Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

CCSS
COPING WITH CRISES
IN COMPLEX SOCIO-ECONOMICS SYSTEMS
A Competence Center of ETH Zurich

Using Laboratory Experiments to Study Law and Crime

Christine Horne, Heiko Rauhut

CCSS Working Paper Series

CCSS-10-010

Using Laboratory Experiments to Study Law and Crime

Christine Horne, Heiko Rauhut

Abstract

The 19th and 20th centuries produced breakthroughs in physics, chemistry, and the biological sciences. Laboratory research played an important role in the rapid advances made in these fields. Laboratory research can also contribute progress in the social sciences and, in particular, to law and criminology. To make this argument, we begin by discussing what laboratory experiments can and cannot do. We then identify three issues in the criminological and legal literature: why violence is higher in the southern United States than in the North, the relation between the severity of punishment and crime, and the expressive effects of law. We describe the relevant data from laboratory experiments and discuss how these data complement those gained through other methods.

Keywords: Laboratory experiments, experimental methods, law, crime

Classifications:

URL: `http://web.sg.ethz.ch/wps/CCSS-10-010`

Notes and Comments: Status: Submitted

Using Laboratory Experiments to Study Law and Crime

Christine Horne[1] and Heiko Rauhut[2]

[1] Washington State University, Department of Sociology, Pullman, WA 99164, 204 Wilson-Short Hall, chorne@wsu.edu

[2] ETH Zurich, Chair of Sociology, in particular of Modeling and Simulation, CLU E6, Clausiusstrasse 50, 8092 Zurich, Switzerland, rauhut@gess.ethz.ch

July 20, 2010

Abstract

The 19th and 20th centuries produced breakthroughs in physics, chemistry, and the biological sciences. Laboratory research played an important role in the rapid advances made in these fields. Laboratory research can also contribute progress in the social sciences and, in particular, to law and criminology. To make this argument, we begin by discussing what laboratory experiments can and cannot do. We then identify three issues in the criminological and legal literature – why violence is higher in the southern United States than in the North, the relation between the severity of punishment and crime, and the expressive effects of law. We describe the relevant data from laboratory experiments and discuss how these data complement those gained through other methods.

Key Words: Laboratory experiments, experimental methods, law, crime

Acknowledgments: We would like to thank Ryan O. Murphy, Karl-Dieter Opp, Jim Short, and Fabian Winter for valuable comments and suggestions on the manuscript, and Dirk Helbing for organizing this scientific collaboration.

Christine Horne is an associate professor in the Department of Sociology at Washington State University. Her work develops theoretical understanding of norm enforcement, tests theoretical predictions using laboratory experiments, and applies theoretical insights to explain naturally occurring norms. She is editor (with Michael Lovaglia) of *Experiments in Criminology and Law* and author of *The Rewards of Punishment: A Relational Theory of Norm Enforcement*.

Heiko Rauhut is a Post-Doc researcher at the ETH Zurich, Chair of Sociology, in particular of Modeling and Simulation. His work focuses on the theoretical understanding of the emergence of social norms, normative conflict, criminal behavior and their interrelation with the wisdom of crowds. He conducted a number of laboratory experiments whose results were published in Rationality and Society, Social Science Research, Journal of Mathematical Psychology, Zeitschrift für Soziologie and elsewhere. He is currently working on a book entitled *Crime and Punishment from a Game Theoretical Perspective*.

The 19th and 20th centuries produced breakthroughs in physics, chemistry, and the biological sciences. Laboratory research played an important role in the rapid advances made in these fields. Can laboratory research also contribute to progress in the social sciences? Can it contribute to progress in law and criminology?

Open any social science methods textbook and you will see evaluations of the strengths and weaknesses of different methods. These evaluations typically include a claim that laboratory experiments are high in internal validity and low in external validity. This claim sometimes is followed by the conclusion that, in practice, laboratory experiments are of little use in the social sciences. But this conclusion is premature. Laboratory experiments can be used to develop and test not just theories about the physical world, but also about the *social* world. By relying on multiple methods – including laboratory experiments – we can increase our knowledge.

To make this argument, we begin by discussing what laboratory experiments can and cannot do. We then describe three issues in the criminological and legal literature – why violence is higher in the southern United States than in the North, the relation between the severity of punishment and crime, and the expressive effects of law. We describe the relevant data from laboratory experiments and discuss how these data complement those gained through other methods.

WHAT LABORATORY EXPERIMENTS CAN AND CANNOT DO

Historically, criminologists have most commonly relied on surveys, data from official records and other sources of quantitative data, along with ethnographic studies (for some exceptions see Steffensmeier and Terry 1975). Recently, field experiments have garnered attention because of their strengths in determining causality (see, for example Farrington and Welsh 2005 for a review of the last two decades; see also Welsh and Farrington 2001). An entire journal (the *Journal of Experimental Criminology*) is now devoted to reporting experimental

work. Despite this emphasis on experimental approaches, laboratory experiments remain rare.

In criminology we are typically interested in testing causal theories. Causal theories include two components – statements about causal relations and statements about causal mechanisms. A causal relation asserts that by varying one factor (X) we can affect another factor (Y). Mechanisms are the processes through which one variable influences the other (Hedström, 2005; Horne and Hechter, 2009). Laboratory experiments have unique strengths, not just in assessing causality (as field experiments do) but also in disentangling the mechanisms responsible for causal relations. Thus they are valuable for testing causal theories.

Laboratory experiments are less useful for establishing the substantive importance of particular causal mechanisms and relations in the field. In other words, they are good for testing causal theories, but not for studying properties of natural settings (Webster and Sell, 2007). Laboratory experiments cannot tell us what the correlation is between disenfranchisement and recidivism in the United States (Manza and Uggen, 2006). But they can provide insight regarding the effects of exclusion from group leadership on obedience to the group's rules. Laboratory experiments cannot tell us how important a factor collective efficacy is for controlling crime in Chicago (or Seattle or Phoenix or Los Angeles) relative to other factors (Sampson, Raudenbush, and Earls, 1997). But they can provide data showing that social support for sanctioning efforts affect levels of control in a group (Horne, 2009). Laboratory experiments cannot tell us how much guards abuse prisoners or why particular people at Guantanamo behaved as they did. But they can show us how prison-like power structures can lead normal people to behave in abusive ways (Zimbardo, 2008).

In sum, laboratory experiments cannot tell us how substantively important particular causal factors and mechanisms are for predicting patterns in the field. They cannot tell us how much of the variance a particular causal factor explains in a substantive setting outside the laboratory. But they can help us to understand the mechanisms that link a causal factor to an observed outcome. They help us to understand *why* we see the correlations we do. In other words, they increase our theoretical understanding. They increase (or decrease) our confidence

in a theory and our expectation that the theory will be upheld in situations in which it is meant to apply. In turn, when we understand the mechanisms at work we are in a better position to develop innovations for the field.

Laboratory experiments have particular characteristics that make them valuable for testing causal theory and for disentangling causal mechanisms. Below we discuss some of these characteristics and their associated opportunities and pitfalls.

RANDOM ASSIGNMENT

Laboratory and field experiments share a common characteristic – the use of random assignment to an experimental condition. Random assignment is a feature of experiments that increases our confidence in drawing conclusions about causality. Uggen argues that "the disciplinary field of 'criminology' is largely organized around a single dependent variable. We observe the effect – conditions of crime and noncrime – in an observational sample, and then make heroic statistical efforts to disentangle the myriad forces that give rise to is. . . . [I]t is much more sensible for a researcher to actively manipulate a cause and then to observe its effects – to do something and watch what happens. In doing so, researchers can compare the effects of the cause they subject to treatment (t) with the counterfactual case in which some other cause or a control condition is applied (c)" (Uggen, 2008: 182; see also Freedman, 1991: 292, 306 f; and Berk, 2004: 98, 102). Random assignment allows us to conclude that the effects we observe are due to the manipulation and not to some unobserved characteristic of individuals. This is because individual traits are randomly distributed across the experimental conditions (and the researcher can be confident that subjects were exposed to the intended treatment).

ARTIFICIALITY AND SIMPLICITY

While both laboratory and field experiments share a reliance on random assignment, laboratory experiments are unique in their artificiality. Theories are simplifications of reality (Hechter and Horne, 2009). Thus experiments that are designed to test theory will be simpler than a field setting. One of the challenges of testing theories in the field is that many things are covarying simultaneously. The world is complicated. It can be difficult therefore to study the particular processes identified by a theory. In survey studies we deal with this difficulty by statistically controlling for possibly relevant factors. By contrast, in the laboratory, because we create artificial settings, we can eliminate extraneous factors by design.

The fact that laboratory experiments control for factors extraneous to the theory by design rather than through statistics means that they do not include as many variables as we are used to seeing in survey analyses. When using survey data we try to include all of the potentially relevant variables in order to maximize the amount of variation in the outcome that we can explain. By contrast, in a laboratory experiment, we focus on only a few causal factors in order to explain why they have an effect – that is, in order to understand the mechanisms at work. Thus experiments will be as simple as the theories they are designed to test (see Willer and Walker 2007; Helbing and Yu 2010). If the empirical implications of the theory are supported or disconfirmed in the laboratory, then the theory has been appropriately tested. Such a test does not tell us about all of the factors that might affect an outcome, but rather increases our understanding of the mechanisms linking a small number of causal factors to that outcome.

The goal of understanding mechanisms might even lead us to create situations that do not exist naturally in the world (Webster and Kervin, 1971). For example, we know that neighborhoods with higher levels of mobility have higher levels of crime. We also know that neighborhoods with higher proportions of renters relative to home owners have higher rates of crime. Suppose that we identify two possible mechanisms that may be operating – financial investment in a neighborhood or social attachment. How would we know whether one or both of these mechanisms is operating? It would be difficult to disentangle these in the field (where

financial and social investments may covary and where possible indicators – stability and homeownership – may also covary). But in the laboratory we can manipulate financial investment and social attachment separately and observe the levels of deviance and/or informal control that emerge.

People often believe that experiments will be more useful if they look more "real" – if they have higher levels of mundane realism. But making a laboratory setting more realistic may actually reduce our ability to examine the theoretical factors and mechanisms of interest (Cook and Campbell, 1979; Zelditch, 1969; see also Shadish, Cook, and Campbell 2002). As we see in physics and chemistry, highly abstract laboratory studies provide evidence for theories that are then applied outside of the laboratory in numerous substantive contexts. In the social sciences, as well as the physical sciences, the theoretical knowledge gained in artificial settings can be applied in the field.

COLLEGE STUDENT SUBJECTS

A common feature of laboratory experiments is that they involve college student participants. College students obviously differ from gang members, drug dealers, and others in multiple ways. What can experiments with college students tell us about the behavior of criminals?

The issue here is whether the theoretical causal factor of interest interacts with the characteristics of the subject.[1] Suppose that we created an experiment to test Manza's and Uggen's (2006) argument that disenfranchisement of felons increases recidivism. We do this by creating groups in the laboratory in which subjects are either granted or denied opportunities to vote for leaders who make the group's rules. We could then look to see whether subjects were more likely to break the rules in conditions in which they could not vote for the group's leaders. What if we found no correlation? How would we interpret the result? We might decide that

college students differ from felons in theoretically relevant ways and that those differences interact with the causal factor of interest. For example, felons are socially stigmatized in ways that college students are not. Perhaps disenfranchisement has an effect for stigmatized people but not for others. If so, then we need to modify our theory to include stigma, and conduct a new experiment to test this possibility. We might also conduct an experiment with convicted felons and see if the experimental results with felons are consistent with patterns observed in the field. The result of these processes is that our knowledge grows.

In many instances, however, the characteristics of subjects do not interact with the theoretical causal factors of interest. Medical researchers use mice (despite their obvious differences from humans) because they are similar to human beings in theoretically relevant ways (Campbell and Campbell, 2006). They can therefore help scientists understand the mechanisms that contribute to diseases such as cancer. Similarly, experiments with college students can help us understand mechanisms that lead to deviant behavior. As long as college student subjects are similar to others in theoretically relevant ways, we would expect similar patterns of responses to the experimental manipulations. Subjects who are different from inner city residents, high-school drop-outs, or white collar criminals may engage in more or less of the behavior to be explained. Their responses to the experimental manipulations should reveal the same patterns, however. In other words, as long as the characteristics of the subject do not interact with the causal factor of interest, we might see higher or lower overall levels of the outcome behavior depending on the subject pool, but the effects of the experimental manipulations would be observed regardless of the subjects.

TRIVIALITY

One criticism of laboratory experiments is that they can only examine trivial behaviors that we do not care about – like stealing points from another subject or cheating (for example of some experimental studies of crime see Fetchenhauer, Simon, and Fetchenhauer, 2008; Kalkhoff

and Willer, 2008). We cannot have subjects in the laboratory engage in serious crimes like murder or rape. Nor can we impose severe punishments like execution or long-term prison sentences. Nor can we place subjects in life-threatening situations.

We can, however, study general mechanisms that may be relevant for explaining a range of crimes. For example, if one has a theory that says that low self-control leads to criminal behavior of all types (robbery, mugging, cheating, and so forth), then we ought to be able to test the effects of self-control on a range of behaviors that can ethically and practically be studied in the laboratory. Such tests would increase (or decrease) our confidence in the theory or suggest theoretical refinements. Laboratory evidence in combination with evidence regarding serious crimes in the field would provide even stronger support.

The implication is that if we have a theory that is specific to a particular crime – rape, for example – then we could not test that theory in the laboratory. That is, if we think that rape is theoretically distinct from other crimes, then lab experiments that use behaviors other than rape as indicators will not contribute to our understanding of rape.

While the utility of lab experiments is limited in such situations, they may still offer some possibilities. For example, methods known from survey methodology can be used to study underlying cognitive mechanisms for severe offences. The central idea is use experimental treatments to activate response biases that in turn allow for inferences regarding causal processes. In the context of rape, one could vary the order in which questions about rape myth acceptance and rape proclivity are posed. Research shows that if respondents are asked about their acceptance of rape myths first, the correlation between rape myth acceptance and rape proclivity are typically higher (Bohner et al, 1998; 2005). Such studies provide evidence that justifications matter in criminal decision-making and are consistent with neutralization theory (Sykes and Matza, 1957). Similarly, wording effects can be used to test arguments about perceptions of probabilities of becoming a victim of a crime (Krumpal et al. 2010).

Typically however, laboratory experiments are most useful for testing general, abstract theories. If we have a theory that applies to crime in general (for example, a theory about the role of power in motivating people to engage in behavior that harms others) then that theory could be tested in the laboratory. In other words, laboratory experiments will be more useful for testing abstract, general theories than concrete, specific theories.

GENERALIZABILITY

Generalizability refers to two things – the ability to generalize from the sample to the population from which it was drawn and the ability to generalize from the sample to other populations and settings (Lucas, 2003). The results of laboratory experiments that rely on convenience samples of college students cannot be generalized to the larger population, because those students are not drawn appropriately from such a population. In that sense, laboratory experiments have low external validity.

But, the *theories* that are tested using laboratory experiments can be applied to other settings. One of the characteristics of a theory is that it is general – it applies across a range of settings. Theories are not about specific individuals or unique situations. If the theory only applies in the laboratory and not elsewhere, then that suggests a problem with the theory, not with the laboratory. Similarly, a theory of crime tested in Chicago ought to be applicable to other American cities. If it is not, it suggests that the theory should be revised – not that Chicago is a bad place to test theories of crime.

If the theory is general, then it can be applied both in the laboratory and in the field. If we test a theory in one setting (like the laboratory) and our findings are consistent with it, then we have greater confidence in the theory. If these findings are consistent with an application elsewhere, we can have even greater confidence (Singleton and Straits 2010: 226).

It is possible that we could see consistent findings across settings just by coincidence. But, such a conclusion would require us to accept the existence of a lot of coincidence in the world (Guala 2002: 263). Most of the time, consistent results are not the result of coincidence. We can therefore identify three conditions for drawing internally and externally valid causal inferences from data collected in both the field and the laboratory: First, we have a theory that identifies a mechanism connecting a causal factor with an outcome. Second, we have evidence from the field that there is a correlation between the causal factor and the outcome. Third, we have evidence of a mechanism linking the causal factor and the outcome in the laboratory. If these conditions are met, it is reasonable to think that the mechanism that operates in the laboratory is also operating in the field.

SUMMARY

Experiments are particularly well-suited for disentangling causal relations and causal mechanisms. Other methods have other strengths. For example, surveys are very useful for describing the distribution of behaviors in a sample and, if the sample is well-constructed, for generalizing it to a larger population. Further, using sophisticated statistical techniques and survey methods, criminologists have been successful at identifying correlations and causal relations. Qualitative methods provide in-depth understanding of particular settings and can produce new insights and hypotheses about important social processes. Integrating these methods will allow us to take advantage of the strengths of each. Integration may involve reporting the results of multiple methods in a single paper. But it may also involve publishing papers based on laboratory experiments, papers based on surveys, and papers based on qualitative methods – with the authors of each building on the knowledge gained from the others.

To illustrate how data from laboratory experiments and other methods may together increase our understanding more than a single method alone, we focus on three issues in the

crime and law literatures – culture and violence, sentence severity and crime, and the expressive effects of law.

CULTURE AND VIOLENCE

In a series of studies, Cohen, Vandello, and Nisbett try to explain high rates of violence among white men in the American South compared with the North. They argue that Southerners adhere to culture-of-honor norms that lead them to view certain kinds of violence with approval and to respond aggressively to even minor slights. They support their theoretical argument with data from national surveys, analysis of major institutions, field experiments, and laboratory experiments.

Cohen and Nisbett (1994) analyzed survey data. They showed that people in the South were more approving of some kinds of violence (but not others) than those in the North. Southerners and Northerners responded in similar ways to questions about violence in the abstract. But, Southerners were much more supportive of the use of violence to defend themselves, their families, or their homes. They were more supportive of using violence to restore order (as in the case of police dealing with hoodlums and riots). And they tended to approve of violence that was a response to insults or that was used to discipline children.

In addition to analyzing survey data, Cohen (1996) analyzed laws and social policies in the South and the North. This analysis also suggested greater tolerance of violence in the South. It showed that laws in the southern states were more approving of gun ownership, self-defense, domestic violence, corporal punishment, and the death penalty than those in the northern states (Cohen, 1996).

These analyses showed a correlation between geographic region and acceptance of particular kinds of violence. Cohen and Nisbett (1997) went further to see if there was a causal

relationship between type of violence and attitudes. They conducted field experiments in which they manipulated the type of violence and measured the response of potential employers. The researchers sent letters inquiring about employment to companies in the South and the North. Some of the letters included information about an honor-related violence and some described involvement in a felony robbery. Whereas the felony letters were responded to in similar ways across regions, the honor letters were responded to more positively and sympathetically in the South than in the North. The results of the field experiments were consistent with the those of the survey and legal analyses. They provide evidence that a concern with honor drives southern attitudes about violence.

Finally, Cohen and his colleagues conducted a series of laboratory experiments that compared how subjects from the South and the North reacted to insults in the laboratory (Cohen et al. 1996). They argued that if Southerners' approval of violence really is driven by a concern with honor, then in comparison to Northerners, Southerners would view insults as damaging to their reputation, they would be more upset by the insult, and they would behave more aggressively following the insult.

In their first experiment, Cohen and his colleagues brought students from the South and the North into the experimental laboratory at the University of Michigan.[2] There were two conditions – one in which a confederate of the experimenter bumped into the subject and called him an asshole, and one in which no insult occurred. Two other confederates observed how subjects reacted to the insult. Subjects then read a vignette about a man (Larry) and woman (Jill) at a party. Twenty minutes after arriving at the party, Jill told Larry about a man who was making passes at her. A few minutes later, Larry saw the man try to kiss Jill. Subjects were asked to finish the story.

The results showed that Northerners were more likely to be amused by the insult and Southerners were more likely to get angry. There was a statistically significant interaction between region (south versus north) and insult (insult versus no insult) such that Southerners who were insulted became more angry than subjects in the other conditions. Further, 75% of

Southerners who were insulted finished the story by describing Larry injuring or threatening to injure the man – a higher proportion than among Southerners who were not insulted or among northerners.

In a second experiment, the researchers looked to see whether the insult caused different physiological reactions. To do this they measured subjects' cortisol levels before and after the insult. They found that cortisol levels increased 79% among insulted Southerners (compared with 42% for no insult Southerners) and increased 33% for insulted Northerners (compared with 39% for non-insulted Northerners). These results were statistically significant at the 5 % level.

The researchers also measured the extent to which the subjects believed that their status was lowered in the eyes of those who observed the insult. The results showed that Southerners who were insulted were more likely than Northerners to feel that their reputation had been damaged.

Finally, the researchers looked to see whether Southerners would behave in a more aggressive way after being insulted. Subjects were instructed to walk down a hall. A confederate walked from the other end of the hall toward the subject. Researchers measured how long the subject would walk toward the confederate before he stepped aside and gave way to him. The results showed that Southerners who had been insulted walked further before they deferred to the confederate (37 inches from the confederate for those who experienced an insult versus 108 inches for those who did not). Northerners behaved no differently whether insulted or not. Insulted Southerners also interacted with others in a more domineering way than those in other conditions.

The laboratory experiments provided evidence of mechanisms. They showed that, for Southerners, insults to honor led to increased anger, cortisol levels, and perceptions of reputation damage. Insults also increased aggressive behavior. One could criticize the experiment. One might argue, for example, that Southerners in Michigan are not like Southerners who stay in the South. Or, one might argue that behaving aggressively in the laboratory is not the same as

shooting someone in a bar. But, given that theory predicts the results found in the survey analyses, legal analyses, field experiments, and laboratory experiments, and that the results in all these empirical settings are consistent with the theory, it would be an unlikely coincidence if the mechanisms observed in the laboratory were simply a fluke.

The results of the laboratory experiments, field experiments, and survey analyses provide complementary evidence that support the theory. The survey analyses show that Southerners are more supportive of some kinds of violence than Northerners. The field experiments find evidence of a causal relation between type of crime and expressions of support for those who commit crime. The laboratory experiments find evidence of causal relations and mechanisms. They show that insults have different physiological effects on Southerners and Northerners, and have different effects on emotional reactions and aggressive behavior. This evidence of mechanisms and causal relations in the laboratory complements evidence from field experiments and evidence of self-reported attitudes provided by surveys. Taken together, these different sources of data give us greater confidence in the theory that Southerners adhere to a culture of honor that leads them to view violence differently than Northerners and therefore react differently to it. These studies do not tell us how important this cultural factor is in the field – how much it contributes to rates of violence relative to structural factors, for example. But they provide evidence that the meaning of violence is different in the South than in the North and that this variation in meaning affects emotions and behavior. Given these results and the patterns of violence we observe in the field, it would be reasonable to conclude that cultural mechanisms explain at least part of the variation in violence between the South and the North.

PUNISHMENT SEVERITY AND CRIME

Increasing the severity of punishment is, for some, an appealing approach to crime. But while policies such as “zero tolerance” and “three strikes and you’re out” are often politically popular (Austin et al., 1999; Cohen, 1985; Hudson, 2002; Zimring, 2001), the evidence that such

approaches are successful is far from clear.

Three prominent bodies of work explore the issue.[3] Two of them, ecological studies (for the pioneering work see Gibbs, 1968; Sellin, 1961; Tittle, 1969) and interrupted time series analyses (see, for example, Sherman, 1990) find little if any effect of policies that increase the severity of punishment. Work on specific crack-downs on crimes such as drug dealing show initial deterrent effects that decline over time, such that crime rates rebound (MacCoun and Reuter, 1998; Zimring et al., 2001). A third body of work uses surveys to ask respondents about their perceptions of the risks and costs of being caught and their own self-reported deviant and criminal behavior. These studies tend to find correlations between perceptions of punishment and behavior (Cook, 1980; Grasmick and Bryjak, 1980; Jensen, 1969; Lochner, 2007; Matsueda, 2006; Paternoster, 1987).

Why the discrepancy in the findings? One answer is that there is a gap between the severity of punishment demanded by policy changes and perceptions of punishment severity among citizens. Nagin (1998) argues that "for crime control policy to be effective it must alter these perceptions. Evidence on whether and how policy in current and prior time periods . . . affects sanction perceptions is fragmentary. Ecological and interrupted time-series studies have focused only on the relationship of policy to crime. In doing so these studies have treated the intervening policy-to-perceptions linkage. . . as a black box. . . While these studies generally find that policy has at least a temporary effect on crime and thereby somehow influences perceptions, the dearth of evidence on the policy-to-perceptions linkage is a major gap in knowledge. . . . [S]uch knowledge would be of great value in designing effective deterrence policies" (p. 7).

Why might there be a gap between policy and perceptions? One possible reason is that what is stated in policy may not actually be what is implemented. Some argue, for example, that when punishments are too severe, they are less likely to be imposed. There are numerous reasons for this including that because it is often more costly to impose harsher punishments the resources to do so may not be available, and that people may be less willing to impose severe punishments for relatively minor offenses.

In turn, when punishments are not imposed, potential offenders learn that risks are low and that they can commit crimes without much fear of being caught (see, for example, arguments that inexperienced offenders tend to overestimate the effectiveness of law enforcement, cf. Minor and Harry, 1982; Tittle, 1980). Of course, this argument requires that criminals actually learn from their experience. There is some support for this assumption in the criminological literature (see, for example, Horney and Marshall, 1992; Paternoster et al., 1985). Notably one study focuses on serious offenders – those incarcerated for felony offenses. In that study, those who had higher rates of arrest relative to criminal activity also reported higher risk perceptions. In other words, the study suggests that those who are arrested at a higher rate do update their perceptions about the risks associated with engaging in crime.

A second possible explanation for findings regarding the deterrent effect of severe punishment is that it is difficult to disentangle the direction of the causal arrow. The severity of punishment may reduce crime. But crime rates may also lead control agents to increase the severity of punishment. This means that in order to determine the extent to which punishment deters crime, one must take into account the effects of crime rates on sanctioning severity . This is often very difficult to do.

Rauhut's approach (2009; 2010) acknowledges both the endogeneity problem identified by criminologists and the possibility that policy may not directly translate into enforcement, and explicitly builds these possibilities into his theoretical argument. He relies on a mechanism that suggests that relatively simple principles of individual decision-making aggregate to produce counter-intuitive effects on the macro-level (Tsebelis 1989; 1990). The result of the mechanism that Rauhut examines is that the incentives of one set of actors are transmitted to the other.

More specifically, the underlying argument is based on the idea that potential offenders and law enforcement personnel have opposing interests. Citizens are more likely to commit a crime if they think they will not be caught, and less likely to commit a crime if they think they will be caught. Control personnel have the opposite incentive structure; they try to detect crime if they believe that people are actually committing crimes, but they do not make as much effort if

they think that people are not committing crimes. Thus, both citizens and enforcers try to foresee the other's choice.

When this argument is applied to predict the effects of punishment severity, it implies the following logic: If the severity of punishment is increased, enforcers anticipate less crime, and therefore make less effort to catch criminals. The criminals, in turn, anticipate lowered control rates and therefore sustain their criminal activities. This criminal activity creates incentives for control personnel to increase their inspection activities, which in turn deter criminal activities. Thus both potential criminals and control personnel react to what the other party does and what they expect the other party to do. Similar dynamics can be observed in many areas of social life. Consider penalty kicks in a soccer game (Chiappori et al., 2002; Moschini, 2004; Palacios-Huerta, 2003;). The goal-keeper prefers to dive left when the penalty kick goes left and prefers to go to the right side if the kick goes right. The kicker has the opposite preferences.

Is there reason to expect citizens and enforcers behave in this way? To answer this question, it may help to consider criminals and enforcers who pay no attention to the actions of the other. If a criminal commits crime without paying any attention to the likelihood of getting caught, she or he will end up in prison. If enforcers always work equally hard and try their best to catch criminals, crime rates will drop and the enforcers will be wasting their time looking for crime that does not exist. It is reasonable to assume that citizens pay some attention to what enforcers are doing, and that enforcers pay some attention to how much crime is occurring.

The result of this dynamic – in which individual citizens and enforcers make reasonable decisions given the information they have – is counterintuitive. Rauhut predicts that the intersecting incentives of citizens and enforcers level out such that increasing the severity of punishment reduces enforcement activities.[4] That is, increasing the severity of punishment aimed at potential criminals actually affects the behavior of enforcers. Similarly, increasing the rewards aimed at enforcement personnel actually affect the behavior of citizens.

This mechanism is difficult to test in the field. As described above, in field settings, it is easier for researchers to get information about policy than about actual enforcement efforts by control personnel. Relatedly, it is difficult to get data on offenders' reactions to actual enforcement levels. Further, it is difficult to observe the dynamic interplay of behaviors in the field. These likely data limitations in conjunction with the many other factors that may covary means that it is very difficult to track the feedback loops between criminal activity and control efforts. In addition, in the field it is difficult to solve the endogeneity problem: Do high crime rates affect the level of punishment severity or does punishment severity drive the level of crime? In the laboratory, these methodological concerns can be addressed.

Rauhut (2009) conducted one experiment that manipulated the severity of punishment and another one that manipulated the incentives for law enforcement to pursue criminals (Rauhut 2010). Because the laboratory setting allowed him to manipulate punishment severity, we know that the crime rate did not cause changes in punishment severity. In addition, Rauhut was able to measure the behavior of both offenders and control agents – something that is often difficult to do in natural settings – so that he could track the feedback loops between crime and control.

Both experiments had a similar structure. There were two citizens and two law enforcers. Participants were randomly assigned to be in either the citizen or the law enforcer position. The participants interacted with each other for 30 rounds. On each round, citizens had the opportunity to steal money from each other. Also on each round, law enforcers were able to investigate crime – that is, to determine whether theft occurred. Law enforcers earned money by catching criminal citizens and criminal citizens were subject to monetary fines.

What happened? In Experiment 1, increasing the severity of punishment led to lower rates of enforcement – when punishment was low, the average enforcement rate was 54%; when punishment was increased, the average enforcement rate dropped to 36%. This effect was significant at the 5 % level (Rauhut, 2009).

Increasing the severity of punishment also discouraged theft. When punishment severity

was shifted from low to high, theft rates dropped from 60 % to 40 %. But, this effect declined over time and subjects gradually started committing more crime. Note that the severity of punishment affected crime more than expected, suggesting that subjects were not good at anticipating that law enforcers were going to reduce their efforts. Further, theft increased slowly over time – suggesting also that subjects were not good at using their experience to update their perceptions of punishment. That is, they updated their perceptions based on experience, but did so slowly.

The results show that subjects varied in their reactions to punishment severity. Those who originally were the most deterred by strong punishment increased their theft behavior more than those who started out with high theft levels. One interpretation of this finding is that subjects varied in their expectations of being punished. Those whose expectations were the least accurate, updated them over time as they realized that they were not being punished. Gradually, they increased their stealing. Those who started out with more accurate perceptions had less change in their behavior over time. This finding is consistent with criminological arguments suggesting that people are naive about the consequences of law breaking (Minor and Harry, 1982; Paternoster et al., 1987; and Tittle, 1980).

Rauhut's findings are consistent with studies showing that severe punishments initially discourage crime, but that over time, crime rates rebound (see, for example, Ross, 1973; 1975). They are also consistent with research showing that crackdowns have residual deterrent effects that extend after the crackdown stops. If people are slow to update their perceptions based on their experience, then we would expect them to continue to refrain from crime and then slowly return to their usual criminal activities.

Experiment 2 produced complementary results and provided additional support for the proposed mechanism (Rauhut, 2010). It showed that as rewards for enforcement increased, crime rates dropped from 47% to 38%. Enforcement efforts also increased from 39% to 60%. As predicted, strengthening the incentives for law enforcers reduced crime.

These experiments provide evidence that incentives (punishments for criminals and rewards for enforcers) not only have direct effects on behavior; they also have indirect effects. Because criminals and control agents have opposing interests, each is sensitive to the incentives of the other. They change their behaviors in response to the behaviors of others. But the results also suggest that people are not as sensitive as the theory would predict – they are slow to update their perceptions.

Evidence from the field shows inconsistent results for the effect of punishment severity on crime. The laboratory experiments described here explore one possible set of mechanisms that may help to explain these different results. These experiments manipulate the severity of punishment and the incentives for law enforcers and accurately measure the behaviors of both citizens and enforcers. Because the causal factors are manipulated, and subjects are randomly assigned to an experimental condition, evidence of causal relations is clear. The laboratory experiments provide evidence of mechanisms relevant to arguments in the criminological literature – that experience with enforcement can lead people to update their expectations of enforcement and, in turn, to reduce their criminal behavior. They provide evidence of mechanisms that can help to explain the discrepancy between studies that compare policy changes and crime rates with those that ask respondents about their perceptions of enforcement and their own criminal activity.

THE EXPRESSIVE POWER OF LAW

Two prominent approaches explain why people obey the law. One (as in the experiments above) explains obedience in terms of sanctions. Law discourages certain behaviors by threatening punishment. A second focuses on legitimacy (Tyler, 1990). Richard McAdams and Janice Nadler propose a third mechanism through which law may have effects. They argue that the law has expressive effects. That is, it provides a focal point that helps people to coordinate their behavior.

Many situations contain elements of both conflict and coordination. For example, consider two drivers approaching an intersection. Each driver wants to go through the intersection first – thus they have conflicting interests. But each also prefers not to get into an accident – crashing is worse than letting the other person go first. They therefore have some desire to coordinate. McAdams and Nadler (2008) argue that if "there is *any outcome* the disputing parties jointly regard as the worse possible result" then the dispute involves some element of coordination (p. 870).

Research on coordination suggests that people will tend to coordinate around the solution that is most salient, the "focal point." What makes a particular solution salient? Nobel Laureate Thomas Schelling suggests that third party statements have an effect. Imagine two people who have lost each other in a large department store. Suppose that the store has posted signs that say "Lost parties should reunite at the fountain on the first floor." It is likely that people will go there. They do so not because they are afraid of punishment nor because the store is seen as a legitimate rule maker. Rather, the salience of the fountain as a recommended meeting place leads people to go there.

McAdams and Nadler argue that a legal rule can create a focal point. That is, in situations in which conflict exists, but in which people also have a good reason to coordinate their activities (to avoid a mutually undesirable outcome), they will coordinate around what is salient. Third party expressions can affect what is salient. Law is one source of third party messages. Thus law can have effects on behavior that are independent of sanctions or legitimacy.

One reason that social scientists have neglected the expressive effects of law is that these effects are difficult to study in the field (McAdams and Nadler, 2005; 2008). It is possible to measure people's perceptions of the likelihood of punishment, as scholars do in deterrence research. It is also possible to measure people's perceptions of the legitimacy of the law (see, for example, Tyler, 1990). But it is very difficult to disentangle the effects of salience from those of sanctions and legitimacy. Laboratory experiments are therefore a useful first step in testing the theory.

In their first experiment, McAdams and Nadler have participants play a "Hawk-Dove" game – basically a game of chicken. This game includes both elements of coordination and conflict. There are two players, each of whom can choose between one of two strategies – hawk (be aggressive) or dove (be a chicken). The highest payoff goes to the person who is aggressive (hawk) when the other person is a chicken (dove). The next highest payoff occurs when both players choose dove. The worst outcome for both occurs when they both decide to be aggressive.

In this situation there is conflict – each person would prefer to choose hawk when the other chooses dove. But, both of them view the hawk-hawk combination as the worst possible outcome. They both want to avoid this outcome and therefore have a reason to coordinate. How do they know whether to choose hawk or dove?

McAdams and Nadler (2005) compared several conditions: control, ostension (spinner), intentionality (leader), and legitimacy (leader selected by merit). In the control condition, no expressive information was provided. In the spinner condition, the experimenters used a spinner to highlight one outcome. In the leader condition, one of the participants was randomly selected to act as leader and to write a strategy on the board. Finally in the merit-based leader condition, participants took a quiz. Those with the highest quiz scores served as leaders who identified a strategy.

McAdams and Nadler looked to see whether the highlighted outcome (the focal strategy) affected subjects' choices. In the control condition, 51% of subjects chose one outcome, the other 49% chose the other. In the spinner condition, 65% of subjects conformed with the spinner suggestion. In the random leader condition, 74% conformed. And in the merit based leader 76% conformed with the leader's suggestion. Analyses showed that the experimental conditions affected behavior. In the spinner, leader, and merit-based leader conditions, the majority of subjects followed the suggestion. Further, subjects followed the suggestion more when it was made by a leader (as in the leader and merit-based leader conditions) than when it was made by a spinner.

This first experiment tested the idea that a third party statement creates a focal point – that is, makes a particular choice more salient – in a mixed motive situation. McAdams and Nadler (2008) then conducted two additional experiments to test the argument that law is one kind of third party statement that creates a focal point, thereby increasing coordination. Instead of relying on a spinner or a leader to suggest behavior, these experiments manipulated legal rules.

In one of the law experiments, participants were asked to imagine that they were in a dispute with another person about ownership of a cat. In the No Law condition, the law prioritized the claim of neither party. In the Law condition, subjects were told that, based on an old court decision, the law favored the other party's claim. Subjects had to decide whether they would give in to the other person or whether they would insist on their own claim. The situation created an incentive structure like that in the hawk-dove game. The results showed that people were more likely to insist on pursuing their own claim when there was no relevant law. In the no law condition, the rate of insisting was 44%; in the law condition the rate of insisting was only 27%. Even though the law was not associated with any punishment, it still affected people's choices.

McAdams and Nadler then created a second experiment that involved a contract dispute. Again they found that mere exposure to a legal rule changed behavior. In the No Law condition 90% of the participants insisted on a remedy that benefitted them, while in the Law condition only 78% did so.

These two sets of results are consistent with the argument that law creates a focal point – guiding decisions to people about what they should choose in situations that involve coordination and conflict. Of course, these experiments do not tell us how important the focal point mechanism is in the field. McAdams and Nadler acknowledge that sanctions and legitimacy are likely to have substantively large effects. But they also argue that because compliance with the law is such an important issue, we want to fully understand the mechanisms that contribute to compliance – not just those that have the biggest, most obvious effects.

The experimental results provide reason to think that the focal point mechanism operates as theorized. Does this mechanism help to explain what we observe in the field? Ginsburg and McAdams (2004) apply the theory in the context of international law to explain why nations frequently comply with decisions of international bodies such as the International Court of Justice even when those institutions have little power to impose sanctions. Ginsburg and McAdams argue that court decisions provide "focal points" that clarify expectations. These clarifications influence behavior. A long-standing territorial conflict between Denmark and Norway over Greenland, for example, was resolved simply by the Permanent Court of International Justice rendering a decision in favor of Denmark. In response to the decision, Norway immediately withdrew its claim.

Ginsburg and McAdams (2004) review the docket of the International Court of Justice and find support for the theory. In cases in which the structure of incentives was like that in a Hawk-Dove Game – that is, there was some element of both conflict and coordination – they found that the International Court of Justice tended to be successful. It was less successful, however, in situations in which parties had, in effect, already chosen Hawk-Hawk strategies. Thus it was less effective in disputes involving armed conflict. But, as the theory suggests, it affected parties' behavior in situations that had some element of coordination.

In this series of studies, McAdams and his colleagues propose a theory, test the theory using several laboratory experiments, and then apply the theory outside the laboratory. These studies taken together provide stronger evidence of the expressive power of law than any one study alone. Further, the experiments provide evidence that would be very difficult to obtain in the field.

DISCUSSION

Although in many ways human society is increasingly cooperative and civilized, crime persists. Relying on official crime statistics, surveys, observational data, and field experiments, much of the research in criminology and law provides evidence of the correlates of crime.

However, we still have much to learn about the causal mechanisms responsible for these correlations (Sampson 2000). Laboratory experiments have strengths in disentangling causal relations and mechanisms that can provide a useful complement to more traditional methods. Laboratory experiments, for example, can help to illuminate mechanisms at the micro-level and their aggregation to macro- level patterns of behavior.

How might we move forward in taking advantage of the strengths of laboratory experiments to increase our understanding of law and crime? Criminologists have identified key correlates of crime. When we have evidence from the field establishing the correlations between causal factors and outcomes, laboratory experiments could be used to disentangle the mechanisms responsible for those correlations (as Rauhut, and Cohen and his colleagues did in their experiments). Laboratory experiments might also be used to provide an initial test of a new theory. Once there is some supportive evidence from the laboratory, the theory could then be applied in the field (the approach of McAdams and his colleagues.

While laboratory experiments have high internal validity and are very good for disentangling causal relations and mechanisms, they tend to have lower external validity. In particular, laboratory results typically do not provide data regarding the size of effects that we will observe in the field, or the relative substantive importance of different causal factors. Other methods are more useful for addressing those issues. Thus laboratory experiments and other methods have complementary strengths. Adding laboratory experiments to our methodological repertoire has the potential to substantially increase our understanding of law and crime.

REFERENCES

Austin, James, John Clark, Patricia Hardyman, and D. Alan Henry. 1999. The impact of "three strikes and you're out." Punishment and Society 1(2):131-162.

Berk, Richard A. 2004. Regression Analysis: A Constructive Critique. Thousand Oaks, CA: Sage Publications.

Bohner, Gerd, Marc-Andre Reinhard, Stefanie Rutz, Sabine Sturm, Dagmar Effler. 1998. Rape myths and neutralizing conditions. Evidence for a causal impact of anti-victims attitudes on men's self-reported likelihood of raping. European Journal of Social Psychology 28:256-268.

Bohner, Gerd, Christopher Jarvis, Friederike Eyssel, and Frank Siebler. 2005. The causal impact of rape myth acceptance on men's rape proclivity: Comparing sexually coercive and noncoercive men. European Journal of Social Psychology 35(6):819-828.

Campbell, T. Colin, and Thomas M. Campbell. 2004. The China Study. Dallas, TX: BenBella Books.

Cohen, Dov. 1996. Law, social policy, and violence: The impact of regional cultures. Journal of Personality and Social Psychology 70:961-978.

Cohen, Dov, and Richard E. Nisbett. 1994. Self-protection and the culture of honor: Explaining southern violence. Personality and Social Psychology Bulletin 20:51-567.

Cohen, Dov, and Richard E. Nisbett. 1997. Field experiments examining the culture of honor: The role of institutions in perpetuating norms. Personality and Social Psychology Bulletin 23(11):1188-1199.

Cohen, Dov, Richard E. Nisbett, Brian F. Bowdie, and Norbert Schwarz.1996. Insult, aggression,

and the southern culture of honor: An "experimental ethnography." Interpersonal Relations and Group Processes 70(5): 945-960.

Cohen, Stanley. 1985. Visions of Social Control: Crime, Punishment, and Classification. New York: Polity Press.

Cook, Philip J. 1980. Research in criminal deterrence: Laying the groundwork for the second decade. Crime and Justice: An Annual Review of Research, 2.

Cook, Thomas D., and Donald D. Campbell. 1979. Quasi-experimentation: Design and analysis Issues for Field Settings. Boston: Houghton-Mifflin.

Doob, Anthony N., and Cheryl Marie Webster. 2003. Sentence severity and crime: Accepting the null hypothesis. Crime and Justice: A Review of Research 30:143-195.

David P. Farrington, and Brandon C. Welsh. 2005. A half century of randomized experiments on crime and justice. Crime and Justice 34: 55-132.

Fetchenhauer, Detlef, Josef Simon, and Felix Fetchenhauer. 2008. Gottfredson and Hirschi in the lab: An experimental test of the general theory of crime. In Experiments in Criminology and Law, eds. Christine Horne and Michael J. Lovaglia, Lanham, MD: Rowman and Littlefield.

Freedman, David A. 1991. Statistical models and shoe leather. Sociological Methodology 21: 291-313.

Gibbs, Jack P. 1968. Crime, punishment and deterrence. Social Science Quarterly 58:15-28.

Ginsburg, Tom, and Richard H. McAdams. 2004. Adjudicating in anarchy: An expressive theory of international dispute resolution. William and Mary Law Review 45:1229-1339.

Grasmick, Harold, and George J. Bryjak. 1980. The deterrent effect of perceived severity of punishment. Social Forces 59:471-491.

Guala, Francesco. 2002. On the scope of experiments in economics: Comments on Siakantaris. Cambridge Journal of Economics 26:261-267.

Hechter, Michael and Christine Horne, eds. 2009. Theories of Social Order. Stanford, CA: Stanford Social Sciences.

Helbing, Dirk and Yu, Wenjian. 2010. The future of social experimenting. Proceedings of the National Academy of Sciences USA. 107(12):5265-5266.

Hudson, Barbara. 2002. Punishment and control. In The Oxford Handbook of Criminology, eds M. Maguire, R. Morgan and R. Reiner, Oxford: Oxford University Press.

Horne, Christine. 2009. The Rewards of Punishment: A Relational Theory of Norm Enforcement. Stanford, CA: Stanford University Press.

Jensen, Gary. 1969. Crime doesn't pay: Correlates of a shared misunderstanding. Social Problems 17: 189-201.

Kalkhoff, William and Robb Willer. 2008. Deterring deviance: Rationality and self-control. In Experiments in Criminology and Law, edited by Christine Horne and Michael Lovaglia, Lanham, MD: Rowman and Littlefield.

Krumpal, Ivar, Heiko Rauhut, Dorothea Böhr, and Elias Naumann. In Press. The framing of risks and the communication of subjective probabilities for victimizations. Quality and Quantity.

Lochner, Lance. 2007. Individual perceptions of the criminal justice system. American Economic Review 97(1):444-460.

Lovaglia, Michael J., Jeffrey W. Lucas, Jeffrey A. Houser, Shane R. Thye, Barry Markovsky. 1998. Status Processes and Mental Ability Test Scores. American Journal of Sociology 104(1): 195-228.

MacCoun, Robert, and Peter Reuter. 1998. Drug control. In The Handbook of Crime and Punishment, ed. Michael Tonry, New York: Oxford University Press.

Maguire, Mike. 2002. Crime statistics. The "data explosion" and its implications. In The Oxford Handbook of Criminology, eds, M. Maguire, R. Morgan, and R. Reiner. Oxford: Oxford University Press.

Matsueda, Ross L., Derek A. Kreager, and David Huizinga. 2006. Deterring delinquents: A rational choice model of theft and violence. American Sociological Review 71:95-122.

Manza, Jeff, and Christopher Uggen. 2006. Locked out: Felon Disenfranchisement and American democracy. New York: Oxford University Press.

McAdams, Richard H., and Janet Nadler. 2005. Testing the focal point theory of legal compliance: The effect of third-party expression in an experimental hawk-dove game. Journal of Empirical Legal Studies 2:87-123.

McAdams, Richard H., and Janet Nadler. 2008. Coordinating in the shadow of the law: Two contextualized tests of the focal point theory of legal compliance. Law and Society Review 42:865-898.

Nagin, Daniel S. 1998. Criminal deterrence research at the outset of the twenty-first century. Crime and justice: A review of research 23:1-42.

Paternoster, Raymond. 1987. The deterrent effect of the perceived certainty and severity of punishment: A review of the evidence and issues. Justice Quarterly 42:173-217.

Rauhut, Heiko, and Ivar Krumpal. 2008. Rechtfertigungen und sexuelle Gewalt. Eine experimentelle Studie. Soziale Probleme 19(2):219-237.

Rauhut, Heiko, Ivar Krumpal, and Mandy Beuer. 2008. Rechtfertigungen und Bagatelldelikte: Ein experimenteller Test. In: Peter Kriwy and Christiane Gross: Klein aber fein! Quantitative empirische Sozialforschung mit kleinen Fallzahlen. VS Verlag für Sozialwissenschaften, pp. 373-396.

Rauhut, Heiko. 2009. Higher punishment, less control? Experimental evidence on the inspection game. Rationality and Society 21(3):359-392.

Rauhut, Heiko. 2010. Stronger incentives for control reduce crime: A lab experiment on paradoxical effects of incentives and a game theoretical explanation. Mimeo ETH Zurich.

Rauhut, Heiko, and Marcel Junker. 2009. Punishment deters crime because humans are bounded in their strategic decision-making. Journal of Artificial Social Systems and Societies (JASSS) 12(3), No. 1.

Sampson, Robert J. 2000. Whither the Sociological Study of Crime. Annual Review of Sociology 26:711-14.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. Neighborhoods and violent crime: A multilevel study of collective efficacy. Science 277:918-24.

Sellin, Thorsten. 1961. Capital punishment. Federal Probation 25: 3-10.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston, MA: Houghton Mifflin.

Sherman, Lawrence. W. 1990. Police crackdowns: Initial and residual deterrence. Crime and

Justice: A Review of Research 12:1-48.

Singleton, Royce A., Jr., and Bruce C. Straits. 2010. Approaches to Social Research, 5th edition. New York: Oxford University Press.

Steffensmeier, Darrell J., and Robert M. Terry. 1975. Examining Deviance Experimentally. Port Washington, NY: Alfred Publishing.

Sykes, Gresham M. and D. Matza.1957. Techniques of neutralization: A theory of delinquency. American Sociological Review 22:664-670.

Tierney, John. 1996. Criminology. Theory and Context. Upper Saddle River, NJ: Prentice Hall.

Tittle, Charles R. 1969. Crime rates and legal sanctions. Social Problems 16:409-423.

Tittle, Charles. 1980. Labeling and crime. An empirical investigation. In: Walter R. Gove, (ed.): The labeling of deviance : Evaluating a perspective. Sage Publications. Beverly Hills, New York.

Tsebelis, George. 1989. The abuse of probability in political analysis: The Robinson Crusoe fallacy. American Political Science Review 1:77-91.

Tsebelis, George. 1990. Penalty has no impact on crime. A game theoretic analysis. Rationality and Society 2:255-286.

Tyler, Tom R. 1990. Why People Obey the Law. New Haven, CT: Yale University Press.

Uggen, Christopher. 2008. Thinking experimentally. In Experiments in Criminology and Law, eds. Christine Horne and Michael J. Lovaglia. Lanham, MD: Rowman and Littlefield..

Webster, Murray Jr., and John B. Kervin. 1971. Artificiality in Experimental Sociology. Canadian

Review of Sociology and Anthropology 8(4): 263-272.

Webster, Murray. Jr., and Jane Sell, eds. 2007. Laboratory Experiments in the Social Sciences. Amsterdam: Elsevier.

Welsh, Brandon C., and David P. Farrington. 2001. Toward an evidence-based approach to preventing crime. Annals of the American Academy of Political and Social Science 2001:158-173.

Willer, David, and Henry A. Walker. 2007. Building Experiments: Testing Social Theory. Stanford, CA: Stanford Social Sciences.

Zelditch, Morris., Jr. 1969. Can you really study an army in the laboratory? In A Sociological Reader on Complex Organizations, edited by Amitai Etzioni, Austin, TX: Holt, Rinehart, and Winston.

Zimbardo, Philip. 2008. The Lucifer Effect: Understanding How Good People Turn Evil. New York: Random House.

Zimring, Franklin E., Gordon Hawkins, and Sam Kamin. 2001. Punishment and Democracy: Three Strikes and You're Out in California. New York: Oxford University Press.

ENDNOTES

1. A similar argument can be made regarding the experimental setting. The question is whether the characteristics of the experimental setting interact with the theoretical causal factors.

2. The comparison of Southerner and Northerners does not involve a true experimental manipulation because whether the subject was a Southerner or a Northerner was not manipulated. Similarly, researchers cannot manipulate the race or gender of the subject – though they can manipulate theoretically relevant aspects of race, gender, and so forth (for an example in the context of race see Lovaglia et al. 1998).

3. For reviews of deterrence research see Doob and Webster (2003); Maguire (2002), Nagin (1989), and Tierney (1996).

4. The underlying mechanism is the game theoretical concept of ``mixed Nash equilibria", which requires rational agents to outsmart their opponents. This concept assumes that citizens choose the probability of committing a crime at the indifference point of the inspectors and inspectors choose the probability of inspection at the indifference point of the criminals (see Rauhut and Junker 2009; Tsebelis1989; 1990). This implies that more severe punishment reduces control rather than crime.