specific as possible when describing the outcomes/effect sizes selected for scoring, such as providing page numbers or table numbers.*** This should ensure that the Practice Reviewers score the same outcomes. Finally, Senior Researchers should include the names of two Practice Reviewers and a backup Reviewer who would be appropriate for reviewing the practice.

Reviewing Final Ratings of Programs and Practices

Once a program or practice has returned from review, the DSG Research Assistant in charge of the review will send out a copy of the Reviewers' instruments for a final Senior Researcher review. If a dispute resolution is needed, the relevant information will be provided in an Excel file (see below for more information about addressing disputes). If the Reviewers agree on the final program or practice/outcome ratings, and no dispute resolution is needed, Senior Researchers should still double-check that Reviewers correctly scored the studies or meta-analyses. This is especially important for studies or meta-analyses rated Class 5—Inconclusive Evidence. Senior Researchers should review the section or sections on the instrument that resulted in the Class-5 rating and ensure that the Reviewers scored the items correctly, while considering that Reviewers are allowed discretion on certain items.

If Senior Researchers have any hesitations or concerns about the scoring of a specific item or items, they should reach out to both Reviewers with their concerns. They should also include the Research Assistant on the email, so the process can be properly documented, and all emails are saved as needed. This may result in a brief discussion with the Reviewers and overall agreement in the Class 5 rating, or the Reviewers may end up changing some of their scores and re-rating the program or practice. If they do change their scores, Reviewers must submit a new instrument to the Research Assistant, or they must provide permission for the Research Assistant to make the changes for them. This process is to ensure that all programs and practices receive the correct final ratings, and to ensure that appropriate studies or meta-analyses receive the Class 5 rating.

Dispute Resolution Process for Programs and Practices

In instances where the Study Reviewers have disagreements in classifying the final rating of a study or outcome, the Research Assistant will send information about the dispute to the Senior Researcher, including 1) an Excel spreadsheet that documents both Study Reviewers' scores for each item on the instrument and clearly marks which scores differ and contribute to the disparity in the overall study or outcome rating, 2) copies of the Study Reviewers' scoring instruments, and 3) a copy of the study or meta-analysis in dispute. Programs may also require a dispute resolution on the Study Design scores, to determine whether a program receives the tag for the Randomized Controlled Trial (RCT); see p. 24 below for more details.

With this information, the Senior Researcher will take the following steps, in sequential order:

STEP 1. The Senior Researcher will consult with the Study Reviewers to resolve the disparity on those dimensions in dispute. The Senior Researcher will review the information provided by the Research Assistant to determine where there is disagreement between the Study Reviewers that must be resolved to reach consensus on the final study rating. The Senior Researchers do have discretion on how the dispute may be resolved. For instance, the Senior Researcher may point out the differences in scores to the Reviewers and ask that they resolve the dispute between themselves. Or, the Senior Researcher may also provide any guidance to the Reviewers that he or she thinks is pertinent to the scores in question and be a part of the conversation between the Reviewers.

Senior Researchers should encourage a dialog between the Reviewers, in which both Reviewers explain their reasoning behind their different scores, and then reach a consensus on a final study or outcome rating. Ultimately, Senior Researchers should try to have both Reviewers provide input in the dispute process and avoid having only one Reviewer provide an explanation for their scores and the other Reviewer simply agreeing without elaboration. If the disparity can be resolved, the Senior Researcher will instruct the Reviewer (or Reviewers) to submit a corrected Program or Practice Scoring Instrument to the DSG Research Assistant coordinating the review, reflecting the resolved score (or scores). The dispute resolution can occur in two forms:

By email with the Reviewers: The Senior Researcher will discuss and resolve all disputes in the scoring of the program's evidence base through email. The Senior Researcher will then forward the email thread between the Reviewers to the Research Assistant, so that the written documentation about how the problem was resolved can be added to the program's file in the *CrimeSolutions* Administrative System.

By phone with the Reviewers: The Senior Researcher will hold a conference call with the Reviewers to discuss and resolve all disputes in the scoring of the program's evidence base. The Senior Researcher will then send an email to the Research Assistant, noting exactly which dimensions were discussed, what was decided, and who will be submitting new scoring instruments. This email will be added to the program's file in the *CrimeSolutions* Administrative System.

STEP 2. If Step 1 fails to resolve the disparity, the Senior Researcher will score the study or meta-analysis and serve as a tiebreaker. He or she will submit a completed scoring instrument, which includes notes concerning why he or she agreed with one Reviewer's score over the other Reviewer's score. The Senior Researcher will replace the Reviewer and therefore will need to complete the entire scoring instrument for the study or meta-analysis under dispute (and any other studies or meta-analyses included in the program's or practice's evidence base). The Senior Researcher's instrument will be added to the program or practice documentation, along with the instrument of the Reviewer whom the Senior Researcher agrees with.

Reviewing Profiles of Accepted Programs and Practices

All accepted programs and practices (i.e., those rated Class 1–4) will receive a profile to be posted to *CrimeSolutions*. The profiles are developed by the DSG Research Assistants, and go through many rounds of writing, reviewing, and editing before completed. At the end of the development process, after the profile has received a final edit, Senior Researchers will receive a copy for review, along with a copy of the scored study or meta-analysis. They may also receive copies of the Reviewers' instruments if they want to review those as well. Senior Researchers will need to give a final review of the profile, to ensure information is accurate. They should pay special attention to the Program/Practice Description, the Methodology, and the Outcome Evidence, to ensure thoroughness and accuracy, compared with the scored study or meta-analysis. They

should also pay particular attention to any questions from the Research Assistant, if there was a part of the study that was hard to understand or translate for the profile. Senior Researchers should also ensure that the profiles are written in nontechnical, user-friendly language for the primary target audience of *CrimeSolutions* (i.e., policymakers and practitioners).

If there are any edits or questions, Senior Researchers should put these into “track changes” or comment boxes and send the document back to the Project Director or Research Assistant for review. No program or practice profiles can be posted to the *CrimeSolutions* live site without a final Senior Researcher review, so we ask that Senior Researchers please keep reviews timely.

Programs receiving the RCT and/or Multisite Tags. Programs may receive one or both tags that appear at the top of the program profile: the RCT Tag or the Multisite Tag.

For *CrimeSolutions*, to receive the Multisite Tag a program must be evaluated in more than one site across multiple studies or evaluated at more than one site within a single study. The term “site” refers to the location of an evaluation that examines the effectiveness of a specific program. For purposes of evaluation and assessment in *CrimeSolutions*, the term “site” includes the following three elements: 1) geographic location (determined by factors such as physical location or boundaries or contextual variations); 2) jurisdictional or organizational independence of implementation (determined by factors such as independence of decisionmakers); and 3) population independence or uniqueness (determined by factors such as race/ethnicity and socioeconomic status).

To receive the RCT Tag, at least one study in the program’s evidence base must be an RCT. Further, the RCT study must have a Design Quality score of 2.0 or higher from both Study Reviewers. The RCT study must also have outcome evidence in the same direction as the program’s overall rating (for example, if a program has multiple studies in the evidence base and an overall rating of Promising, but the RCT study is rated Class 4 | No Effects, then the program would *not* receive the RCT Tag). A dispute resolution may be necessary if the two Study Reviewers differ on the Design Quality score of an RCT study. For instance, if a program has an overall rating of Promising, and one Reviewer scored the Design Quality of the RCT study in a program’s evidence as 2.0 and the other Reviewer scored the Design Quality as 1.8, a dispute resolution would be needed. Ordinarily, these scores would not require a dispute resolution with regard to a final rating of Promising. But because the program could receive the RCT Tag if the second Reviewer upgraded their score to 2.0 or higher, a dispute resolution would now be needed.

This information will be checked off on the program template by the Research Assistants developing the program profile.

Replacing programs with practices. In some instances, a new practice may replace programs currently rated on *CrimeSolutions*. The following steps should be used to determine whether a program should be removed and replaced by a practice:

1. **Use the same standards for determining when two (or more) studies are examining the same program (as listed below) to compare the program and practice descriptions.** If the six criteria (from the Program Scoring Instrument) match across the program and practice profiles, then the program should be subsumed under the practice. This designation is favored because meta-analyses are considered more rigorous than individual studies. These six criteria are as follows:
 - a. Logic of the program.

- b. Details of all key components: Consider these on a case-by-case basis.
 - c. Frequency/duration of program activities.
 - d. Targeted population.
 - e. Targeted behaviors.
 - f. Settings: The emphasis should be on the delivery setting, but the geographical setting may also be considered.
2. **Consider the value-added information in the program profile.** Even if the program and practice descriptions match on the six items above, the program profile may provide additional pertinent information that could help users better understand the intervention (i.e., the program profile provides a more detailed intervention description, or the program evaluation looked at additional outcomes that the practice meta-analysis did not). If so, then the program should *not* be subsumed under the practice.
 3. **Consider the number of studies in the program's evidence base.** If a program has a large number of studies in its evidence base (for instance, more than three studies), DSG Research Assistants may search for applicable or relevant meta-analyses that could replace the program as a practice. Programs that have a large evidence base may then be considered for removal if there are eligible meta-analyses for review and after the Research Assistants have gone through steps 1 and 2.

Although the DSG Research Assistants in charge of developing practice profiles may make suggestions on which programs to remove, all final decisions on removal will be made jointly by NIJ and the Senior Researcher in charge of the practice review.

Contact Information

Project Manager

Marcia Cohen, President, Development Services Group (DSG), Inc., 301.951.0056,
mcohen@dsgonline.com

Project Director

Rachel Stephenson, Project Director, DSG, Inc., 301.951.5385, rstephenson@dsgonline.com

Deputy Project Director

Rachel Keller, Deputy Project Director, DSG, Inc., 301.951.5384, rkeller@dsgonline.com

Senior Researchers

- **Juvenile Justice**—**Stephen Gies**, Senior Research Scientist, DSG, Inc., 301.951.6600, sgies@dsgonline.com (for juvenile practices), and **Fran Harmon**, Senior Research Scientist, DSG, Inc., 301.951.5371, FHarmon@dsgonline.com (for juvenile programs).
- **Crime & Crime Prevention**—**Rob Guerette**, Professor, Florida International University, 305.348.0424, rtguerette@gmail.com
- **Corrections & Reentry and Courts**—**Alex Holsinger**, Associate Dean and Professor, Criminal Justice and Criminology, University of Missouri | Kansas City, 816.235.1132, holsinger.alex@gmail.com, and **Eoin B. Healy**, Senior Research Scientist, DSG., Inc., 240.330.4691, ehealy@dsgonline.com
- **Law Enforcement**—**Edward Maguire**, Professor, School of Criminology and Criminal Justice, Arizona State University, 703.606.5752, edmaguire@asu.edu
- **Drugs and Substance Abuse**—**Carol Petrie**, DSG, Inc., former Director, Committee on Law and Justice, National Research Council, and former Acting Director, National Institute of Justice, 301.585.3128, cpetrie@dsgonline.com
- **Victims and Victimization**—**Alissa Worden**, Professor, School of Criminal Justice, University of Albany, SUNY, 518.442.5213, aworden@albany.edu
- **National Mentoring Resource Center**—**David DuBois**, Professor, University of Illinois at Chicago, 312.413.9806, dldubois@uic.edu

Research Assistants

Amanda Bobnis, Research Associate, DSG, Inc., 301.951.5382, abobnis@dsgonline.com

Brittney Clemons, Research Assistant, DSG, Inc., 240.330.4680, BClemons@dsgonline.com

Scottie Whiteley, Research Assistant, DSG, Inc., 301.951.6606, swhiteley@dsgonline.com

Appendix A.

Outline for Including New Types of Training Programs on CrimeSolutions

Introduction

The National Institute of Justice (NIJ) tasked Development Services Group, Inc. (DSG), with reconsideration of a CrimeSolutions (CS) policy regarding the inclusion of training programs. Currently, the Senior Researcher Procedures Manual states: “Police- or correctional-officer wellness programs are excluded. Training programs for police officers, correctional officers, or other justice staff are also excluded, unless the study examines the impact of training on officers’ behaviors” (p. 8).

However, in keeping with the agency’s efforts to promote the safety, health, and wellness of individuals affected by, or employed within, the criminal justice system, as described in NIJ’s *Safety, Health, and Wellness Strategic Research Plan 2016–2021* (2016), there may be other types of programs in that plan that NIJ would consider for inclusion on CS in the future.

Main Discussion

The lists below on target populations and outcomes of interest pertaining to training, safety, health, and wellness programs will need to be considered during discussions between NIJ and DSG. NIJ has stated that it does not wish to change the current scope of CS; thus, any new types of programs must fall within the scope of CS, as follows:

- Aim to prevent or reduce crime, delinquency, or related problem behaviors (such as aggression, gang involvement, or school attachment).
- Aim to prevent, intervene, or respond to victimization.
- Aim to improve justice systems or processes.
- Target an offender population or an at-risk population (that is, individuals who have the potential to become involved in the justice system).

Definition

The goal is to think about the behavioral impacts of programs focused on the training, safety, health, and wellness of police/correctional officers and other criminal justice actors, as such programs relate to the scope of “improving justice systems or processes.” To improve the justice system, the behaviors, health, safety, and wellness of the professionals in that system should also be improved because of the potential adverse outcomes unique to this population (for example, high stress or aggression in police officers could lead to an overuse of force against suspects/civilians, or trauma/posttraumatic stress disorder [PTSD] symptoms in officers could lead to subsequent suicide ideations and attempts). Evaluations of these programs must provide clear and explicit explanations about the goals and intentions of the program, and how the theory of change is unique to the characteristics of the targeted criminal justice actors. The potential target populations and outcomes of interest are listed below. The eligibility of some outcomes of interest will need to be determined on a case-by-case basis.

Target Populations

Potential target populations for training programs include

1. Police officers

2. Judges
3. Attorneys (e.g., prosecutors, defense attorneys)
4. Corrections officers
5. Probation/parole/community supervision officers
6. Other criminal/juvenile justice personnel
7. Individuals contracted by a justice agency to serve in an official capacity in the justice system or to provide onsite services to justice-involved individuals

Training programs that target prevention, intervention, treatment, or response by persons not employed by the justice system would *not* be eligible for review (unless the training program targets individuals listed in target population no. 7).

Outcomes

An evaluation of a training program is eligible for review only if it includes *at least one primary outcome of interest that is a measure of a changed systemic, behavioral outcome.*

Primary outcomes of interest include

1. Behavioral outcomes related to the population targeted by the training (e.g., improving interactions between police officers and youths or improving other aspects of work performance)
2. Behavioral outcomes related to the justice-involved population affected by the training (e.g., offenders, at-risk juveniles, victims)
3. Safety-related outcomes (e.g., reduction in traffic accidents involving police officers)
4. Suicide rates of police/correctional officers

Secondary outcomes of interest (see footnote)* include

1. Measures related to behavioral/mental health, physical health, or wellness of criminal/juvenile justice personnel (determined on a case-by-case basis), such as
 - a. Measures of stress (including self-report or medical/physical/biological indicators)
 - b. Symptoms or diagnoses of PTSD/trauma
 - c. Symptoms or diagnoses of internalizing and externalizing behaviors (e.g., depression, aggression)
 - d. Measures related to physical well-being (e.g., cardiovascular health, sleep quality)
2. Attitudes, beliefs, perceptions
3. Knowledge
4. Job satisfaction

*If an evaluation includes *only* secondary outcomes of interest, then the evaluation is *not* eligible for review. For example, a study of a stress-reduction training program for police officers that measures reductions in stress but does not tie those reductions to a behavioral outcome, such as work performance, would not be eligible.

Appendix B.

DISCLOSURE OF POTENTIAL CONFLICT OF INTEREST (PRACTICES)

As a *CrimeSolutions* reviewer, I agree that I will immediately notify Development Services Group, Inc. (DSG), if I am aware of **any** of the following potential conflicts of interest with practices, programs, or meta-analyses I am asked to review.

I understand that checking items 1 through 8 below does not automatically preclude my participation in a particular review. However, if I have indicated a potential conflict of interest relating to a practice, program, or meta-analysis, I understand that I must immediately alert DSG and obtain clearance from the Office of Justice Programs (OJP), before I begin a review of that practice, program or meta-analysis.

I am asked to review meta-analysis(es) of: (insert practice name)

For each item you check, please describe the nature of the potential conflict following that item.

1. ___ I cannot review a practice or meta-analysis objectively or feel that others would perceive that I cannot review a practice or meta-analysis objectively.
2. ___ I, my spouse, child, or other family member or business partner is listed as a staff member, consultant, or advisor on any practices or meta-analyses I am asked to review.
3. ___ I have or have had a close personal or familial relationship with the author(s)/staff on any of the practices or meta-analyses studies I am asked to review.
4. ___ I have, or within the past year have had a financial interest with the author(s)/staff or organization on any of the practices or meta-analyses I am asked to review.
5. ___ I have or within the past year have had a professional relationship with the meta-analysis/meta-analyses' author(s)/ on evaluations I have reviewed, including serving as a trustee, board member, officer, or past officer of the meta-analysis author's organization. If checked, please state the nature of the professional relationship: _____
6. ___ I am or within the past year have been a faculty member or employee of the organization submitting or funding the practices or meta-analyses I am asked to review.
7. ___ I have within the past year received a gift or other similar gratuity from the author(s)/staff from an organization that submitted a meta-analysis I am asked to review.
8. ___ I, my spouse, child or other family member or business partner, am employed by or will be seeking employment with an organization that funded, conducted, or currently employs a Principal Investigator (PI) associated with a meta-analysis I am asked to review.
9. ___ **I do not have a conflict with the programs or evaluation studies identified above.**

NAME: _____

DATE: _____

SIGNATURE: _____

If you have checked #9, indicating no potential conflicts of interest, you may return this form with your review packet. If you determine that you have a potential conflict of interest, please contact Marcia Cohen at mcohen@dsgonline.com and do not review the practice or meta-analysis until clearance is obtained from OJP.

Appendix C.

DISCLOSURE OF POTENTIAL CONFLICT OF INTEREST (PROGRAMS)

As a *CrimeSolutions* reviewer, I agree that I will immediately notify Development Services Group, Inc. (DSG), if I am aware of **any** of the following potential conflicts of interest with programs or evaluation studies I am asked to review.

I understand that checking items 1 through 8 below does not automatically preclude my participation in a particular review. However, if I have indicated a potential conflict of interest relating to a program or evaluation study, I understand that I must immediately alert DSG and obtain clearance from the Office of Justice Programs (OJP), before I begin a review of that program or evaluation study.

I am asked to review study(ies) of: (insert program name)

For each item you check, please describe the nature of the potential conflict following that item.

1. ___ I cannot review a program or evaluation study objectively or feel that others would perceive that I cannot review a program or evaluation study objectively.
2. ___ I, my spouse, child, or other family member or business partner is listed as a staff member, consultant, or advisor on any programs or evaluation studies I am asked to review.
3. ___ I have or have had a close personal or familial relationship with the author(s)/staff on any of the programs or evaluation studies I am asked to review.
4. ___ I have, or within the past year have had a financial interest with the author(s)/staff or organization on any of the programs or evaluation studies I am asked to review.
5. ___ I have or within the past year have had a professional relationship with the study's/studies' author(s)/ on evaluations I have reviewed, or with the subject program, including serving as a trustee, board member, officer, or past officer of the evaluation author's organization or the subject program. If checked, please state the nature of the professional relationship: _____
6. ___ I am or within the past year have been a faculty member or employee of the organization submitting or funding the programs or evaluation studies I am asked to review.
7. ___ I have within the past year received a gift or other similar gratuity from the author(s)/staff from an organization that submitted a program or evaluation study I am asked to review.
8. ___ I, my spouse, child, or other family member or business partner, am employed by or will be seeking employment with an organization that is the subject program or an organization that funded, conducted, or currently employs a Principal Investigator (PI) associated with an evaluation study I am asked to review.
9. ___ **I do not have a conflict with the programs or evaluation studies identified above.**

NAME: _____

DATE: _____

SIGNATURE: _____

If you have checked #9, indicating no potential conflicts of interest, you may return this form with your review packet. If you determine that you have a potential conflict of interest, please contact Marcia Cohen at mcohen@dsgonline.com and do not review the program or evaluation until clearance is obtained from OJP.

Appendix D.

CRIME SOLUTIONS PROGRAM SCREENING FORM

NAME OF PROGRAM:	
SENIOR RESEARCHER:	
DATE OF SCREENING:	

PROGRAM GOALS			
GOAL #	PROGRAM GOAL	SOURCE	PAGE OR TABLE #
1			
2			
3			
4			
5			

STUDY GOALS			
GOAL #	PROGRAM GOAL	SOURCE	PAGE OR TABLE #
1			
2			
3			
4			
5			

STUDY 1 CITATION:					
PRIMARY OUTCOMES	OUTCOME NAME	PAGE OR TABLE #	CATEGORIZATION FROM TIERED OUTCOMES LIST	TIER #	GOAL #
1					
2					
3					
4					
5					
SECONDARY OUTCOMES	OUTCOME NAME	PAGE OR TABLE #	CATEGORIZATION FROM TIERED OUTCOMES LIST	TIER #	GOAL #
1					
2					

3					
4					
5					
ELIGIBLE OUTCOMES NOT SELECTED FOR SCORING (PROVIDE JUSTIFICATION FOR EACH OUTCOME NOT SELECTED)					
ADDITIONAL NOTES					

STUDY 2 CITATION:					
PRIMARY OUTCOMES	OUTCOME NAME	PAGE OR TABLE #	CATEGORIZATION FROM TIERED OUTCOMES LIST	TIER #	GOAL #
1					
2					
3					
4					
5					
SECONDARY OUTCOMES	OUTCOME NAME	PAGE OR TABLE #	CATEGORIZATION FROM TIERED OUTCOMES LIST	TIER #	GOAL #
1					
2					
3					
4					
5					
ELIGIBLE OUTCOMES NOT SELECTED FOR SCORING (PROVIDE JUSTIFICATION FOR EACH OUTCOME NOT SELECTED)					
ADDITIONAL NOTES					

STUDY 3 CITATION:					
PRIMARY OUTCOMES	OUTCOME NAME	PAGE OR TABLE #	CATEGORIZATION FROM TIERED OUTCOMES LIST	TIER #	GOAL #
1					
2					
3					
4					
5					

SECONDARY OUTCOMES	OUTCOME NAME	PAGE OR TABLE #	CATEGORIZATION FROM TIERED OUTCOMES LIST	TIER #	GOAL #
1					
2					
3					
4					
5					

ELIGIBLE OUTCOMES NOT SELECTED FOR SCORING (PROVIDE JUSTIFICATION FOR EACH OUTCOME NOT SELECTED)

ADDITIONAL NOTES

ADDITIONAL STUDIES (NOT SENT FORWARD FOR REVIEW AND REASONS WHY)

SUPPLEMENTAL DOCUMENTS (CITATIONS)

STUDY REVIEWERS	
REVIEWER 1	
REVIEWER 2	
BACK-UP REVIEWER	

Appendix E.

CRIME SOLUTIONS PRACTICE SCREENING FORM

META-ANALYSIS TITLE:	
AUTHORS:	
YEAR:	

Step 1. Eligibility Requirement

All meta-analytic studies are screened to identify an eligible practice. An eligible practice is defined as a set of interventions* that share one or more active components† designed to address a justice-related problem. Specifically, to be eligible for inclusion in CrimeSolutions (CS), the meta-analytic study **must**

1. Focus on at least one of the following:
 - 1.1. A **justice**‡ or **eligible justice-related**§ **problem** or
 - 1.2. Any problem involving a **justice-involved**** **population**
2. Target one or more **active components** through
 - 2.1. The specified inclusion criteria or
 - 2.2. A defined moderator analysis
3. Calculate an effect size for the targeted component from at least two studies.

Further, all meta-analytic studies are screened to identify inclusion criteria that target a **specific population**, a unique **setting**, or a specific **age range**.

Step 2. Minimum Criteria Requirement

*An intervention may refer to a specific program (e.g., the Perry Preschool Project), a more general strategy (e.g., drug courts, community policing), or a government policy (e.g., drunk driving laws).

†For the purposes of CS, an active component is the element of an intervention that is expected to produce a therapeutic (or beneficial) effect. Conversely, an inactive component is the inert element of an intervention that by itself is not expected to have any effect.

‡For the purposes of CS, a justice problem is any issue that involves the preventing, detecting, prosecuting, sentencing, or punishing of illegal behavior (including behavior related to status offenses) and any issue (i.e., victim assistance, officer training, system improvement, etc.) that involves at least one justice subsystem (i.e., pretrial services; child welfare agencies [juvenile justice system]; trial, family [juvenile justice], and appellate courts; victim service agencies; prosecution and public defender offices; probation and parole agencies; and custodial institutions [jails, prisons, reformatories, halfway houses, etc.]).

§For the purpose of CS, a justice-related problem is limited to the following issues: violent and disruptive behavior, relational and social aggression, maltreatment, and legal substance use (e.g., alcohol and tobacco).

**For the purpose of CS, a justice-involved individual includes a victim of crime and any person who has had interactions with a justice subsystem, including pretrial services; child welfare agencies (juvenile justice system); trial, family (juvenile justice), and appellate courts; prosecution and public defender offices; probation and parole agencies; and custodial institutions (jails, prisons, reformatories, halfway houses, etc.).

3.1. All meta-analyses included in the evidence base for practices must meet the certain minimum criteria (see below for details).

ELIGIBILITY			
Item	Yes	No	Notes
A. The meta-analytic study focuses on at least one eligible justice or justice-related problem.	<input type="checkbox"/>	<input type="checkbox"/>	
B. The meta-analytic study focuses on a problem involving a justice-involved population.	<input type="checkbox"/>	<input type="checkbox"/>	
C. The study includes specified eligibility criteria or a moderator analysis that targets one or more active components.	<input type="checkbox"/>	<input type="checkbox"/>	
D. The meta-analytic study calculates an effect size for the targeted component from at least two studies.	<input type="checkbox"/>	<input type="checkbox"/>	
<i>Inclusion Criteria:</i>			
Notes: At least one of A or B must be true for the study to be eligible. C and D must both be true for the practice to be eligible.			

OTHER POTENTIAL PRACTICES		
	Active Component	Number of Studies (k)
1.		
2.		
3.		
4.		
5.		

TARGET POPULATION^a The study contains inclusion criteria that target individuals with any of the following specific characteristics:			
Item	Yes	Item	Yes
Criminal behavior	<input type="checkbox"/>	Children exposed to violence	<input type="checkbox"/>
Delinquent behavior	<input type="checkbox"/>	Maltreated/neglected youth	<input type="checkbox"/>
Disruptive behavior	<input type="checkbox"/>	Victims of crimes	<input type="checkbox"/>
Aggressive behavior	<input type="checkbox"/>	Mentally ill youths/individuals	<input type="checkbox"/>
Relational/social aggression (bullying)	<input type="checkbox"/>	Homeless youths/individuals	<input type="checkbox"/>
Internalizing behaviors (depression, anxiety, suicide)	<input type="checkbox"/>		
Trauma- and stress-related disorders	<input type="checkbox"/>	Females	<input type="checkbox"/>
Co-occurring disorders (substance use and mental health)	<input type="checkbox"/>	Lesbian, gay, bisexual, transgender, and questioning (LGBTQ)	<input type="checkbox"/>
Cognitive disorders	<input type="checkbox"/>	College students	<input type="checkbox"/>
		Tribal members	<input type="checkbox"/>
Substance use	<input type="checkbox"/>	Parents	<input type="checkbox"/>
Alcohol use	<input type="checkbox"/>	Families	<input type="checkbox"/>
Tobacco use	<input type="checkbox"/>		
Cocaine use	<input type="checkbox"/>	Other	<input type="checkbox"/>
Marijuana use	<input type="checkbox"/>	Specify:	
Opioid use	<input type="checkbox"/>		
First-time offenders	<input type="checkbox"/>		
Moderate/medium-risk offenders	<input type="checkbox"/>		
Serious offenders	<input type="checkbox"/>		
Sex offenders	<input type="checkbox"/>		
Gang offenders/members	<input type="checkbox"/>		
Incarcerated offenders	<input type="checkbox"/>		
Parolees	<input type="checkbox"/>		
Probationers	<input type="checkbox"/>		

^aThe selected characteristic (or characteristics) *must* appear in the practice title. The title can, however, reflect a more specific subgroup of the characteristic. For example, if appropriate, depression can be used instead of internalizing.

TARGET SETTING ^a The study contains inclusion criteria that target a unique setting.			
Item	Yes	Item	Yes
Correctional (secure facility)	<input type="checkbox"/>	High-crime neighborhood (hot spot)	<input type="checkbox"/>
Residential (nonsecure facility)	<input type="checkbox"/>	Disorganized neighborhood	<input type="checkbox"/>
Court	<input type="checkbox"/>		
Community	<input type="checkbox"/>	Urban	<input type="checkbox"/>
Home	<input type="checkbox"/>	Suburban	<input type="checkbox"/>
School	<input type="checkbox"/>	Rural	<input type="checkbox"/>
Tribal (reservation)	<input type="checkbox"/>		
Workplace	<input type="checkbox"/>	Other	<input type="checkbox"/>
		Specify other:	

^aThe selected setting (or settings) *must* appear in the practice title.

TARGET AGE ^a The study contains inclusion criteria that target a specific age group.			
Item	Yes	Item	Yes
Juveniles	<input type="checkbox"/>	Adults	<input type="checkbox"/>
Both	<input type="checkbox"/>		

^aThe selected age group *must* appear in the practice title.

POTENTIAL PRACTICE TITLE(S) ^{a, b}				
	Active Component	Target Population	Setting	Age
1.				
2.				
3.				
E.g.	COGNITIVE-BEHAVIORAL THERAPY FOR DISRUPTIVE BEHAVIOR	JUVENILES		

^aAny meta-analysis that produces an equivalent practice title can be included in the same evidence review. A meta-analysis may produce one or more possible practices.

^bThe evidence base *cannot* contain two or more related meta-analyses (e.g., parent and sibling meta-analysis, two sibling meta-analyses, original and updated meta-analysis).

PRACTICE:

1. **Primary aim of the intervention.** The programs included in the meta-analysis must aim to address the goals identified in the initial screening stage.

Finding:

2. **Literature search.** The literature search that guided the inclusion of primary studies in the meta-analysis must include at least two sources and must provide evidence that unpublished literature was sought in the search.

Finding:

3. **Primary outcomes.** The meta-analysis must report on at least one eligible outcome related to crime, delinquency, overt problem behaviors (e.g., aggression, gang involvement, substance misuse), crime victimization, justice system practices or policies, or risk factors for crime and delinquency.

Finding:

4. **Control groups.** All studies included in the meta-analysis must include an appropriate control, or, in cases where both appropriate and inappropriate controls are included, the meta-analysis must analyze appropriate controls separately from inappropriate controls. (Please note the CS will focus exclusively on the appropriate control analyses).

Finding:

5. **Reporting of results.** The meta-analysis must report effect sizes that represent the magnitude of the treatment effect.

Finding:

6. **Combining effect sizes.** When an average effect size is reported for multiple studies, all effect sizes in the combination must address the same type of relationship (i.e., active and inactive comparisons* are not combined to calculate an average effect size).

*An inactive comparison is one in which the comparison group receives services that are considered inert (treatment as usual). In contrast, an active comparison is one in which the comparison group receives services that are expected to produce an outcome

Finding:	
-----------------	--

7. **Publication date.** At least 50 percent of the studies included in the meta-analysis must be published or otherwise available during or after 2000.

Finding:	
-----------------	--

8. **Age of samples.** Samples included in the meta-analysis must be restricted to either adults or juveniles, or mean effect sizes for adults and juveniles must be reported separately.

Finding:	
-----------------	--

Recommendation	Yes
REVIEW	<input type="checkbox"/>
DO NOT REVIEW	<input type="checkbox"/>
Notes:	

similar to that of the intervention under investigation (another intervention). Combining these different comparisons will likely produce blurred and biased effect-size estimates.