

## What Makes a Research Study Reliable?

Thousands of studies have been conducted on topics related to the justice system, and their results often inform policies and practices. However, not all research studies are conducted with the same “methodological soundness,” or level of quality, which can impact the reliability of their findings. This brief helps consumers of research studies evaluate a study’s quality and understand some common terminology.

■  
Justice system agencies use not only research studies but also data that they have collected to drive policies and practices. For data to be valuable, it should be of high quality (i.e., accurate, complete, reliable, relevant, and timely) and ideally reflect a longer time frame.

### Research Standards

- **Was the study peer-reviewed?** Using subject matter experts to evaluate a study’s design and the researchers’ conclusions adds credibility to the findings. Peer-reviewed studies are often published in prestigious and other recognized journals.
- **Is there general agreement among experts on the study’s conclusions?** When there is disagreement, reviewing the critics’ qualifications and the validity of the criticism can help determine whether the input should be considered.
- **Is there general agreement among studies?** Studies with similar conclusions to other studies generally have increased validity. Studies that are outliers should be scrutinized with increased vigor.
- **Have the results been replicated?** Studies that are replicated in other settings and have similar conclusions are generally more valid.
- **What sample size was used?** Generally, the larger the study sample, the more robust the findings and the less likely they are a result of chance.
- **Were other variables considered?** Researchers should evaluate how variables other than those being examined may have impacted the study’s outcomes.
- **Do study results reflect causation or correlation?** With causation, also known as “cause and effect,” one event is the result of another. With correlation, two events are linked to a greater or lesser extent.
- **Was bias present?** Researchers should discuss possible biases—for example, if they knew who was in the experimental and control groups or if they had a personal interest in the study’s outcomes—and how these might have affected study results.
- **What limitations, if any, should be considered?** Quality studies should identify any flaws or limitations in the study design or in the findings.

peer-  
reviewed  
generalizable  
replicable  
transparent  
unbiased

Based on the answers to the previous and other questions, research gradients have been established. **Gold studies** have the greatest scientific vigor. However, few justice system studies are gold standard because of the difficulty in creating control groups due to equal protection rights and other factors preventing a pure experimental–control design. However, many studies fall in the **silver-standard** gradient.

### STANDARDS OF RESEARCH<sup>1</sup>

<b>Gold</b>	<b>Studies use a randomized controlled trial; demonstrate significant, sustained positive results; and have results that have been replicated at multiple sites.</b>
<b>Silver</b>	<b>Studies use a quasi-experimental research design, with appropriate statistical controls for the control group; demonstrate significant, sustained positive results; and have results that have been replicated at multiple sites.</b>
<b>Bronze</b>	<b>Findings show promise but require more rigorous empirical study. For example, studies may have used a matched comparison group without complete statistical controls, or results may need to be examined in greater detail to have full confidence in expected outcomes.</b>
<b>Iron</b>	<b>Studies consistently show conflicting or inconclusive findings (i.e., one study shows that something works while another study shows that it does not work).</b>
<b>Dirt</b>	<b>Studies use methods and criteria associated with gold and silver standards, except the findings consistently demonstrate that an intervention is not effective.</b>

### TYPES OF STUDIES

The following are various types of studies in order of the strength of their evidence, beginning with the most reliable:<sup>2</sup>

- A **meta-analysis** is a type of quantitative, formal study of multiple research studies that are assessed using a statistical method.
- A **systematic review** identifies, appraises, and summarizes all relevant research on a particular topic.
- A **randomized controlled trial (RCT)** is a single study where subjects are randomly assigned to an experimental or a control group. The experimental group receives the treatment or intervention, while the control group does not.
- A study conducted using a **quasi-experimental design** is similar to an RCT, but subjects are not randomly assigned to the experimental or control group.
- In a **cohort study**, researchers identify a topic that they wish to study and then follow research participants who share some characteristics, such as their location, age, or profession, over a period of time. A cohort study can be prospective (i.e., forward-looking) or retrospective (i.e., backward-looking).
- A **case control study** compares two groups—one with a condition/trait and one without it—and looks back to determine how frequently a risk factor was present in each group. This helps determine the factor’s relationship to the condition/trait.
- In a **case study**, information about a subject is gathered either through observation, interviews, or official records.

<sup>1</sup> National Institute of Corrections and Crime & Justice Institute. (2004). *Implementing evidence-based practice in community corrections: The principles of effective intervention*. <https://s3.amazonaws.com/static.nicic.gov/Library/019342.pdf>

<sup>2</sup> The George Washington University Himmelfarb Health Sciences Library. (n.d.). *Study design 101*. <https://guides.himmelfarb.gwu.edu/studydesign101>