Research in Brief

ROBINA INSTITUTE
OF CRIMINAL LAW AND CRIMINAL JUSTICE
UNIVERSITY OF MINNESOTA

To write the *Research in Brief* series, researchers at the Robina Institute reviewed the best available research on topics of interest to criminal justice practitioners and policymakers. Researchers assessed the strength of the evidence in support of specific outcomes by evaluating the rigor of the research. To standardize evaluation researchers created and followed the Evidence Assessment Criteria and the Hierarchy of Study Design.

Evidence Assessment Criteria

Strong Evidence	Where **two or more high-quality studies** (i.e., high-quality meta-analyses, randomized controlled trial, or quasi-experimental) exist that, over time and/or geography, consistently show a direct relationship between the program/policy/practice and improvements in primary outcomes (i.e., public safety or correctional population).
Promising Evidence	Where at least **one high-quality study** (i.e., high-quality meta-analyses, randomized controlled trial, or quasi-experimental) exists, that shows a direct relationship between the program/policy/practice and improvements in primary outcomes (i.e., public safety or correctional population) and no substantial evidence of negative or null effects exists, or; Where there are **two or more studies of medium quality** (i.e., observational studies) that point in the same positive direction on primary outcomes (i.e., improvements in public safety or correctional population) and no substantial evidence of negative effects exists, or; Where there is strong evidence of **success in tackling intermediate outcomes** (e.g., reductions in anti-social behavior, improvements in pro-social behavior), and these outcomes have been shown to be linked to improvements in the primary outcomes.
Mixed Evidence	Where strong or promising evidence shows **varying outcomes for similar populations**, so that it is difficult to find consensus regarding effectiveness.
Insufficient Evidence	Where **no evaluation** has been found on the program/policy/practice or; Where some attempt has been made to evaluate the program/policy/practice, but this is of **unknown quality** or **below the standards of promising evidence**, such that it is difficult to identify impact.
Negative Evidence/ No-Effect	Where there is substantial evidence, from **one or more high** (i.e., high-quality meta-analyses, randomized controlled trial, or quasi-experimental) or **two or more medium quality studies** (i.e., observational), that the program/policy/practice has **negative impacts** or **no effect** on the primary outcomes, and no conflicting outcomes exist from other studies.

Adapted from Centre for Justice Innovation, 2016. *Problem-solving courts: An evidence review*. London, England.

Hierarchy of Study Design

High Quality Studies			Medium Quality Studies	Low Quality Studies
Systematic Reviews or Meta-Analysis	**Randomized Controlled Trials**	**Quasi-Experimental Studies**	**Observational Studies**	**Descriptive Studies**
Using strict inclusion and exclusion criteria, systematic reviews and meta-analyses combine results across studies to determine an overall effect. High quality meta-analyses generally exclude observational studies in order to protect against internal bias.	The simplest form of randomized controlled trial is known as the parallel group trial which randomizes eligible participants to two or more groups, treats according to assignment, and compares the groups with respect to outcomes of interest. Participants are allocated to groups using both randomization (allocation involves the play of chance) and concealment (ensures that the intervention that will be allocated cannot be known in advance). There are different types of randomized study designs, such as: • *Randomized cross-over trials:* Where all participants receive all the interventions; for example in a two arm cross-over trial, one group receives intervention A before intervention B, and the other group receive intervention B before intervention A. It is the sequence of interventions that is randomized. • *Cluster randomized trials:* A cluster randomized trial is a trial where clusters of people rather than single individuals are randomized to different interventions. For example, whole clinics or geographical locations may be randomized to receive particular interventions, rather than individuals.	The main distinction between randomized and quasi-experimental studies is the way in which participants are allocated to the intervention and control groups; quasi-experimental studies do not use random assignment to create the comparison groups. However, they can use other methods to mimic randomization. Examples: • *Non-randomized controlled studies:* Individuals are allocated to a concurrent comparison group using methods other than randomization and are matched based on demographics and other characteristics. The lack of concealed randomized allocation increases the risk of selection bias. • *Interrupted time series:* Interrupted time series designs are multiple observations over time that are 'interrupted,' usually by an intervention or treatment. With a strong method to mimic randomization, quasi-experimental studies can be of high-quality, but can also be of medium quality if there are non-major threats to internal validity.	A study in which natural variation in interventions or exposure among participants (i.e., not allocated by an investigator) is investigated to explore the effect of the interventions or exposure on health outcomes. Examples: • *Cohort study:* A defined group of participants is followed over time and comparison is made between those who did and did not receive an intervention. (The majority of criminological research falls into this category.) • *Case-control study:* Groups from the same population with (cases) and without (controls) a specific outcome of interest, are compared to evaluate the association between exposure to an intervention and the outcome. If adequate controls or matching techniques are not used, this type of study falls down to low quality.	Studies that are not comparative and that do not seek to establish a causal relationship between practice/policy/program and outcome. Descriptive studies are used to describe characteristics of a population or the intervention being studied. They do not answer questions about how/when/why the outcome occurred.

Adapted from *Systematic Reviews: Centre for Reviews and Dissemination Guidance for Undertaking Reviews in Health Care*, 2008

GLOSSARY

Primary outcomes:

- Reduction in criminal behavior or recidivism (e.g., arrest, supervision failure, conviction, re-incarceration, self-reported criminal involvement)
- Reduction in the size of the correctional population (for large-scale policies)
- Reduction in the rate of victimization or crime rate (for large-scale policies)

Intermediate outcomes:
These are outcomes that have been shown through research to affect recidivism, victimization, and the crime rate. A non-exhaustive list would include drug/alcohol use, employment, housing, and criminal thinking.

Unintended consequences or auxiliary benefits:
These are not considered in the evidence assessment but will be mentioned in the executive summary if substantive evidence can be found.

- Cost-benefit compared to status quo
- Effect on institutional legitimacy (e.g., justice involved person view process as fair)
- Effect on racial disparity
- Net-widening (i.e., increases in the number of people under criminal justice system)

Improvement:
Any statistically significant and substantively meaningful improvement in the intended outcomes.

Process Evaluation:
A process evaluation documents and explains the goals, key program elements, and operations (i.e., processes) of the program/policy/practice. It evaluates whether the project was implemented as intended, without necessarily tracking the effect on primary outcomes. A high-quality process evaluation generally involves some level of observation, interviews, and basic quantitative analysis (e.g., participant characteristics, changes in sanctions, program completion) (Karlstein, 2011).