

# Experimental Methods in Criminology

Rylan Simpson, Simon Fraser University

<https://doi.org/10.1093/acrefore/9780190264079.013.841>

**Published online:** 20 March 2024

## Summary

Experimental methods have been a hallmark of the scientific enterprise since its inception. Over time, experiments have become much more sophisticated, complex, and nuanced. Experiments have also become much more diverse, and their use within research settings has expanded from the physical sciences to the social sciences, including criminology.

Within criminology, experimental methods can manifest in the form of laboratory experiments, field experiments, and quasi-experiments, each of which present their own strengths and weaknesses. Experimental methods can also be applied in the context of between-subject and within-subject paradigms, both of which exhibit unique characteristics and implications. Experimental methods—as a *research method*—are unique in their ability to help establish causal relationships among variables. This article introduces the topic of experimental methods in criminology, with a specific focus on the subfield of policing.

**Keywords:** criminology, experimental criminology, experiments, field experiment, laboratory experiment, policing, quasi-experiment, randomization, social science research

**Subjects:** Policing, Research Methods

## Introduction

---

Experimental methods have been a hallmark of the scientific enterprise since its inception. For much of history, experiments have been celebrated as a symbol of science. As early as primary school, children are taught about experiments and their implications for understanding various phenomena. Most of this early education, at least as it relates to experiments, regards the physical sciences. Memories of combining vinegar and baking soda to make volcanoes erupt, or creating makeshift geysers by adding Mentos to Coca-Cola, remain salient among the minds of many people.

Over time, experiments have become much more complex, nuanced, and diverse. Over the life span, experiments also typically become much more technical and sophisticated, especially for those who pursue a career in research. Instead of combining cupboard ingredients to create homemade lava lamps, researchers may apply strategic changes to policing techniques to assess their effects on crime (Braga et al., 2019). The sophistication of experiments becomes particularly prominent when considered in the context of the social sciences, such as criminology, where researchers often examine concepts that can be difficult to measure and/or manipulate, including perceptions, attitudes, and behaviors.

For many people, experimentation is a salient part of life. Such salience can come in the form of active participation, whereby people participate in experimentation, as well as passive participation, whereby people are subjected to experimentation. A cursory search of research in the news highlights both active and passive engagement with experimentation. As government budgets tighten and calls for accountability strengthen, especially in the criminal justice context, public groups have called for greater evaluations of programs, strategies, and ideas, and many of the responses to these calls have involved experimental methods. From this perspective, experimental methods have become increasingly relevant in fields like criminology (Braga et al., 2014; Sampson, 2010; Weisburd, 2003).

In light of their methodological benefits, some scholars have gone so far as to suggest that experimental methods may be the “gold” standard of research methods (e.g., see Weisburd, 2003). This conclusion has been met with contention, however, and justifiably so (e.g., see Sampson, 2010). Each set of methods present their own strengths and weaknesses, and experimental methods are no exception. Although experiments can present many benefits for testing hypotheses and assessing causality, they, too, are not perfect, especially if not conducted well. Moreover, not every research question is amenable to every research method, experimental methods included. Identifying the implications of experimental methods as they relate to criminology forms part of the goal of this article.

This article begins by first introducing the relationship between experiments and causality and then describing the three different types of experiments: laboratory experiments, field experiments, and quasi-experiments. Next, the article reviews two different types of experimental paradigms: between-subject and within-subject paradigms. Within each aforementioned section, the article provides examples of policing studies that have employed the associated methods. Finally, the article concludes with a discussion of the implications of experimental methods for criminology, with specific attention to their prevalence, reception, and future.

Before proceeding, several caveats must be noted. First, this article is intended to provide an introductory overview of experimental methods rather than a comprehensive description of all elements related to experimental methods. Indeed, the topic of experimental methods in criminology could form full-length textbooks, which would provide much more space than an entry-length article as this one. Second, when describing examples of experiments, this article focuses largely on recent work as opposed to historical work, which has otherwise been referenced and described elsewhere. Third, the examples provided in this article derive from the subfield of policing (which previous research has observed houses most of the experiments published in leading criminology journals; Dezhnev et al., 2021). Nonetheless, this discussion could apply to other subfields of criminology where experimental methods are also used, including corrections, courts, and psychology and law.

## Experiments and Causality

---

Experiments can come in many different forms—from laboratory to field to quasi-experiments, with between- and within-subject paradigms. One element that all such forms of experiments generally share is a focus on establishing causality. Causality occurs when a change in one variable produces a direct change in another variable, independent of any other variables (Blomberg et al., 2013; Sampson et al., 2013). It can be denoted—most simply—as “ $X \rightarrow Y$ ,” whereby  $X$  is the independent variable and  $Y$  is the dependent variable. The independent variable is the variable that is changed or manipulated, whereas the dependent variable is the variable that such change or manipulation is tested upon.

Causality can only be assessed when a researcher achieves three criteria: empirical association, temporal sequence, and no spuriousness. Causality is very difficult to establish, especially in the social sciences, in large part because of its stringent requirement to demonstrate the latter two criteria: temporal sequence and no spuriousness.

Empirical association requires that two variables be related to each other in at least some capacity: If  $X$  caused  $Y$ , then  $X$  and  $Y$  must at minimum exhibit some relationship. For example, a change in policing technique must associate with a change in crime. If there is no relationship between the two variables, then there would be no causal effect to identify.

Temporal sequence requires unidirectional ordering in the relationship between two variables. To demonstrate that  $X$  caused  $Y$ ,  $X$  must come before  $Y$ . For example, a change in policing technique must occur before the change in crime if the policing technique caused the change in crime. This criterion generally requires multiple measurements of the dependent variable.

No spuriousness requires that the change in  $Y$  caused by  $X$  is only due to  $X$  and not some other unmeasured variable. Ruling out extraneous explanations can be challenging in the social sciences because the social world is a complex world. For example, the change in crime must only be caused by the change in policing technique and not changes in funding, resources, law, or any other unmeasured variable. If the research is not able to tightly control all elements of the research context, it is not possible to confidently rule out spuriousness. Experiments thus help to achieve this third—and arguably most difficult—criterion of causality via their careful control and effective use of either randomization, in the case of laboratory and field experiments, or advanced statistical techniques, in the case of quasi-experiments.

## Types of Experiments

---

### Laboratory Experiments

Laboratory experiments are arguably the foundation of experimental research (e.g., see Falk & Heckman, 2009). Most often conducted in physical laboratories, this type of experiment allows for meticulous control and careful testing as it relates to all kinds of phenomena. Researchers employing laboratory experiments begin with a clearly formulated hypothesis, which they then

test within the environment of their own creation. Laboratory experiments can be used to test hypotheses regarding perceptions, attitudes, and behaviors. They can be elaborate or simple, fast or slow, and expensive or inexpensive.

A signature feature of laboratory experiments is randomization. Randomization can come in two different forms: random sampling and random assignment. Whereas the former is not unique or limited to experiments (but when employed bolsters the strength of experiments), the latter is. Random sampling involves the selection of participants for the research: When the selection process is random, all members of a population should have an equal chance of being included in the sample. Random assignment involves the assignment of participants (once sampled) to groups, in the case of a between-subject design, or to the order of “treatment” (sometimes also referred to as an “intervention”), in the case of a within-subject design. In all cases, random assignment helps to ensure that no systematic differences that could otherwise affect the dependent variable exist between groups or among participants.

For example, in one line of laboratory experiments, Simpson (2017, 2020) assessed the effects of officer appearance on participants’ perceptions of police under the guise of a memory study. As part of his research, Simpson (2017, 2020) presented participants with carefully manipulated images of police officers, of different genders and racial backgrounds, in various aesthetic capacities (e.g., when wearing different forms of attire and accoutrements), in randomized order, and asked them to rate such officers along several dependent variables, such as aggressiveness, approachability, and friendliness. Not only did all participants experience exactly the same research procedures in the same physical space, as confirmed by the attending research assistant, but randomization and mild deception were both embedded within the study design to help rule out spuriousness.

In another line of experiments, researchers employed laboratory methods to assess officers’ decisions to use deadly force. In one experiment of this genre, Taylor (2020) strategically manipulated the content of dispatch transmissions provided to participating officers before assessing their decision to shoot (or not shoot) a suspect during a manipulated video scenario. In a similar experiment, Potts et al. (2022) used a virtual reality simulator to test the effects of dispatch priming on both sworn and nonsworn participants’ decisions to shoot (or not shoot) a suspect. Careful control and randomization in both experiments allowed for the researchers to isolate the effect of dispatch priming on participant behavior in a safe setting that mirrored what might be a real interaction in the field, but without the real consequences that could come from discharging a firearm.

Although laboratory experiments are frequently conducted in a physical laboratory, it is important to acknowledge that advancements in technology and web-based software have allowed for the development of online laboratories as well: Many experiments that arguably conform to laboratory standards are now conducted remotely. Popular experimental software, such as Inquisit, and survey tools, such as Qualtrics, allow for the manipulation and presentation of variables that facilitate the identification of causality. For example, these online platforms allow for randomization both between- and within-subjects in terms of group assignment as well as treatment exposure. These online platforms also allow for the precise measurement of

participant behavior, including as it relates to explicit actions (e.g., participants' selections) and passive actions (e.g., participants' time to selection). In recent years, many laboratory experiments have employed online frameworks. For example, Yesberg et al. (2021) experimentally tested the effects of firearm presence on public perceptions of police among a British sample, Simpson and Sandrin (2022) experimentally evaluated the impact of wearing personal protective equipment on public perceptions of police during the COVID-19 pandemic, and Sandrin et al. (2023) experimentally explored public perceptions of police canine units.

Laboratory experiments are most often celebrated for their internal validity and associated ability to identify causality (internal validity refers to the ability to isolate the effect of an independent variable on a dependent variable, whereas external validity refers to the ability to extrapolate the findings regarding that effect to other contexts/environments/groups). Indeed, the level of control embedded within this type of experiment is unparalleled. Researchers are able to control all elements of the participant's research experience in a laboratory experiment. In the online environment, this control can include everything from the font of the text to the size of the text to the order of the questions presented to participants. In the physical environment, this control can extend to include the color of the paint on the laboratory's walls to the height of the participant's chair to the size of the participant's computer monitor. When combined with randomization, researchers can effectively expel extraneous explanations for otherwise meaningful relationships via this type of experiment.

Nonetheless, laboratory experiments are not without their limitations. For example, their artificial nature can generate concerns about the external validity of findings. As some argue, the world is not a laboratory, even though much of the scholarly understanding about the world derives from the laboratory. The stakes of laboratory experiments are also arguably trivial: A participant's action in a laboratory, whether good or bad, does not generally have "real" consequences. Moreover, the participants who complete such experiments are sometimes inexperienced in the subject matter of which they are tested. This is particularly relevant in the criminal justice context, where laboratory experiments may test specific criminal justice phenomena among non-criminal justice-exposed participants. From this perspective, laboratory experiments' lack of realism can present genuine concerns.

The characteristics of participants who complete laboratory experiments may also be unique, insofar that a self-selection effect may exist among those who complete laboratory experiments—a concern that has been raised in the context of student samples (e.g., see Henrich et al., 2010). When compounded by criticisms of small sample sizes, these concerns warrant some consideration, although recent experimental research in the area of policing has found similar results among different sampling frames (e.g., see Simpson, 2021). Finally, there are concerns that laboratory experiments may be subject to the Hawthorne effect, whereby participants change their behavior as a function of being watched. All of these limitations weigh against the benefits of laboratory experiments and speak to the need for researchers to select their methods based on their specific research questions and the demands of such questions for research design.

## Field Experiments

Field experiments offer an interesting alternative to laboratory experiments. By definition, field experiments are conducted in the *field*, whereas laboratory experiments are conducted in the *laboratory*. The field facilitates greater external validity of findings as it moves the experimental paradigm out of the artificial environment and into the real environment. This type of experiment allows researchers to experimentally examine phenomena and test hypotheses as they occur in naturalistic settings. This type of experiment has been historically more common in criminology than its laboratory counterpart (Dezember et al., 2021).

Field experiments blend the benefits of control with the naturalism of the field. For example, rather than explore perceptions of police in a hypothetical context, as would be done in a traditional laboratory experiment, a field experiment would allow researchers to test perceptions as they occur among citizens in actual policing contexts. Field experiments are frequently conducted in collaboration with practitioner agencies, such as police departments, especially in criminology. This collaboration is often necessary given that the kinds of topics amenable to field experiments typically require third-party assistance in manipulating the variables of interest—whether it be police actions, court procedures, or prison programming.

Like with laboratory experiments, randomization is central to field experiments. Depending on the research question, random sampling may involve the selection of locations, such as blocks, beats, or cities. It could also involve the selection of citizens, officers, groups of officers, or community groups. On the other hand, random assignment involves the delivery of the treatment to the sampled unit. For example, whereas random sampling would identify which blocks to include in the field experiment, random assignment would determine which blocks actually receive the additional patrols by police. Random assignment could also determine which officers wear a new uniform design or participate in a new training program or not.

Randomization can become challenging when working with third-party agencies who may not wish for researchers to randomly sample their personnel or randomly assign the experimental manipulation. This can especially be the case if the experimental manipulation may delay the delivery of a treatment to all locations or all personnel until after the experiment is completed or if the experiment may involve a treatment that could negatively affect some locations or personnel. In these instances, agencies may wish to choose which units to include and assign to the treatment rather than allow the researcher to do so randomly.

For example, hot spots policing has received tremendous attention in the context of field experimentation (Braga et al., 2019). In one recent field experiment, Groff et al. (2015) randomly assigned hot spots to receive foot patrol, problem-oriented policing, offender-focused policing (i.e., experimental), or business-as-usual policing (i.e., control), and then compared violent crime in such hot spots over time. In another recent field experiment, Ariel et al. (2016) randomly assigned hot spots to receive foot patrol by uniformed (but unarmed) police community support officers (i.e., experimental) or not (i.e., control), and then compared calls for service and crime in such hot spots over time. Other genres of field experiments have assessed the effects of

procedural justice dialogue (Mazerolle et al., 2013), nonenforcement public–police contacts (Peyton et al., 2019), and police uniform alterations (Johnson et al., 2015) on public perceptions of police.

The benefit of field experimentation is the researcher’s ability to experimentally test phenomena as they occur in the real world. Such benefit is a requirement in the context of some research questions that otherwise could not be tested in a laboratory setting. As alluded to by the examples of hot spots policing, a researcher could not create and manipulate an entire neighborhood in the confines of a physical laboratory. Testing the effects of a policing technique of this kind would require a change in the physical environment in which it would actually occur. When paired with rigorous, statistical methods, field experiments can create laboratory-like conditions in the natural environment, which can then yield important findings for scholarship and practice.

The greatest strength of field experiments—the naturalistic environment in which they are conducted—can also be their greatest weakness. Unlike a laboratory experiment, where even the color of the paint on the walls can be controlled, the field setting presents many more barriers for researcher control. Such barriers can be environmental and/or organizational. Weather and natural events can occur unexpectedly. Organizational leaders can change without notice. Funding can evaporate quickly. Priorities can change rapidly. Competing reforms can create overlapping changes within the study environment that affect the dependent variable in unforeseen ways. Differential attrition, whereby participants from different groups withdraw from the study at unequal rates or for systematic reasons, can pose threats to randomization. Noncompliance with assigned behavior can also occur among field participants. All of these events can occur without the researcher’s control and can negatively impact the ability of the researcher to rule out spuriousness: a key criterion of causality.

## **Quasi-Experiments**

Unlike laboratory and field experiments, quasi-experiments (sometimes called “natural experiments”) do not involve random assignment. Rather than relying on randomization, researchers employing quasi-experiments craft their control and experimental groups by using other techniques, such as matching “participants” (which could include people, groups, places, etc.) based on their characteristics to establish equivalence between groups or assessing change within a unit before and after some kind of an event using statistical techniques. In this way, quasi-experiments attempt to isolate the specific effect of a given treatment on a tested outcome without the help of random assignment.

For example, in one recent quasi-experiment, Wheeler and Phillips (2018) evaluated the effects of using automated license plate readers at traffic roadblocks on crime, calls for service, and traffic accidents in a series of police-identified hot spots. As part of their design, the authors employed propensity score matching to match each “treated” hot spot (i.e., experimental) with an equivalent “untreated” location (i.e., control) before then comparing differences in the outcomes. In another quasi-experiment, Jonathan-Zamir and Weisburd (2013) examined the relative importance of police performance and procedural justice for police legitimacy in situations of

acute security threats. As part of their design, the authors compared survey responses from people living in a town affected by severe security threats (i.e., experimental) to survey responses from people living in other towns and cities not affected by such threats (i.e., control).

Given that quasi-experiments do not involve random assignment, this type of experiment exhibits weaker internal validity than laboratory and field experiments. It is much more difficult to ensure no spuriousness in a quasi-experiment because it is much more difficult to control all elements of the research context. Whereas laboratory experiments offer the greatest control, quasi-experiments offer the least control given the “natural” process by which they occur. Moreover, even though effective matching of participants, as one example, can bolster the methodological strength of quasi-experiments, it is very difficult in practice to create “identical” groups that only vary in their exposure to the treatment. This is in part because doing so requires knowledge and measurement of all possible variables that could confound an observed relationship, which is seldom possible.

With that being said, quasi-experiments offer a suitable alternative for exploring relationships among variables that cannot be manipulated (and hence randomized) because doing so would otherwise be impossible, unethical, or illegal. For example, researchers could not terminate half of the police department, force officers to engage in misconduct, or create a new criminal law in the name of research. But, if such changes were to occur naturally in a field setting, researchers could assess the effects of these changes on specified outcomes using a quasi-experiment. From this perspective, quasi-experiments generally exhibit greater ecological validity than laboratory or even field experiments.

## Experimental Paradigms

---

### Between-Subject Paradigms

Experiments can employ between-subject and/or within-subject paradigms. Between-subject paradigms are generally more synonymous with experimental criminology, in part because criminology has favored field experiments—which often employ between-subject paradigms—throughout much of its history. In a between-subject paradigm, participants are subdivided into at least two different groups: the experimental group and the control group. Participants assigned to the experimental group receive the treatment, whereas participants assigned to the control group do not, noting that “participants” could include any unit in question—from people to groups to places. Researchers then assess differences in the dependent variable induced by the independent variable by comparing scores for the dependent variable among the groups. For example, a researcher wishing to test the effects of a new police uniform design (i.e., independent variable) on public perceptions of officer approachability (i.e., dependent variable) could assign one half of the participants to rate an image of an officer wearing the new uniform design (i.e., experimental group) and the other half of the participants to rate an image of the *same* officer wearing the existing uniform design (i.e., control group), holding all else about the images constant. The researcher would then measure the difference in participants’ ratings of officer approachability by group.



Between-subject paradigms are methodologically the strongest when random sampling is used to generate the sample and random assignment is used to assign the participants to their groups. As suggested by Weisburd and Gill (2014), it is important to acknowledge the nuance among random assignment strategies. Random assignment is often conceptualized in terms of simple randomization, whereby every participant has an equal chance of being assigned to either the experimental or control group, regardless of their characteristics. This simple strategy exhibits the least risk with large samples, when it is more reasonable to assume that no systematic differences exist among participants. In between-subject paradigms with smaller samples, where such risk of differences among participants is higher—as can be the case in place-based experiments—blocked randomization can help to ensure equivalence among groups. As part of this strategy, researchers create pairs of similar participants and then administer the random assignment within each pair.

Like with different types of experiments, different experimental paradigms exhibit different strengths and weaknesses. In the context of strengths, between-subject paradigms generally require less time from participants given that participants must only complete “one half” of the study. Between-subject paradigms also present little risk of testing effects given that participants only provide a single measurement of the dependent variable.

In the context of weaknesses, between-subject paradigms generally require larger samples because each participant only completes “one half” of the study, which can make these paradigms costly and resource intensive. If random assignment fails, between-subject paradigms also present greater risk of spuriousness among findings. Contamination among groups can present challenges for between-subject paradigms as well if participants are alerted to differences in their experiences by group, as can experimenter bias if the researcher treats the groups differently as a function of their assignment (the use of blind and/or double-blind procedures, which hide assignment information from participants and/or researchers, can help to minimize these effects). Finally, differential attrition can present inferential issues if participants systematically withdraw from the experiment as a function of their group assignment.

### Within-Subject Paradigms

In contrast to between-subject paradigms, which assign participants to either the experimental or control group, within-subject paradigms expose all participants to the treatment and allow them to act as their own controls (“participants,” again, could include any unit in question—from people to groups to places). Researchers then assess differences in the dependent variable induced by the independent variable by comparing participants’ scores for the dependent variable from before and after the delivery of the treatment. To use the same example of a new police uniform design, a researcher could ask *all* participants to rate two images of the *same* officer: One image of the officer wearing the new uniform design (i.e., experimental) and one image of the officer wearing the existing uniform design (i.e., control), holding all else about the images constant. The researcher would then measure the difference in participants’ ratings of officer approachability by image.

Similar to between-subject paradigms, randomization is important in within-subject paradigms. These paradigms are methodologically the strongest when random sampling is used to generate the sample and random assignment is used to assign the order of the treatment to participants. Herein, though, lies the distinction between paradigms: Given that all participants in a within-subject design receive the treatment (i.e., there is no separate control group), the order by which participants provide their measurement on the dependent variable as it relates to the treatment and control should randomly vary to account for order effects (a process sometimes referred to as “counterbalancing”). In the police uniform design example, some participants would rate the experimental image first, whereas other participants would rate the control image first.

Within-subject paradigms exhibit many strengths. For example, they generally require fewer participants given that all participants are exposed to the treatment and the control, which can make these paradigms less expensive and resource intensive. Given that participants act as their own controls, there is also no concern about the potential for systematic differences among groups that could introduce spuriousness. For the same reason, contamination among groups, at least in the traditional sense, is not possible because there are no groups in a within-subject paradigm.

With that being said, testing effects can present concern in within-subject paradigms: Because participants provide a measurement of the dependent variable twice, once for the treatment and once for the control, learning can occur, which can affect their responses. If not effectively balanced, participants’ exposure to previous information may affect their interpretation of subsequent information, which can induce an inferential challenge. The potential for both of these negative effects makes randomization so important as randomization helps to minimize these risks. Within-subject paradigms also generally require more time from participants given that all participants must complete “both halves” of the study.

Before transitioning to a broader discussion of experimental methods in criminology, it is important to highlight that some experiments integrate both between- and within-subject paradigms into their research designs. Such combination can maximize the strengths and minimize the weaknesses associated with each paradigm. For example, Simpson and Sandrin (2022) employed a combination of between- and within-subject paradigms when evaluating the effects of personal protective equipment (PPE) on public perceptions of police in the context of PPE awareness. Participants were randomly assigned to read one of three different news articles (i.e., between-subject) before rating images of the *same* officer wearing versus not wearing different items of PPE (i.e., within-subject).

## Discussion

---

There has been much growth in experimental criminology in recent decades, especially within the subfield of policing (Braga et al., 2014; Dezhnev et al., 2021). Although experimental methods are still used less frequently in criminology than other methods, more experiments are being conducted now than before. Experiments in criminology are also becoming much more robust over time. This growth is promising: Experimental methods offer many benefits and can help

criminologists in their mission to better understand crime, criminal justice, and related phenomena. In addition to their technical benefits, experimental methods can also present important practical benefits and policy implications (Nagin & Weisburd, 2013; Petersilia, 2008). From this perspective, the findings from experiments have helped to shape policing, as well as other subfields of criminology, in important and meaningful ways (e.g., see Lum & Koper, 2017).

On that note, it is helpful to clarify here that this article has adopted a broader view of experimental criminology than some other related articles that have conceptualized experiments only in terms of field settings. For that reason, it is important to reaffirm that “randomized experiments” and “randomized control trials” can refer to laboratory experiments as well. These types of experiments present the same kinds of benefits as field experiments for assessing causality, with an added layer of researcher control.

In light of the benefits of experimental methods, some might ask why these methods have not become more commonplace in criminology. The answer to such question extends beyond the scope of this article, which sought to merely introduce the relevance of experimental methods for criminology and—especially—policing research. Nonetheless, a number of potential explanations may help to explain this discrepancy, five of which are briefly introduced in the subsequent paragraphs and all of which should continue to receive attention among the field.

A first potential explanation regards practicality. Experiments can be costly and time-intensive. They require technical knowledge and sometimes sophisticated software. In the context of field experiments, they generally require third-party cooperation. Such cooperation is not always easy to obtain and, if obtained, is not always easy to manage. It can be difficult for agencies to forego control of their own personnel or practices in the name of research. This can make experiments seem more daunting and less appetizing than other research methods, especially if in the pursuit of rapid publications.

A second potential explanation relates to criminology’s reliance on secondary data (Kleck et al., 2006). Although it is possible that secondary data may be experimental data, many of the data sets that criminologists frequently use to conduct their analyses do not contain experimental data. Without experimental data, it is not possible to conduct experimental analyses.

A third potential explanation regards the nature of research questions posed by criminologists. In many respects, criminology is still a young field. Much research within the field remains at the exploratory phase, and hence testing causal mechanisms has not yet become the focus of that work. Given the emphasis of experimental methods on causal testing (as opposed to theory development), this makes them less enticing as a method to some criminologists: Without adequate theory to test, some would argue that experiments are not yet necessary. Moreover, even when causality is identified, experiments do not always explain mechanisms or provide—what some might term—“black box” insight. Lastly, some questions of interest among criminologists do not lend themselves well to experimental manipulation for practical reasons, as outlined in the following point.

A fourth potential explanation relates to the challenges of assessing causality. As noted at the outset of this article, a defining feature of experiments is their ability to identify causality, which derives in large part from the ability of experiments to achieve the criteria of temporal sequence and no spuriousness. If researchers interested in a topic are unable to achieve these criteria, one would not expect them to use experimental methods. For example, many macro-level phenomena are not amenable to randomization or present ethical complications if randomized. This issue is particularly salient in the context of field experimentation, where randomization may require that a potentially beneficial treatment be withheld from some or a potentially harmful treatment be imposed on others. However, as Weisburd (2003) notes, these challenges should not rule out the possibility of randomization altogether.

A fifth potential explanation involves the pace by which experimental development may occur. Research generally moves slow, as does change in trends among research, and experiments are still a relatively new addition to criminology. Until recently, a “vital few” (Braga et al., 2014, p. 3) experimental criminologists have been responsible for sustaining experimental methods in criminology and training the next generation of experimental scholars. For experimental criminology to grow, networks must expand and exposure to experimental methods must increase. Indeed, Lum and Yang (2005) observed that students of researchers who had completed experiments were more likely to later conduct their own experiments, and Braga et al. (2014) argued that “the best way to increase the number of experimental criminology converts is to increase the number of graduate students who are involved in randomized experiments and trained in the classroom by seasoned experimentalists” (p. 22). On this front, it is worth noting that specific courses on experimental methods remain rare in criminology departments at both the undergraduate and graduate levels.

### Expanding the Use of Experimental Methods in Criminology

Experimental criminology remains rather insulated within the field of criminology. With that being said, several professional networks dedicated to experimental criminology have sought to both highlight the importance of experimental methods in criminology as well as advance their use, integration, and recognition among the field. For example, the Division of Experimental Criminology (DEC) was established by the American Society of Criminology in 2009. As per its description (verbatim), the Division of Experimental Criminology (2023)

seeks to advance the development of experimental scholarship within all subject areas, as well as promote the use of evidence-based crime policy . . . the DEC has worked to advance applied and field research methods, as well as to increase the visibility of experimental and non-experimental evidence. Today DEC remains the only Division focused on a particular set of methodological approaches, as opposed to a substantive area of research.

As part of their efforts, the DEC distributes several awards each year, including the Jerry Lee Lifetime Achievement Award for “lifetime achievement in the field of experimental criminology,” the Award for Outstanding Experimental Field Trial for “a single research project

or program that contributes significantly to criminological research and experimental science,” and the Student Paper Award for “an outstanding paper involving experimental field trials, or theoretical or policy questions in the area of experimental criminology.”

In addition to the DEC, the Academy of Experimental Criminology (AEC), which was founded in 1998, “recognizes scholars who have successfully led randomized controlled, field experiments in criminology.” Similar to the DEC, the AEC distributes several awards each year, including the Joan McCord Award for “distinguished experimental contributions to criminology and criminal justice” and the Outstanding Early Career Experimental Criminologist Award for “exceptional early career scholarship.” Another organized effort that promotes the use, application, and review of experimental methods is the Campbell Collaboration Crime and Justice Group (Farrington & Petrosino, 2001; Wilson et al., 2021). As noted on their website, this group’s mission is “to coordinate, facilitate and encourage the production, updating and accessibility of high quality systematic reviews,” with the intent to inform policy, reduce crime, and increase justice in society.

The emphasis on the promotion and application of experimental methods in criminology is not limited to organizational efforts. Such emphasis is also reflected in various publication outlets, including the *Journal of Experimental Criminology* (JOEX), which is the official journal of the DEC. The JOEX consistently ranks among premier journals within the field, with an Impact Factor of 3.0 in 2022. As per its aims and scope (verbatim), the *Journal of Experimental Criminology* (2023)

focuses on high quality experimental and quasi-experimental research in the advancement of criminological theory and/or the development of evidence based crime and justice policy. The journal is also committed to the advancement of the science of systematic reviews and experimental methods in criminology and criminal justice. The journal seeks empirical papers on experimental and quasi-experimental studies, systematic reviews on substantive criminological and criminal justice issues, and methodological papers on experimentation and systematic review. The journal encourages submissions from scholars in the broad array of scientific disciplines that are concerned with criminology as well as crime and justice problems.

Given its linkage with the DEC, the JOEX shares important connections and overlap with the DEC membership.

## Conclusion

---

As the term *research methods* implies, methods are critical to the pursuit of research. Each method offers its own benefits and weaknesses, and experimental methods—as one set of methods—are no exception. Useful for identifying causality, experiments come in many different forms, including laboratory experiments, field experiments, and quasi-experiments. Experimental paradigms can also be of the between-subject and within-subject variety. Despite their nuance, experiments generally share a similar goal of helping to isolate the effect of a given variable on a

tested outcome via careful control and rigorous design. As the field continues to grow, criminologists should continue to consider the use of experimental methods to test their research questions.

### Further Reading

Blomberg, T. G., Mestre, J., & Mann, K. (2013). Seeking causality in a world of contingency: Criminology, research, and public policy. *Criminology & Public Policy*, 12(4), 571–584.

Braga, A. A., Welsh, B. C., Papachristos, A. V., Schnell, C., & Grossman, L. (2014). The growth of randomized experiments in policing: The vital few and the salience of mentoring. *Journal of Experimental Criminology*, 10(1), 1–28.

Chalmers, I. (2003). Trying to do more good than harm in policy and practice: The role of rigorous, transparent, up-to-date evaluations. *The Annals of the American Academy of Political and Social Science*, 589(1), 22–40.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Houghton Mifflin.

Farrington, D. P., & Petrosino, A. (2001). The Campbell Collaboration Crime and Justice Group. *The Annals of the American Academy of Political and Social Science*, 578(1), 35–49.

Farrington, D. P., & Welsh, B. C. (2005). Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology*, 1(1), 9–38.

Lum, C., & Yang, S.-M. (2005). Why do evaluation researchers in crime and justice choose non-experimental methods? *Journal of Experimental Criminology*, 1(2), 191–213.

Nagin, D. S., & Weisburd, D. (2013). Evidence and public policy: The example of evaluation research in policing. *Criminology & Public Policy*, 12(4), 651–679.

Sampson, R. J. (2010). Gold standard myths: Observations on the experimental turn in quantitative criminology. *Journal of Quantitative Criminology*, 26(4), 489–500.

Sampson, R. J., Winship, C., & Knight, C. (2013). Translating causal claims: Principles and strategies for policy-relevant criminology. *Criminology & Public Policy*, 12(4), 587–616.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Shepherd, J. P. (2003). Explaining feast or famine in randomized field trials: Medical science and criminology compared. *Evaluation Review*, 27(3), 290–315.

Sherman, L. W. (2009). Evidence and liberty: The promise of experimental criminology. *Criminology & Criminal Justice*, 9(1), 5–28.

Sherman, L. W. (2010). An introduction to experimental criminology. In A. R. Piquero & D. Weisburd (Eds.), *Handbook of quantitative criminology* (pp. 399–436). Springer.

Sherman, L. W., Farrington, D. P., Welsh, B. C., & MacKenzie, D. L. (2002). *Evidence-based crime prevention*. Routledge.

Sherman, L. W., & Strang, H. (2004). Verdicts or inventions? Interpreting results from randomized controlled experiments in criminology. *American Behavioral Scientist*, 47(5), 575–607.

Weisburd, D. (2003). Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review*, 27(3), 336–354.

Weisburd, D., & Gill, C. (2014). Block randomized trials at places: Rethinking the limitations of small N experiments. *Journal of Quantitative Criminology*, 30(1), 97–112.

Weisburd, D., Lum, C. M., & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *The Annals of the American Academy of Political and Social Science*, 578(1), 50–70.

## References

Ariel, B., Weinborn, C., & Sherman, L. W. (2016). “Soft” policing at hot spots—Do police community support officers work? A randomized controlled trial. *Journal of Experimental Criminology*, 12(3), 277–317.

Blomberg, T. G., Mestre, J., & Mann, K. (2013). Seeking causality in a world of contingency: Criminology, research, and public policy. *Criminology & Public Policy*, 12(4), 571–584.

Braga, A. A., Turchan, B. S., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing and crime reduction: An update of an ongoing systematic review and meta-analysis. *Journal of Experimental Criminology*, 15(3), 289–311.

Braga, A. A., Welsh, B. C., Papachristos, A. V., Schnell, C., & Grossman, L. (2014). The growth of randomized experiments in policing: The vital few and the salience of mentoring. *Journal of Experimental Criminology*, 10(1), 1–28.

Dezember, A., Stoltz, M., Marmolejo, L., Kanewske, L. C., Feingold, K. D., Wire, S., Duhaime, L., & Maupin, C. (2021). The lack of experimental research in criminology—Evidence from *Criminology* and *Justice Quarterly*. *Journal of Experimental Criminology*, 17(4), 677–712.

Division of Experimental Criminology. (2023). *Official website of the Division of Experimental Criminology* <<http://www.expcrim.org/>>.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, 326(5952), 535–538.

Farrington, D. P., & Petrosino, A. (2001). The Campbell Collaboration Crime and Justice Group. *The Annals of the American Academy of Political and Social Science*, 578(1), 35–49.

Groff, E. R., Ratcliffe, J. H., Haberman, C. P., Sorg, E. T., Joyce, N. M., & Taylor, R. B. (2015). Does what police do at hot spots matter? The Philadelphia Policing Tactics Experiment. *Criminology*, 53(1), 23–53.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83.

- Johnson, R. R., Plecas, D., Anderson, S., & Dolan, H. (2015). No hat or tie required: Examining minor changes to the police uniform. *Journal of Police and Criminal Psychology*, 30(3), 158–165.
- Jonathan-Zamir, T., & Weisburd, D. (2013). The effects of security threats on antecedents of police legitimacy: Findings from a quasi-experiment in Israel. *Journal of Research in Crime and Delinquency*, 50(1), 3–32.
- Journal of Experimental Criminology. (2023). Official website of the Journal of Experimental Criminology <<https://www.springer.com/journal/11292>>.
- Kleck, G., Tark, J., & Bellows, J. J. (2006). What methods are most frequently used in research in criminology and criminal justice? *Journal of Criminal Justice*, 34(2), 147–152.
- Lum, C., & Koper, C. S. (2017). *Evidence-based policing: Translating research into practice*. Oxford University Press.
- Lum, C., & Yang, S.-M. (2005). Why do evaluation researchers in crime and justice choose non-experimental methods? *Journal of Experimental Criminology*, 1(2), 191–213.
- Mazerolle, L., Antrobus, E., Bennett, S., & Tyler, T. R. (2013). Shaping citizen perceptions of police legitimacy: A randomized field trial of procedural justice. *Criminology*, 15(1), 33–63.
- Nagin, D. S., & Weisburd, D. (2013). Evidence and public policy: The example of evaluation research in policing. *Criminology & Public Policy*, 12(4), 651–679.
- Petersilia, J. (2008). Influencing public policy: An embedded criminologist reflects on California prison reform. *Journal of Experimental Criminology*, 4(1), 335–356.
- Peyton, K., Sierra-Arevalo, M., & Rand, D. G. (2019). A field experiment on community policing and police legitimacy. *Proceedings of the National Academy of Sciences of the United States of America*, 116(40), 19894–19898.
- Potts, J., Hawken, A., Hillhouse, M., & Farabee, D. (2022). Virtual reality for law enforcement training: A demonstration and implication for dispatch priming. *Police Practice and Research*, 23(5), 623–632.
- Sampson, R. J. (2010). Gold standard myths: Observations on the experimental turn in quantitative criminology. *Journal of Quantitative Criminology*, 26(4), 489–500.
- Sampson, R. J., Winship, C., & Knight, C. (2013). Translating causal claims: Principles and strategies for policy-relevant criminology. *Criminology & Public Policy*, 12(4), 587–616.
- Sandrin, R., Simpson, R., & Gaub, J. E. (2023). An experimental examination of the perceptual paradox surrounding police canine units. *Journal of Experimental Criminology*, 19(4), 1021–1031.
- Simpson, R. (2017). The Police Officer Perception Project (POPP): An experimental evaluation of factors that impact perceptions of the police. *Journal of Experimental Criminology*, 13(3), 393–415.
- Simpson, R. (2020). Officer appearance and perceptions of police: Accoutrements as signals of intent. *Policing: A Journal of Policy and Practice*, 14(1), 243–257.
- Simpson, R. (2021). When police smile: A two sample test of the effects of facial expressions on perceptions of police. *Journal of Police and Criminal Psychology*, 36(2), 170–182.



- Simpson, R., & Sandrin, R. (2022). The use of personal protective equipment (PPE) by police during a public health crisis: An experimental test of public perception. *Journal of Experimental Criminology*, 18(2), 297–319.
- Taylor, P. L. (2020). Dispatch priming and the police decision to use deadly force. *Police Quarterly*, 23(3), 311–332.
- Weisburd, D. (2003). Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review*, 27(3), 336–354.
- Weisburd, D., & Gill, C. (2014). Block randomized trials at places: Rethinking the limitations of small N experiments. *Journal of Quantitative Criminology*, 30(1), 97–112.
- Wheeler, A. P., & Phillips, S. W. (2018). A quasi-experimental evaluation using roadblocks and automatic license plate readers to reduce crime in Buffalo, NY. *Security Journal*, 31(1), 190–207.
- Wilson, D. B., Mazerolle, L., & Neyroud, P. (2021). Campbell Collaboration systematic reviews and the *Journal of Experimental Criminology*: Reflections on the last 20 years. *Journal of Experimental Criminology*, 17(4), 539–544.
- Yesberg, J. A., Bradford, B., & Dawson, P. (2021). An experimental study of responses to armed police in Great Britain. *Journal of Experimental Criminology*, 17(1), 1–13.

### Related Articles

Qualitative Methods in International and Comparative Criminology

Historical Methods in Criminology