

Cambridge Books Online

<http://ebooks.cambridge.org/>



Cambridge Handbook of Experimental Political Science

Edited by James N. Druckman, Donald P. Green, James H. Kuklinski, Arthur Lupia

Book DOI: <http://dx.doi.org/10.1017/CBO9780511921452>

Online ISBN: 9780511921452

Hardback ISBN: 9780521192125

Paperback ISBN: 9780521174558

Chapter

5 - Economics versus Psychology Experiments pp. 58-70

Chapter DOI: <http://dx.doi.org/10.1017/CBO9780511921452.005>

Cambridge University Press

Economics versus Psychology Experiments

Stylization, Incentives, and Deception

Eric S. Dickson

In this chapter, I follow other authors (e.g., Kagel and Roth 1995; McDermott 2002; Camerer 2003; Morton and Williams 2010) in focusing on a few key dimensions of difference between experiments in the economic and psychological traditions.

Section 1 considers the level of *stylization* typical in economics and psychology experimentation. Although research in the political psychology tradition tends to place an emphasis on the descriptive realism of laboratory scenarios, work in experimental economics tends to proceed within a purposefully abstract, “context free” environment.

Section 2 considers the kinds of *incentives* offered to subjects by experimentalists from these two schools of thought. Experimental economists generally offer subjects *monetary incentives* that depend on subjects’ choices in the laboratory – and, in game-theoretic experiments, the choices of other subjects as well. In contrast, psychology research tends not to offer inducements that are conditional on subjects’ actions, instead giving subjects fixed cash payments or fixed amounts of course credit.

Section 3 considers the use of *deception*. The psychological school tends to see deception as a useful tool in experimentation, and, at times, a necessary one. In contrast, the economic school by and large considers deception to be taboo.

These basic differences in research style highlight the historical divide between psychological and economic – alternatively, behavioral and rational choice – scholarship in political science. Over the years, scholars have tended to peer across this divide with more mistrust than understanding, and intellectual interchange between the different schools has been lamentably limited in scope. However, the difference in approaches between psychologists and economists reflects more than the sociology of their respective traditions; many of the norms characteristic of each field have evolved in response to the specific nature of theory and of inquiry within the separate disciplines.

To say that each school of experimentation has categorical strengths and weaknesses would perhaps be too strong a claim. Rather, in this chapter, I argue that the advantages

and disadvantages associated with specific design choices may play out differently, depending on the nature of the research question being posed, the theory being tested, and even the results that are ultimately obtained.

In this chapter, I organize my discussion around the logic of inference in economics- and psychology-style experiments. Going down this path leads me to several conclusions that may at first seem counterintuitive. For example, I will argue that stylized, economics-style experimentation can sometimes be particularly valuable in the study of essentially psychological research questions. Contrary to the way our discipline has traditionally been organized around separate schools of methodological practice, strategy and psychology are inextricably bound together in virtually all political phenomena that we desire to understand. The multifaceted nature of our objects of study, along with the varying strengths and weaknesses of different research methods in attacking different problems, highlight the advantages of methodological pluralism in building an intellectually cumulative literature in experimental political science.

1. Stylized versus Contextually Rich Experimental Scenarios

A first salient dimension of difference between economics and psychology experiments is rooted in the basic nature of the experimental scenarios presented to subjects. With some exceptions, economics experiments tend to be carried out in a highly stylized environment in which the scenarios presented to subjects are purposefully abstract, whereas experiments in psychology tend to evoke more contextually rich settings. Because the economic style of experimentation is likely to be more foreign to many readers, this discussion begins by describing some arguments that have been given in support of stylization in laboratory experiments.

Logic of Stylization

Research in the economic style tends to frame experimental scenarios in an abstract

rather than in a naturalistic manner. The roles assumed by subjects, and the alternatives that subjects face, are generally described using neutral terminology with a minimum of moral or emotional connotations; experimental instructions are written in a technocratic style. For example, in their landmark study of punishment in games of public goods provision, Fehr and Gächter (2000) employ an experimental frame using strictly neutral language, never once mentioning the word “punishment” or other potentially charged terms such as “fairness” or “revenge.” In a similar way, Levine and Palfrey (2007) use the labels X and Y, rather than terms like “vote” and “abstain,” in their experimental study on voter turnout; the cost of voting is translated into a “Y bonus” accruing only to individuals who choose Y, that is, do not vote. In their study of deliberation, Dickson, Hafer, and Landa (2008) model individual decisions to communicate in a stylized environment; the “arguments” exchanged during deliberation are represented using simple single-digit numbers.

The abstract experimental tasks associated with this form of stylization are used in part because of a desire to maintain experimental control. Researchers in this tradition generally believe that the use of normatively charged terms such as punishment, fairness, or revenge may evoke reactions in subjects whose source the analyst cannot fathom and that the analyst cannot properly measure. Experimental economists would generally argue that such loss of control would limit the generalizability, and thus the usefulness, of their findings in the laboratory.

According to this argument, the descriptively appealing complexity of highly contextual experiments comes with strings attached when it comes to inference. Suppose that a particular effect is measured in a contextually rich setting. More or less by definition, contextually rich settings contain many features that could potentially claim subjects’ attention or influence subjects’ behavior or cognition. Given this, how could we know *which* feature of the setting – or which combination of features – led to the effect that we observed?

In contrast, it is argued that a similar effect measured in a stylized setting may have wider lessons to teach. One argument for this claim can be explicated through the use of two examples. First, consider the Fehr and Gächter (2000) experiment, which demonstrated that many experimental subjects are willing to undertake costly punishment of counterparts who fail to make adequate contributions to a public good, even under conditions where such punishment is costly and no benefit from punishment can accrue to the punisher. Because this result was obtained in such an abstract choice environment, which did not directly prime subjects to think in terms of punishment or fairness, the result seems unlikely to be merely an artifact of some abstruse detail of the experimental frame presented to subjects. A more natural interpretation of the study's findings is that a willingness to punish the violation of norms is a basic feature of human nature that comes to be expressed even in novel settings in which subjects lack experience or obvious referents. As such, the use of an abstract, stylized environment in the study arguably strengthens rather than weakens the inferences we make from its result.

Second, Dickson et al. (2008) demonstrate that many subjects "overspeak" compared to a benchmark equilibrium prediction; that is, subjects often choose to exchange arguments during the course of deliberation even when they are more likely to alienate listeners than persuade them. This finding suggests that deliberation may unfold in a manner more compatible with the deliberative democratic ideal of a "free exchange of arguments" than a fully strategic model would be likely to predict. In their study, stylization has at least two distinct advantages. The use of a stylized, game-theoretic environment allowed for the definition of a rational choice benchmark in the first place – without which overspeaking could not have been defined or identified. And the finding that individuals overspeak, even in a stylized environment without obvious normative referents, underscores the behavioral robustness of individual willingness to exchange arguments with others.

Such arguments in favor of stylization have, in fact, even been employed from time to time within social psychology itself. The minimal group experimental paradigm (Tajfel et al. 1971; Tajfel and Turner 1986) demonstrated that social identities can motivate individual behavior even when those social identities were somewhat laughable constructs artificially induced within a stylized setting, for instance, dividing subjects based on their tendency to overcount or undercount dots on a screen or their preference for paintings by one abstract painter (Klee) over another (Kandinsky). The finding that even these social identities could affect behavior helped establish social identity theory and motivate a vast field of research.

Limits of Stylization

The first and perhaps most obvious point is that certain research questions – particularly certain research questions in political psychology – cannot reasonably be posed both in stylized and in contextually rich settings. Just to take one clear-cut example, Brader (2005) studies the effects of music within political advertisements on voters' propensities to turn out, on voters' seeking additional political information, and on other dependent variables. It would obviously make little sense to attempt to translate such a study into a highly stylized setting because the psychological mechanisms Brader explores are so deeply rooted in the contextual details of his experimental protocol.

Many other research questions, however, could potentially lend themselves to exploration either in stylized or in highly contextual contexts. In considering the advantages and disadvantages of stylization in such cases, a natural question to ask is whether experimental results obtained using both methods tend to lead to similar conclusions.

For at least some research questions, the evidence suggests that stylization may lead to conclusions that are misleading or at least incomplete. A classic example comes from the psychology literature on the Wason selection task. In Wason's (1968) original study, subjects were given a number of cards, each of

which had a number on one side and a letter on the other, and a rule that had to be tested, namely, that every card with a vowel on one side has an even number on the other side. Given a selection of cards labeled E, K, 4, and 7, subjects were required to answer *which* cards must be turned over in order to test the rule. In this study, only a small fraction of subjects gave the correct answer (E and 7); especially few noted that the rule could be falsified by turning over the 7 and finding a vowel, whereas others included 4 in their answer in an apparent search for information confirming the rule. This finding is often taken as clear evidence for a confirmatory bias in hypothesis testing.

The Wason (1968) selection task became a popular paradigm in the aftermath of the original study, and parallel versions have been carried out in many different settings. Interestingly, subjects' performance at the task appears to be highly variable, depending on the context in which the task is presented. In another well-known study, Griggs and Cox (1982) present subjects with a selection task logically equivalent to Wason's, but rather than using abstract letters and numbers as labels, the task is framed as a search for violators of a social norm: underage drinking. In this study, most subjects are readily able to answer correctly that people who are drinking and people who are known to be underage are the ones whose age or behavior needs to be examined when searching for instances of underage drinking.

Results such as these suggest that subjects may sometimes think about problems quite differently, depending on the frame in which the problem is presented, an intuition that seems natural to scholars with a background in psychology. At the same time, such results by no means imply that stylized studies yield different results from highly contextual ones more generally. To take framing effects themselves as an example, parallel literatures within economics and psychology suggest that frames can affect choice behavior in similar ways both in stylized and highly contextual environments.

As of now, there is nothing like a general theory that would give experimentalists

guidance as to when stylization might pose greater problems for external validity. Many scholars find that stylization can be beneficial, given their research questions – because of a perceived higher degree of experimental control, because stylization can sometimes allow for a clearer definition of theoretical benchmarks than might be the case in a highly contextual environment, or because stylized environments can sometimes pose a “tough test” for measuring behavioral or psychological phenomena, as in the Fehr and Gächter (2000) and Dickson et al. (2008) studies. At the same time, a literature consisting wholly of such studies would be widely met with justifiable skepticism about external validity. At least for many research areas within political science, the best progress is likely to be made most quickly when research in both traditions is carried out – and when scholars communicate about their findings across traditional dividing lines. When research using different techniques tends to point in the same direction, we can have more confidence in the results than we could have if only one research method had been employed. When research using different techniques instead points in different directions, the details of these discrepancies may prove invaluable in provoking new theoretical explanations for the phenomenon at hand, as scholars attempt to understand the discrepancies' origins.

2. Use of Monetary Incentives

In most economics experiments, subjects receive cash payments that depend on their own choices in the laboratory and, in the case of game-theoretic experiments, on the choices of other people. In contrast, subjects who take part in political psychology experiments are generally compensated in a way that does not depend on the choices they make, typically either a fixed cash payment or a fixed amount of course credit. What motivates experimentalists from these two traditions to take different approaches to motivating subjects?

The most obvious point to make is that many research studies in political psychology

are not well suited to the use of monetary incentives because the relevant quantities of interest cannot be monetized in a reasonable way. For example, in a framing study by Druckman and Nelson (2003), subjects report their attitudes on political issues after exposure to stimuli in the form of newspaper articles. Clearly, in studies with a dependent variable, offering subjects financial incentives to report one opinion as opposed to another would be of no help whatsoever in studying framing effects or the formation of public opinion.

Of course, the same is not true of *all* research questions of interest to political experimentalists, political psychologists included. As such, experimenters sometimes have a real choice to make in deciding whether to motivate subjects with monetary incentives. In considering the implications of this choice, it is useful to review some of the varied purposes for which monetary incentives have been used in experiments.

Monetary Incentives as a Means of Rewarding Accuracy or Reducing Noise

One potential use for monetary incentives in experiments is to reward accuracy. Experimentalists want to ensure that subjects actually pay attention and properly engage the tasks that they are meant to perform. In settings where a “right answer” is both definable and, at least in principle, achievable by the subject – a setting very unlike the Druckman and Nelson (2003) article cited previously – financial inducements can help fulfill this role. For example, in a survey experiment on political knowledge, Prior and Lupia (2008) find that monetary rewards motivate subjects to respond more accurately and to take more time considering their responses. This result suggests that financial inducements can sometimes help elicit more accurate measures of knowledge and reduce levels of noise in survey responses.

A natural, and related, setting for the use of such methods in political experiments involves the study of political communication. Scholars want to understand how indi-

viduals learn from the political communications to which they are exposed and whether citizens are actually able to learn what they need to in order to make reasoned choices (Lupia and McCubbins 1998). In pursuit of these objectives, a number of scholars have devised stylized experimental settings in which subjects receive messages whose informational value can be objectively weighed using Bayes’ rule in the context of a signaling game equilibrium (e.g., Lupia and McCubbins 1998; Dickson in press). Subjects then receive monetary rewards that depend on the degree of fit between their own posterior beliefs and the “correct” beliefs implied by Bayesian rationality in equilibrium.

Monetary Incentives as a Means of Controlling for Preferences

Many experiments in political economy focus on the effects of institutions in shaping individual behavior. Such experiments are typically organized as tests of predictions from game-theoretic models. Of course, actors’ preferences over different possible outcomes are primitive elements of such models. As such, to expose a game-theoretic model to an experimental test, it must be that there is some means of inducing subjects to share the preferences of actors in the theoretical model. In economics experiments, this is done through the use of monetary incentives for subjects.

It is instructive to highlight the difference between this approach and typical research methods in the psychological tradition. In political psychology experiments, direct inquiry into the nature of individual motivations, preferences, and opinions is often the goal. In contrast, for the purposes of testing a game-theoretic model, economics experiments generally prefer to control for individual motivations by manipulating them exogenously, to the extent that this is possible. By controlling for preferences using monetary incentives, experimental economists attempt to focus on testing other aspects of their theoretical models, such as whether actors make choices that

are consistent with a model's equilibrium predictions, or the extent to which actors' cognitive skills enable them to make the optimal choices predicted by theory.

Monetary Incentives as a Means of Measuring Social Preferences

Finally, it might also be noted that the use of monetary incentives can be beneficial for the study of subjects' intrinsic motivations. Consider, for example, the Fehr and Gächter (2000) study cited previously. Subjects interacted within a stylized environment, making public goods contributions decisions and choosing whether to punish others based on their behavior. In the experiment, both kinds of decisions were associated with monetary incentives; a decision to punish another subject, for example, came at a (monetary) cost to the punisher. That individuals are willing to engage in punishment, even when this has a monetary cost and when no future monetary benefit can possibly accrue, strengthens our sense of how strong subjects' intrinsic motivations to punish may be. Certainly, this finding is more telling than would be a parallel result from an analogous experiment in which subjects' decisions were hypothetical and did not bear any personal material cost for punishing others. In principle, this methodology can potentially allow us to measure the strength of this intrinsic motivation by varying the scale of the monetary incentives. Thus, studies such as Fehr and Gächter can allow us to learn about individuals' intrinsic motivations by observing deviations from game-theoretic predictions about how completely (monetarily) self-interested actors would behave.

Other studies have taken a similar approach, allowing for inquiry into traditionally psychological topics within the context of game-theoretic experiments. A prominent example is Chen and Li (2009), who translate the study of social identities into a lab environment where subjects play games for monetary incentives, thereby offering a novel tool for measuring the strength of identities and the effects of identities on social preferences.

Does the Scale of Monetary Incentives Matter?

If an experimentalist decides that motivating subjects with monetary incentives is appropriate for his or her study, one basic question of implementation involves the appropriate *scale* for monetary incentives. It is not unusual for experimental economics labs to have informal norms that subjects' expected earnings should not fall below some minimum rate of compensation; the maintenance of a willing subject pool requires that "customers" be reasonably happy overall with their experiences in the lab. Morton and Williams (2010) summarize existing norms by estimating that payments are typically structured to average around 50 to 100 percent above the minimum wage for the time spent in the lab. Such considerations aside, resource constraints give experimentalists a natural incentive to minimize the scale of payoffs in order to maximize the amount of data that can be selected – as long as the payments that subjects receive are sufficient to motivate them in the necessary way.

A recent voting game study by Bassi, Morton, and Williams (2007) suggests that the scale of financial incentives can affect experimental results. In their study, the inducements offered to subjects varied across three treatments, involving a flat fee only, a scale typical of many experimental economics studies, and a larger scale offering subjects twice as much. The fit between subjects' behavior and game-theoretic predictions became monotonically stronger as incentives increased; suggestively, this pattern was found to be most prominent for the most cognitively challenging tasks faced by subjects. These results suggest that, at least in some settings, higher rates of payment to subjects can increase subjects' level of attention to the experiment in a way that may affect behavior, a result consistent with intuitions derived from Prior and Lupia (2008), as well as related studies in economics (e.g., Camerer and Hogarth 1999).

Gneezy and Rustichini (2000) carried out a study on IQ test performance that

communicates a compatible message. Their experiment varied financial incentives for correct answers across four distinct treatments. They found performance to be identical in the two treatments offering the least incentives for performance (one of which simply involved a flat show-up fee) and identical in the two treatments offering the highest incentives, but performance in the higher incentive treatments exceeded that in the lower incentive treatments. This finding, along with Prior and Lupia (2008) and Bassi et al. (2007), suggests that a higher scale of incentives can increase attention, at least up to a point, and that higher attention can increase performance, at least up to a point that is determined in part by the difficulty of the problem.

This pattern has implications for the kinds of inferences that can be made from studies employing monetary incentives. The nature of these implications can be reasonably expected to differ depending on the nature of the experimental findings. Consider some of the political communication studies cited previously. In the scenarios of Lupia and McCubbins (1998), for example, subjects are quite good at inferring the informational content of communications they receive from strategically motivated speakers. In such instances, confidence in a result's external validity may depend to some extent on the "calibration" between the financial incentives in play and the stakes involved in receiving analogous communications in the real world. The incentives offered by Lupia and McCubbins appear to be quite appropriate in scale. However, consider a counterfactual experiment in which the monetary stakes for subjects were much larger. If, in this counterfactual experiment, subjects were substantially more motivated to pay attention and make proper inferences by the monetary inducements in the laboratory than they would have been by naturalistic considerations in the real world, then a clear issue would arise in extrapolating from "good" performance in the laboratory to predictions about real-world performance. In contrast, in the "cheap talk and coordination" scenario of Dickson (in press), subjects systematically fail to

account fully for a speaker's strategic incentives when inferring the information content of communications. Of course, proper calibration of financial incentives to real-world motivations would always be an ideal. However, for a study whose central result demonstrates "poor" performance or the existence of a "bias" in subject behavior, confidence in external validity is likely to be stronger when the experimenter errs on the side of making financial incentives too large rather than too small. That is, our confidence that a particular form of bias actually exists will be stronger if it persists even when subjects have extra incentives to perform a task well in the laboratory relative to the weaker incentives they face in real-world settings. This logic underscores the extent to which simple decisions of experimental design may have powerful effects on the inferences we can draw from an experiment, even when the results are the same across different designs. A given finding will generally be more impressive when the experimental design is more heavily stacked against the emergence of that finding.

Potential Problems with Use of Monetary Incentives

As noted previously, monetary incentives may be a nonstarter for some research questions, but there may be arguments in favor of their use for other research questions. Are there potential problems with the use of monetary incentives that may argue against their use in certain settings?

One potential issue involves interactions between subjects' intrinsic motivations and the external motivation that they receive from financial incentives. Some research in psychology suggests that financial incentives can "crowd out" intrinsic motivations, leading to somewhat counterintuitive patterns of behavior. One of the best-known examples of crowding out comes from Titmuss (1970), who showed that offering financial compensation for blood donations can lead to lower overall contribution levels. The standard interpretation is that individuals who donate blood are typically motivated to do so for altruistic reasons; when financial

incentives are offered, individuals' mode of engagement with the blood donation system changes, with marketplace values coming to the fore while intrinsic motivations such as altruism are crowded out.

Whether crowding out poses a problem for the use of monetary incentives is likely to depend on the nature of the research question being explored. For the purposes of game theory testing, the crowding out of intrinsic motivations can actually be considered desirable because the experimenter wants to exogenously assign preferences to subjects in order to instantiate the experimental game in the laboratory. In contrast, suppose that social interactions within some real-world setting of interest are believed to depend heavily on individuals' intrinsic motivations. In translating this real-world setting into the laboratory, injudicious use of monetary incentives could potentially crowd out the intrinsic motivations that are central to the phenomenon being studied.

This potential problem with the use of monetary incentives is in some instances a challenging one because it can be difficult to anticipate to what extent such incentives might cause a transformation in subjects' modes of engagement with the experimental scenario. This concern goes hand in hand with understandable questions about the extent to which stylized economic and contextually rich psychological experiments actually investigate the same cognitive mechanisms, an important and understudied matter that may be illuminated more thoroughly in the future by across-school collaborations, as well as by neuroscientific and other frontier research methods.

3. Use of Deception

In few regards is the difference between the economic and psychological schools as stark as in attitudes about deceiving subjects. The more-or-less consensus view on deception in the experimental economics subfield is simple: just don't do it. In contrast, deception has been and remains fairly commonplace within the political psychology research tradition.

This section describes potential advantages and disadvantages of using deception from a methodological and inferential perspective. Ethical considerations are not discussed here due to space limitations (for a recent review, see Morton and Williams 2010).

Lack of Deception in Experimental Economics

Deep-seated opposition to the use of deception has become a feature of various institutions within the economics discipline. It is common for experimental economics laboratories to publicize and enforce bans on deceiving subjects; in fact, a strong norm among practitioners and journal editors makes experiments employing deception de facto unpublishable in major economics journals.

Before describing the motivations for these norms, it is worth describing what "deception" means, and does not mean, to experimental economists. A rough distinction can be made between sins of *commission* and sins of *omission*. Describing features of the experimental scenario in a way that is either explicitly dishonest or actively misleading – a sin of commission – would be straightforwardly considered a taboo act of deception by experimental economists. In contrast, a failure to fully describe some features of the experimental scenario – a sin of omission – would not necessarily be counted as a deceptive act. As Hey (1998) puts it, "There is a world of difference between not telling subjects things and telling them the wrong things. The latter is deception, the former is not" (397). Thus, in several studies of public goods provision, experimentalists employ a "surprise restart," in which a second, previously unannounced public goods game is played after the completion of the first. As long as subjects are not actively misled by the wording of the experimental protocol, such a procedure is not considered to be deceptive. And, of course, few scholars would argue that it is necessary to explicitly inform subjects about the purpose of the study in which they are taking part.

What arguments do experimental economists present against the use of deception?

Both Bonetti (1998) and Morton and Williams (2010) cite Ledyard (1995) as offering a standard line of reasoning:

It is believed by many undergraduates that psychologists are intentionally deceptive in most experiments. If undergraduates believe the same about economists, we have lost control. It is for this reason that modern experimental economists have been carefully nurturing a reputation for absolute honesty in all their experiments. . . . [i]f the data are to be valid. Honesty in procedures is absolutely crucial. Any deception can be discovered and contaminate a subject pool not only for the experimenter but for others. Honesty is a methodological public good and deception is not contributing. (134)

At the heart of this case is the fear that the use of deception will lead to a loss of experimental control; as we have seen, many features of economics-style experimentation, including the use of stylized experimental scenarios and monetary incentives, are designed to help maintain experimental control of different kinds. Hey (1991) articulates the specific nature of this concern:

[I]t is crucially important that economics experiments actually do what they say they do and that subjects believe this. I would not like to see experiments in economics degenerate to the state witnessed in some areas of experimental psychology where it is common knowledge that the experimenters say one thing and do another. . . . [O]nce subjects start to distrust the experimenter, then the tight control that is needed is lost. (171–73)

This kind of concern about experimental control is quite natural given the typical nature of research questions in experimental economics. As noted previously, most economics experiments either test the predictions of game-theoretic models or explore the nature of behavior in game-theoretic settings. Crucially, the most common concepts of equilibrium in games, from which predictions are derived, assume that actors share common knowledge about basic features of the game being played. Of course, experimental subjects learn about “the rules of the game” through the experimenter. If

researchers indeed do, as Hey fears, develop a reputation for employing deception in their experiments, then subjects may develop heterogeneous beliefs about what is really going on in the laboratory – while also being aware that other subjects are doing the same. At the end of the day, subjects could effectively find themselves playing a wholly different game than the one the experimenter had intended. The conjectures within subjects’ minds about the true nature of the game would, of course, be essentially unknowable not only to one another, but also to the analyst.

Ledyard’s (1995) opinion also reflects a common viewpoint among experimental economists, namely, that a lab can benefit from maintaining a reputation for transparency with its subject pool. Such a reputation, it is argued, could quickly be squandered if deception takes place in the laboratory; the subject pool may become “tainted” with subjects who have themselves experienced deception or who have been told about it by friends.

This argument is reasonable, but the question it bears on is ultimately an empirical one. Relatively little systematic research has explored this point, but there is some evidence that the experience of deception in the laboratory may affect individual subjects’ propensities to participate in future experiments as well as their behavior in future experiments (Jamison, Karlan, and Schechter 2008). To my knowledge, there has been no systematic research into a related issue: the extent to which experimental economics laboratories who do ban deception actually attain the reputations to which they aspire – that is, to what extent subjects are aware of lab policies on deception in general or actually believe that they are never being deceived while taking part in particular experiments in “no deception” labs. Economists’ arguments about the sanctity of subject pools further tend to presuppose that psychology departments do not exist, or at least that they draw from a disjointed set of participants. If psychology and economics labs operate simultaneously at the same university, to what extent do undergraduate subjects actually perceive them as separate entities with distinct reputations? Does the physical proximity of the

labs to each another affect subject perceptions (e.g., if they are in the same building as opposed to different buildings)? It would appear that such questions remain to be answered.

Use of Deception in Experimental Political Psychology

In contrast, the use of deception is quite common in political psychology, as it is in social psychology. As we have seen, the reasons for this difference can be understood as springing from the distinctive natures of inquiry and theory testing in the two schools. Importantly, the ability to induce common knowledge of an experimental scenario within a group of subjects is usually not nearly so crucial for experiments in the political psychology tradition, which typically do not involve tests of game-theoretic models. This subsection reconsiders the advantages and disadvantages of deception in the context of political psychology research questions.

One prominent class of examples can be found in the study of political communication, in which scholars quite frequently present subjects with stimuli that are fabricated or falsely attributed. Thus, Brader (2005) presents experimental political advertisements to subjects as though they were genuine ads from a real, ongoing campaign; meanwhile, Druckman and Nelson (2003) present experimental newspaper stories to subjects as though they came from well-known outlets such as the *New York Times*.

In the following paragraphs, I use these articles as examples in discussing potential advantages of deception. Throughout, I take as the salient alternative an otherwise identical experimental design in which the same stimuli are presented to subjects, but explicitly labeled as “hypothetical” campaign ads, newspaper stories, etc. Of course, in certain circumstances, different counterfactual designs might also be reasonably considered.

In judging the potential usefulness of deception, then, a natural question to ask is whether an individual’s mode of psychological engagement with a stimulus depends

on whether that stimulus is framed as being “real” as opposed to hypothetical. If the answer to this question is “yes” – and if this would make a substantial enough difference for measurements of the quantities of interest – then at the least a benefit from deception will have been identified. Ultimately, of course, in any given setting it is an empirical question whether the answer will be “yes” or “no.” To my knowledge, however, no systematic studies have been carried out measuring the effects, if any, of choosing deceptive as opposed to explicitly hypothetical experimental scenarios.

Taking Druckman and Nelson’s (2003) design as an example, though, it at least seems plausible that the difference may sometimes be considerable. An individual picking up what he or she believes to be an article from the *New York Times* will respond to frames and other cues in a way that depends directly on his or her relationship with the *New York Times* – his or her sense of the newspaper’s reliability, the fit of its ideology with his or her own, and so forth. In contrast, a hypothetical exercise of the form “suppose the *New York Times* reported . . .” could insert in the subject’s mind a mysterious intermediary between the newspaper and the subject. Who is it that is doing this supposing, and what are they up to? Alternatively, the subject may simply attend differently to the article, paying it less heed or greeting it with less trust, if he or she knows from the offset that it is a fiction. Under such circumstances, it would not be unreasonable to suppose that a given article might have less of an effect than it would have had it been described as a “real” article. Although economically inclined scholars might tend to doubt whether experiments employing deception can ever gain a full measure of experimental control, it is arguable in this setting that more control might be lost with an explicitly hypothetical stimulus than with a deceptive one. Whether this is true, of course, depends on the extent to which subjects were actually successfully deceived. This, however, is the sort of question that can often be addressed through the use of simple manipulation checks by the experimenter. At least in this example, the treatment effects

in Druckman and Nelson's findings strongly suggest that the deceptive manipulation did indeed have the desired effect on subjects.

In a similar way, it seems plausible that deception may be a useful element of Brader's (2005) design. In part, this is arguable because of the nature of some of Brader's dependent variables. Among other things, Brader shows that the use of music in contrived political advertising can affect subjects' self-reported level of inclination to seek more information about an election campaign, whereas the idea of asking subjects to report their level of inclination to seek more information about a hypothetical campaign that means nothing to them seems straightforwardly problematic.

These examples suggest that deception may offer access to certain research questions that would remain inaccessible in its absence. Psychologists also claim that deception may be necessary at times to conceal the purpose of an experiment from subjects (Bortolotti and Mameli 2006), and they are frequently concerned about the possibility of "Hawthorne effects," through which subjects attempt to meet what they perceive to be the experimenter's expectations. Such effects can be particularly worrisome in sensitive research areas, such as the study of racial politics.

Finally, it could be argued that the use of deception can sometimes strengthen the inferences that are possible from a given piece of research. Among the most famous experiments in social psychology is the seminal Milgram (1974) experiment on obedience and authority. In the experiment, subjects were deceived into believing that they could, with the twist of a knob, deliver electric shocks of increasing magnitude to another person; an authority figure urged subjects to deliver such shocks in the context of a staged scenario. In the end, a large fraction of subjects did conform to the authority figure's commands, to the point of delivering highly dangerous voltages.

This is a rather shocking result, one that had a profound effect on the study of authority specifically and on social psychology more generally. Its power, of course, comes from our sense that subjects really did

believe – at least to a considerable extent – that their actions were causing bodily harm to another human. An otherwise comparable study involving an explicitly hypothetical scenario would, for obvious reasons, have been far less convincing, even if it yielded the same results. It could be easily argued that Milgram's (1974) act of deception was central to the lasting influence of Milgram's study.

References

- Bassi, Anna, Rebecca Morton, and Kenneth Williams. 2007. "The Effects of Identities, Incentives, and Information on Voting." Paper presented at the Political Economy Seminar Series, Kellogg School of Management, Northwestern University, Evanston, IL. Retrieved from www.kellogg.northwestern.edu/research/fordcenter/documents/morton0502072.pdf (October 30, 2010).
- Bonetti, Shane. 1998. "Experimental Economics and Deception." *Journal of Economic Psychology* 19: 377–95.
- Bortolotti, Lisa, and Matteo Mameli. 2006. "Deception in Psychology: Moral Costs and Benefits of Unsought Self-Knowledge." *Accountability in Research* 13: 259–75.
- Brader, Ted. 2005. "Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions." *American Journal of Political Science* 49: 388–405.
- Camerer, Colin F. 2003. *Behavioral Game Theory*. Princeton, NJ: Princeton University Press.
- Camerer, Colin F., and Robin M. Hogarth. 1999. "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework." *Journal of Risk and Uncertainty* 19: 7–42.
- Chen, Yan, and Sherry Xin Li. 2009. "Group Identity and Social Preferences." *American Economic Review* 99: 431–57.
- Dickson, Eric S. in press. "Leadership, Followership, and Beliefs about the World: An Experiment." *British Journal of Political Science*.
- Dickson, Eric S., Catherine Hafer, and Dimitri Landa. 2008. "Cognition and Strategy: A Deliberation Experiment." *Journal of Politics* 70: 974–89.
- Druckman, James N., and Kjersten R. Nelson. 2003. "Framing and Deliberation: How Citizens' Conversations Limit Elite Influence." *American Journal of Political Science* 47: 729–45.

- Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90: 980–94.
- Gneezy, Uri, and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All." *Quarterly Journal of Economics* 115: 791–810.
- Griggs, Richard A., and James R. Cox. 1982. "The Elusive Thematics Material Effect in Wason's Selection Task." *British Journal of Psychology* 73: 407–20.
- Hey, John D. 1991. *Experiments in Economics*. Oxford: Blackwell.
- Hey, John D. 1998. "Experimental Economics and Deception: A Comment." *Journal of Economic Psychology* 19: 397–401.
- Jamison, Julian, Dean Karlan, and Laura Schechter. 2008. "To Deceive or Not to Deceive: The Effect of Deception on Behavior in Future Laboratory Experiments." *Journal of Economic Behavior and Organization* 68: 477–88.
- Kagel, John H., and Alvin E. Roth. 1995. *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.
- Ledyard, John O. 1995. "Public Goods: A Survey of Experimental Research." In *The Handbook of Experimental Economics*, eds. John H. Kagel and Alvin E. Roth. Princeton, NJ: Princeton University Press, 111–94.
- Levine, David K., and Thomas R. Palfrey. 2007. "The Paradox of Voter Participation: A Laboratory Study." *American Political Science Review* 101: 143–58.
- Lupia, Arthur, and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge: Cambridge University Press.
- McDermott, Rose. 2002. "Experimental Methodology in Political Science." *Political Analysis* 10: 325–42.
- Milgram, Stanley. 1974. *Obedience to Authority*. New York: Harper & Row.
- Morton, Rebecca B., and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. New York: Cambridge University Press.
- Prior, Markus, and Arthur Lupia. 2008. "Money, Time, and Political Knowledge: Distinguishing Quick Recall and Political Learning Skills." *American Journal of Political Science* 52: 169–83.
- Tajfel, Henri, Michael Billig, R. Bundy, and Claude L. Flament. 1971. "Social Categorization and Inter-Group Behavior." *European Journal of Social Psychology* 1: 149–77.
- Tajfel, Henri, and John Turner. 1986. "The Social Identity Theory of Intergroup Behavior." In *Psychology of Intergroup Relations*, eds. Stephen Worchel and William G. Austin. Chicago: Nelson-Hall.
- Titmuss, Richard M. 1970. *The Gift Relationship: From Human Blood to Social Policy*. London: George Allen & Unwin.
- Wason, P. C. 1968. "Reasoning about a Rule." *Quarterly Journal of Experimental Psychology* 20: 273–81.

