

Field Experimental Work on Political Institutions

Christian R. Grose

Department of Political Science, University of Southern California, Los Angeles, California 90089-0044; email: cgrose@dornsife.usc.edu

Annu. Rev. Polit. Sci. 2014. 17:355–70

First published online as a Review in Advance on February 21, 2014

The *Annual Review of Political Science* is online at polisci.annualreviews.org

This article's doi:
10.1146/annurev-polisci-072012-174350

Copyright © 2014 by Annual Reviews.
All rights reserved

Keywords

field experiments, representation, accountability, legislatures, courts, bureaucracy

Abstract

A nascent but growing research area examines political institutions through the use of field experiments. I consider why field experimentation has been used infrequently in the study of political institutions and note that some research questions are not amenable to field experimentation. I review areas of research inquiry where field experimentation has enhanced scholarly knowledge about political institutions and representation. These areas include the study of race, representation, and bias in legislatures and courts; and policy responsiveness and legislative accountability. I synthesize this research by examining puzzles that emerge between the field experimental and observational work. I conclude with suggestions for promising research avenues, including the use of field experiments to study the bureaucracy. The discipline's understanding of political institutions could be improved with a greater emphasis on field experimental work.

INTRODUCTION

An experimental revolution is occurring in contemporary political science. Research utilizing experimental methods—as well as attempts to examine relationships with natural or quasi-experiments—is common in most leading journals in political science. To some social scientists, a randomized experiment is the gold standard for empirical researchers interested in causal identification (though for contrary views, see Heckman & Smith 1995, Heckman 2005, and Deaton 2010). Because of this promise of causal identification, as well as an increasing acceptance of experimental methods in certain subfields in the discipline, there has been a dramatic increase in political science experiments (Druckman et al. 2006, 2011; Morton & Williams 2010).

The experimental revolution in political science has extended to work conducted in the field. Field experiments randomly assign treatment and control conditions to subjects in settings outside of the laboratory (Gerber 2011, Gerber & Green 2012). Some field experiments are conducted in highly realistic settings, such as when researchers embed experiments in cooperation with political campaigns (e.g., Alvarez et al. 2010, Gerber et al. 2011), whereas other experiments are researcher-led interventions in the field. Field work has historically been conceived of as participant observation, qualitative interviews, or the collection of observational or survey data in field locations. The explosive growth of field experiments in political science has allowed scholars to rethink what going into the field means; and field experiments are a useful bridge between case studies and more systematic hypothesis testing and data collection.

Until the late 2000s, most political science studies utilizing field experiments examined the political behavior of citizens. Within the political behavior subfield, field experiments have achieved the greatest prominence in studies of voter mobilization (de Rooij et al. 2009, Michelson & Nickerson 2011). Without question, scholars of political behavior were the early adopters of field experimental methods, and scholars of political institutions have been slower to embrace field experiments.

A nascent but growing research area, however, examines political institutions through the use of field experiments. Much of this work examines political representation and elite decision making within specific political and electoral institutions. Usually, most empirical studies of political institutions use observational data (though there is also a strong research tradition of laboratory experiments examining decision making within political institutions; e.g., Fiorina & Plott 1978, Wilson 1986, Palfrey 2009, Boudreau et al. 2010).

I discuss the vast potential for field experiments in the study of political institutions and point to ways that field experiments can be used in future research. Not all questions about political institutions—or in political science—can be answered via field experimentation or via experimentation generally. However, our understanding of political institutions is greatly enhanced when scholars use field experiments. In a theory-rich subfield like political institutions, scholarly advances can be achieved by leveraging field experiments to test implications of theories of institutions.

In this article, I discuss the benefits of field experiments and consider why field experimentation has been used infrequently in the study of political institutions. In the first section, I also suggest that scholars use field experiments to study the bureaucracy and local and subnational political institutions. Second, I review the two substantive areas of inquiry where field experimentation on political institutions has been most common. These areas include the study of race, representation, and bias in legislatures and courts; and policy responsiveness and legislative accountability. I synthesize this research by noting puzzles that emerge between the field experimental and observational work on these topics. Finally, I conclude with suggestions for promising research questions that could be analyzed with field experiments in future work.

WHY WE SHOULD USE FIELD EXPERIMENTS TO STUDY POLITICAL INSTITUTIONS, AND WHY MANY HAVE NOT

The study of political institutions is theory rich, and arguably more theory rich than other subfields in political science. For instance, numerous theories of the effect of electoral institutions on the behavior of public officials (e.g., Ashworth 2012) and numerous theories of the effect of institutional arrangements and the separation of powers on policy outcomes and elite behavior (e.g., Barnes 2007, Gailmard & Patty 2012, Epstein & Knight 2013) exist. Whereas observational data have been used to test some of the implications of these theories, there is the exciting possibility of more work that can be done that leverages field experimentation. Because scholars studying institutions have developed so many theories, there are significant opportunities to design and conduct field experiments on institutions where there are clear predictions and resultant hypotheses. Some empirical work on political institutions using observational data suffers from causal identification problems, often with empirical results confirming multiple and competing theoretical predictions (e.g., observational studies of legislative shirking, which can be defined as elected officials exerting little effort on behalf of constituents; see discussion in Grose 2010). These research areas provide the most obvious opportunities for scholars of political institutions to conduct field experiments.

A number of political scientists have already made the case for the use of experiments in general and field experiments in particular (McDermott 2002, Humphreys & Weinstein 2009, Morton & Williams 2010, Gerber & Green 2012), so I will only briefly remind readers of some advantages. One advantage of experimental methods is that, in expectation, confounding factors are controlled for through randomization. Another key advantage of a well-designed and carefully implemented field experiment is that it will have greater external validity than laboratory studies, yet it will also have more internal validity than a correlational or observational study, given that the researcher is able to control the assignment of subjects to treatment and control groups. There are also disadvantages to field experiments, including high costs of implementing field studies and a reduction in internal validity relative to a laboratory experiment. Even with field experiments, external validity is often limited, as the case examined may not generalize to other settings.

Despite the advantages of the field experimental approach, scholars of institutions have not frequently embraced field experiments in political science for a number of reasons:

1. A researcher is simply not able to design treatments that alter rules and institutions as would be needed to address some questions; or, the researcher is interested in general-equilibrium effects that cannot be easily analyzed with the measurement of one outcome following a specific randomized intervention.
2. Samples in many institutional settings are too small to achieve sufficient statistical power.
3. There is a belief among scholars that repeated field experiments on the same sample will burden elites in many political institutions. The pool of subjects could be depleted.
4. There is a scholarly perception that political elites in institutions are not systematically accessible and thus are not able to be treated in a field experiment.
5. Scholars of institutions associate experiments with political behavior and psychology, and scholars of political institutions have access to significant amounts of observational data. This path-dependent divide between subfields leads to inertia in methodological approaches in the study of political institutions.

The first two are reasonable explanations for why scholars of political institutions have not conducted field experiments. The third and fourth may be reasonable in some political institutions, though I would suggest that scholars reconsider the inaccessibility of political elites in

institutions in many contexts. The final reason is not a legitimate reason to avoid field experimentation by political-institutions scholars, yet the perception that experiments are the province of behavioralists and psychologists has limited the growth of field experiments on political institutions. I elaborate on these barriers to political-institutions field experiments below.

Some Important Institutional Questions Cannot Be Answered with Field Experiments

There are “questions of central importance” in the subfield of political institutions where “good causal identification is not always possible” (Huber 2013). Relatedly, some questions simply cannot be answered with field experiments. For instance, theories of political control of the bureaucracy suggest that institutional design can affect the levels of delegation and discretion granted to bureaucrats (Miller 2005, Gailmard & Patty 2012).¹ These and other theories posit strategic choices made by institutional actors that are embedded within a specific system of governance, often with a separation of powers across different governmental branches. Although field experiments are useful for assessing partial-equilibrium effects, they are not as useful for assessing general-equilibrium effects. Empirical tests to ascertain the effect of some macro-level institutional characteristics (such as having a presidential versus a parliamentary system, or unified versus divided government in presidential systems) obviously cannot be randomly assigned in the field in established democracies, yet they are vital sources of variation in explaining political and policy outcomes. These are obviously important questions worthy of study, even if field experimentation cannot address them. Even when field experimentation may be possible, the external validity of the study may be limited to the case at hand.

When field experimentation is not practicable but where there is well-developed theory, the pairing of observational analyses and laboratory experiments, or the analysis of natural experiments, may be useful. In addition, when parameters of interest are not easily measured or manipulated by a field experiment (e.g., deep, theoretical parameters like ideal points), structural estimation with observational data is useful.

Sample Sizes and Statistical Power

Some political institutions are simply too small to have enough statistical power for valid inferences in a field experiment. For example, studies of the US Supreme Court can leverage across-time variation with observational data, but one-time field experiments are obviously not practicable with such a small sample (nine justices). Similar sample-size limitations would exist if one were to use field experiments to study the behavior of presidents or top-level executive branch appointees. However, field experiments where trial judges are subjects allow for larger samples (Rachlinski et al. 2009) and greater statistical power.

This sample-size issue may have precluded institutions scholars from using field experiments, but it should not in all instances. The number of potential subjects is large enough in some institutions, particularly in some legislatures and lower levels of bureaucracies. In these settings, statistical power is less of a concern (though samples will not approach the sizes seen in

¹ Predictions from these theories have been tested using ideal point estimates derived from observational roll-call data (e.g., McCarty & Poole 1995; Martin & Quinn 2002; Bailey 2007; Epstein et al. 2007; Bertelli & Grose 2009, 2011; Clinton 2012; Clinton et al. 2012).

get-out-the vote field experimental studies). The relatively larger samples of legislatures may explain why most field experimentation on institutions has thus far been conducted in legislatures (Bergan 2009, Grose 2010, Butler & Broockman 2011, Butler et al. 2012, Dropp & Peskowitz 2012, Richardson & John 2012, Broockman 2013a, Grose et al. 2013, Harden 2013). In the United States, there are more than 7,000 state legislators, and every two to four years, there is a reasonable amount of turnover in these legislatures. In fact, sample size is one reason the largest US state legislative body (New Hampshire's lower chamber, with 400 members) has been studied in more than one field experiment (Bergan 2009, Grose 2010). However, field experimental studies have also been conducted on small legislatures like the New Mexico state legislature (Butler & Nickerson 2011) or subsamples of the US Congress (Chin et al. 2000, Grose et al. 2013).

Repeated Field Interventions Could Burden Political Elites

A related concern that may have kept institutional scholars from using field experiments is the fear that too-frequent intervention via field experimentation by a large number of scholars can place time burdens on legislators and other public officials. For ethical reasons, when possible, scholars should design experiments that yield minimal time costs for elected and public officials and reduce other risks to subjects (Butler 2014, McClendon 2012, Malesky 2013).² However, most field experimental interventions with legislators, for instance, have been messages delivered by mail, phone, or in person that require minimal time costs to the legislators.³ At least in the existing research on legislatures, this concern may be overblown.

Nevertheless, scholars should carefully design their research, and the importance of the research question should outweigh the costs of field experimentation when the subject pool is potentially limited. An academic clearinghouse for field experiments on political elites could be developed, so that other scholars could target legislatures and other political institutions that have not had a large number of interventions. Such a clearinghouse could also encourage the reuse of prior experimental intervention data to examine downstream effects on other dependent variables of interest (Green & Gerber 2002). Given that the use of field experiments to study political institutions has been an emerging area of inquiry with only a handful of scholars intervening in any given legislature, this has not been a significant concern. It may be a greater concern going forward.

In addition to a clearinghouse, other solutions to the repeated interventions issue would include scholars studying institutions in which the likelihood of being a repeat subject is low. Local officials are much less frequently studied with field experiments than are state and national officials [though see Butler (2014), who studies municipal elected officials in the United States, and Richardson & John (2012), who study local councilors in the United Kingdom].

²Typically, field experiments of public officials engaging in public behavior fall under the exempt research categories for institutional review boards (IRBs) (though scholars should seek clarification and exemption or approval for all studies from their local IRBs). IRB research exemption exists for public officials because nearly all of their behavior is, by definition, public. However, behavior of public officials has important gradations, and the ethical obligations for the protection of public officials as human subjects should be considered carefully. For a more in-depth discussion of the ethical considerations of conducting field experiments on public officials, see Humphreys & Weinstein (2009, pp. 375–76), Butler (2014), McClendon (2012), Grose (2013), and Malesky (2013).

³Some also express concern that too many interventions with the same subjects (e.g., legislators) could potentially alter results of future research. For instance, legislators who have been subject to repeated interventions to assess racial or ethnic bias among legislators may learn about the studies once they are published and update their behavior in future actions with fictional constituents (thus possibly yielding null results in future studies when racial bias among legislators may still exist). This problem, of course, exists in any scholarly intervention where the subjects participate in repeated interventions (e.g., panel surveys, convenience samples where subjects participate in numerous laboratory experiments).

One political institution that has a numerically large potential subject population but that has been subject to very little field experimentation is the bureaucracy, at both the federal and the subnational levels [though see Mason (2013), who examines organizations].⁴ By one estimate, there are more than 22 million people who work for federal, state, or local governments in the United States alone (Moore 2011). Even more are administrators in nonprofit and nongovernmental organizations. Depending on the research question, there is a potentially unlimited subject pool where field or survey experiments could be conducted. Studies could examine questions of shirking, discretion, bias, efficiency in decision making, and the provision of public goods by street-level bureaucrats as well as higher-level public officials. The likelihood of significant numbers of burdensome, repeated interventions would be low. One of the only field experiments to examine delegation decisions by executives to agencies is Avellaneda (2013), who finds that mayors in Latin America are less likely to delegate authority in policy areas that have electoral-geographic benefits (infrastructure policy).

I encourage scholars to think creatively about how to use field experimentation to test important theoretical questions of institutions in settings where the subject pool is potentially very large. A by-product of an increased use of field experimentation with these subjects could be an invigoration of the study of local and subnational political institutions. Of course, one external validity limitation of this approach is that the results and conclusions from samples with large potential-subject pools (e.g., state bureaucracies, state and local legislatures) may not generalize to some national institutions (e.g., top levels of the federal bureaucracy, national legislatures).

Inaccessibility of Elites in Political Institutions: A Belief That Treatments Will Not Reach the Subjects

Certain institutions—and those individuals within the institutions—are simply more accessible and thus open to field experimental interventions. Local governmental institutions, legislators, and some bureaucrats/public administrators interact one-on-one with members of the public regularly and thus are quite accessible to researchers as well. Supreme Court justices, most federal judges, and some top executive branch officials would be exceedingly difficult for most scholars to access for these purposes and thus are not prime candidates for field experimental interventions. However, my hunch is that field experimentation has not been as frequent among institutions scholars because of the misperception that public officials are not systematically accessible, when in fact many are.

There is a variant of this concern: the belief among scholars that ordinary citizens are malleable whereas elites in institutions are less so. Scholars with this belief may not adopt field experimental interventions with elites, presuming that no intervention will be strong enough to cause elite behavior or outcomes to change. I view this as an open empirical question. Much of the research I review in the next section has found important differences in decisions by legislators when treated with field experiments, so more research is needed in other institutional domains to see if other field experimental interventions will have effects on decision making. To conduct and implement such studies, the researcher may need to partner with public officials or existing interest groups.

⁴There have been a significant number of policy evaluation field experiments. In these instances, the subjects are often citizens, but bureaucrats may vary the treatment of a policy to the citizens. Much less common are field experiments where the decisions of public administrators and officials are examined following an experimental treatment in order to learn about decision making within bureaucracies and government organizations.

Institutions Scholars Associate Experiments with Political Behavior and Psychology

Finally, one of the reasons a limited number of scholars of institutions have conducted field experiments is that there is not an extensive history of the method in the subfield, though this is changing. A number of scholars of institutions have told me that they simply are not regular consumers of work in the behavior subfield and thus had not really thought about using field experiments because they infrequently read and review scholarship that utilizes them. Scholars also associate experiments with work in political psychology and thus have not gravitated toward their use when studying political institutions.

We should not let subfield and disciplinary barriers or preconceptions keep us from utilizing methods that can provide important empirical leverage on key questions in the subfield of political institutions.

In sum, not all questions addressing political institutions are answerable with experimentation. However, scholars of institutions should be more open to the use of field experiments in research projects going forward.

FIELD EXPERIMENTS BREAK NEW GROUND IN THE STUDY OF POLITICAL INSTITUTIONS

Despite initial reticence, scholars have begun to utilize field experiments with increasing regularity to study political institutions. I consider the two areas of research inquiry where field experimentation has been most common: representation, race, and bias in political institutions; and policy responsiveness and accountability in legislatures.⁵ In some instances, the field experimental work has confirmed expectations from earlier observational work, whereas in other instances there are important differences between the field experimental and observational work.

Representation and Race in Institutions: Do Political Elites Exhibit Racial Bias?

One area where field experiments have made significant contributions is the study of race and representation in legislatures and the courts. An extensive theoretical and empirical literature examines race and representation in legislatures (Griffin 2014), and the empirical literature is mostly based on observational data. A question frequently asked in this literature is, are black legislators more responsive to black constituents than legislators of other races?

Pitkin (1967) theorizes generally that descriptive representation (in this instance, the presence of black legislators in office) does not affect substantive representation (responsiveness on policy), and Swain's (1993) observational study found support for Pitkin's expectations when studying black and white legislators. In contrast, other theorists and empirical scholars using observational data have argued that black legislators are more responsive to black constituents than are white legislators (e.g., Whitby 1997; Canon 1999; Mansbridge 1999; Haynie 2001; Grose 2005, 2011; Grose et al. 2007; Minta 2011). Again based on observational data, others find mixed support

⁵In a number of studies political elites are used to deliver a treatment to regular citizens. For example, Cover & Brumberg (1982) examine the effect of a treatment of congressional mail on mass opinion, whereas Esterling et al. (2011) use the field setting of telephone town hall sessions conducted by members of Congress as experimental treatments to study how citizens deliberate. Another example is Wantchekon (2003), where political parties in Benin were used to deliver clientelism messages to treatment groups of citizens to assess the effect of these messages on mass behavior. I limit my discussion to those field experiments that treat individuals serving within institutions, omitting instances in which the institution is used as a method of treatment delivery to test questions of citizen behavior.

that the presence of black legislators yields outcomes beneficial to black constituents (Tate 2003, Griffin & Flavin 2007) or find that constituency-level factors are important (Lublin 1997).

Those who have found that black legislators are more responsive to black constituents than white legislators have argued that these results are (*a*) due to a linked fate between black legislators and black constituents, or (*b*) due to differing electoral considerations of black and white legislators. The linked-fate perspective suggests that black legislators and citizens have shared histories of discrimination and other experiences (Whitby 1997, Haynie 2001). This linked fate means that black legislators have personal preferences on concerns of importance to the black community that are different from white legislators' preferences. While not rejecting the linked-fate perspective, others emphasize that some black legislators are more likely to draw on electoral support from black voters than are legislators of other races (Canon 1999, Grose 2011). In this latter view, strategic electoral considerations will lead to black legislators better representing black constituents.

These are two distinct arguments. Yet existing works based on correlational regression analyses—usually of legislator outputs as a dependent variable and the race of the legislator as an independent variable—are unable to create research designs and measures that causally distinguish between the two arguments (though Grose 2011 does leverage court-ordered redistricting as a natural experiment).

Butler & Broockman (2011), in the first field experimental study of race and representation in legislatures, have moved this observational literature forward. Similar to an audit study where names of fictional job candidates were changed on otherwise identical resumes in order to assess labor market discrimination (Bertrand & Mullainathan 2004), Butler & Broockman (2011) randomly assigned US state legislators into different treatment groups to determine whether legislators discriminated against black constituents; whether this discrimination, if it existed, was conditional on the race of the legislator; and whether the legislators differentially responded to constituents of different races for strategic, electoral reasons or because of prejudice.

Butler & Broockman (2011) had fictional constituents send each legislator an identical request for assistance and changed only the constituent's race and partisan affiliation in the service requests. In general, white constituents received more responses than black constituents. White legislators responded to white constituents at higher rates than black constituents, and black legislators responded to black constituents at higher rates than white constituents. Most interestingly, this differential in response was driven more by simple bias, or "taste-based discrimination," and less by legislators' strategic partisan considerations (though both mattered) (Butler & Broockman 2011, p. 465).

Empirically, this work is an advance because the researchers exogenously varied the race of the constituent and showed that legislators are more responsive to constituents who share the legislators' racial backgrounds. Their stronger evidence for taste-based discrimination by legislators suggests that personal preferences of legislators may drive differential responsiveness. This provides complementary support for the linked-fate argument in other racial representation studies, given that legislators' biases are presumably a reflection of their personal preferences. It also suggests that electoral and partisan concerns of legislators may be a secondary explanation.

Other field experiments have also addressed questions of race, ethnicity, and representation. McClendon (2013) treated elected officials in South Africa: Some received a treatment message from a constituent from the same ethnic background, and others received a message from a constituent that did not share the same ethnic background. Like Butler & Broockman (2011), she found that public officials were more likely to respond to constituents of the same ethnic background.

Harden (2013) conducted a list experiment of US state legislators to determine whether black legislators were more likely to view one of their roles in legislatures as representing people of

the same ethnic background.⁶ Both white and black legislators who were in the treatment group in the list experiment were more likely to say that they sought to represent the interests of their specific racial groups, but the effect was much larger among black representatives. Broockman (2013a) examines whether the theoretical concept of virtual representation occurs among black legislators. Virtual representation is the idea that representatives of geographic, single-member districts are not constrained only by the electoral goals of their districts but will sometimes seek to represent those who live outside their districts' boundaries (Pitkin 1967, Mansbridge 2003). Supporting intrinsic motivations of legislators and contrary to a view of legislators as purely driven by reelection incentives, black legislators are much more likely to respond to black constituents from outside of their districts.⁷ The observational research has suggested that racial differences in responsiveness by legislators to constituents has been driven by both constituents (demand) and the legislators themselves (supply), yet the experimental work on race and representation has almost exclusively addressed the supply side. Future experiments should examine constituent reactions to legislators of different races to determine whether the legislator or constituent is the primary driver of racial differences in responsiveness in legislatures. Broockman (2013b) does this by randomly assigning black and white legislators to constituents to assess whether constituents are more likely to contact legislators of their own race. He finds support for the constituency-demand side of racial representation. More experimental work in this vein should be conducted to see if these results are robust to other settings.

Although most experimental work on race and political institutions has been conducted in legislatures, there is also evidence that judges exhibit implicit racial bias. Rachlinski et al. (2009) recruited trial judges to take the implicit association test (IAT). The IAT is designed to measure implicit bias against racial, ethnic, and other groups by examining how quickly (in milliseconds) subjects respond to words and images associated with the groups (see also Pérez 2010 for another study using the IAT with ordinary citizens). They subjected the judges to the IAT and found that many judges harbored implicit biases against blacks. They conducted further experiments where they gave judges accounts of hypothetical cases and asked how they would rule, and they randomized implicit racial primes. They did not find that implicit racial primes had a direct effect on judicial decisions in the hypothetical cases, but they did find evidence that implicit racial primes caused those judges who had high antiblack pretreatment attitudes as measured by the IAT to rule more harshly against black defendants.

A few patterns emerge from these studies of race and political institutions. First, many of the findings support results from observational work that preceded the experimental work. However, the evidence suggests that outright bias—and not strategic response to different groups—underlies much of the findings on race and decision making in political institutions. This departs from some of the observational work on the subject.

⁶A list experiment randomly assigns different survey formats to subjects to garner responses that otherwise might not be offered, especially on sensitive questions. For instance, in the control group, there may be four possible responses to a survey question such that the respondent can choose one, some, or all; whereas in the treatment group, there may be five possible responses such that the respondent can choose one, some, or all. Providing the additional response option in the treatment group allows measures of latent opinion on sensitive topics to emerge. In a traditional survey, legislators may be unwilling to state they are more likely to represent constituents of the same racial background.

⁷One consideration with some of these studies is that staff may have responded to the constituent requests and not the legislators themselves. However, the legislative office as the unit of analysis still provides leverage on these questions. The legislator's staff members are part of the broader "enterprise" (Salisbury & Shepsle 1981, p. 559) in which a legislator engages in his job and communicates with his or her constituents. In addition, some—though not all—of the US state legislatures studied have few or no staff for each legislator. In those states, the legislator chose to respond or not.

Policy Responsiveness and Accountability in Legislatures

Another research area where field experiments are used is the study of accountability and responsiveness in legislatures. This accountability, usually to constituents, can occur through roll-call voting or via other forms of legislator behavior. Most field experiments examine these questions with treatments that attempt to measure constituency preferences or constituency influence. A handful of field experiments also focus on the role that interest groups play in influencing legislator decisions.

A central question in the study of legislative accountability has been how well constituents are able to control the policy decisions of legislators. Underlying many theories of political accountability is the assumption that legislators know the policy preferences of their constituents. However, legislator perceptions of their constituents' views may differ from the preferences of their constituents (Fenno 1978).

In a novel field experiment, Butler & Nickerson (2011) test the relationship between constituency preferences and legislator roll-call decisions by treating randomly assigned legislators with poll results indicating their districts' levels of support for a roll call pending before the New Mexico legislature. The experiment's goal was to determine whether legislators who learn the policy preferences of their constituents are subsequently more likely to vote on a roll call in the direction favored by the constituents. The results of the experiment showed that legislators were more likely to vote in favor of the constituent majority's preferred positions. The implication is that increased legislator knowledge about constituency preferences enhances accountability.

In addition to constituency influence, legislators face pressures from interest groups hoping to affect policy outcomes. One method of interest group influence is grassroots lobbying, where an interest group persuades constituents to contact legislators in support of or in opposition to a policy proposal. Bergan (2009) conducted a field experiment where he treated a randomly assigned group of New Hampshire legislators by sending each a message expressing an explicit position on a bill before the legislature. The control group of legislators did not receive the message. On some floor votes, he found a relationship between the constituent grassroots lobbying message and a legislator's roll-call decision. Similarly, Chin et al. (2000) and Chin (2005) were interested in testing the role of direct lobbying by political action committees (PACs) and constituents in legislator responsiveness. They studied members of Congress, and the outcome variable they examined was whether the legislator's office agreed to meet with a person requesting a meeting. The treatment interventions assigned to each legislator included a request from a constituent for a meeting and a request from a PAC donor for a meeting. In both Chin et al. (2000) and Chin (2005), the constituency treatment yielded higher rates of meeting acceptances than did the PAC donor treatment. These lobbying studies suggest that constituency-based or grassroots lobbying strategies are more effective than direct lobbying strategies. However, these three studies had small samples and somewhat weak effects, so more should be done to replicate these studies in other settings.

Other studies have examined shirking and participation in legislative activities following transparency treatments. In three US state legislatures, Grose (2010) induced legislators to participate in legislative activities via a field experiment (see Malesky et al. 2012 for a similar transparency study in a setting without democratic elections, Vietnam). A randomly assigned treatment group of legislators was encouraged to show up to cast roll-call votes by being told in a written message that their frequency of attendance would be made public to the media and thus possibly constituents at the end of the legislative session. Those in this treatment group participated in roll-call voting at greater rates than those legislators in the control group, who received no message, suggesting the power of constituency concerns in increasing legislator participation rates in roll-call voting.

Accountability to constituents by legislators occurs in venues beyond roll-call voting as well. In field experiments where constituents send messages advocating for a public policy, there is evidence that legislators seek to tailor their explanations to constituents' views, though they do not hide their voting records from constituents with whom they disagree (Grose et al. 2013). In contrast, Richardson & John (2012) find less evidence of legislator tailoring of messages to constituents. A field experiment assessing whether legislators are more likely to respond to service-oriented or policy-oriented appeals shows that legislators are more likely to respond to constituent queries about service (Butler et al. 2012), and another shows that legislator electoral security affects response to constituents (Dropp & Peskowitz 2012).

The common denominator underlying nearly all of these field experimental studies of legislator responsiveness is that accountability to constituents occurs with great frequency. When legislators were informed about the positions of their constituents, they were more likely to vote in line with constituent preferences. When requests for meetings came from constituents, legislators responded to these requests more than to those from nonconstituent donors. When transparency was enhanced, legislators showed up for work. When legislators were lobbied by constituents through grassroots e-mail campaigns, their roll-call positions sometimes changed.

The evidence strongly suggests that constituency preferences matter quite a bit in shaping legislative behavior. This accountability has many normatively desirable qualities and also demonstrates basic principal-agent model predictions where the representative is the agent and the constituency is the principal.

Yet when we situate this field experimental work with other work on accountability in legislatures, a puzzle emerges. The field experimental work suggests that legislators are first and foremost reelection seekers (supporting Mayhew 1974). Constituency influence treatments matter, presumably because legislators care about reelection. However, the field experimental work on race and representation overwhelmingly suggests that biases enter into legislative decision making and that these biases are not driven purely by strategic, electoral goals.

In addition, classic work on congressional decision making suggests that constituency preferences are but one factor influencing legislative decisions (Kingdon 1989). Miller & Stokes (1963), in their landmark study, show that legislators only reflect constituency preferences in certain issue domains and not in others. Fenno (1978) and Bianco (1994) argue that legislators who have established trust with their constituents are able to operate with substantial leeway, often voting without regard to constituent preferences. In addition, parties constrain legislator choices (Rohde 1991). Poole & Rosenthal (1997) show that legislators' revealed preferences do not change significantly over time, even as constituency preferences may, whereas Burden (2007) argues that personal preferences have influence on legislative decisions. These arguments and findings stand in contrast to the field experimental work in which constituency treatments exhibited large effects on roll-call voting and other forms of legislative responsiveness.

The research designs used in field experiments on legislative accountability may help explain why this puzzle emerges in the literature. The treatments are almost universally designed to yield responsiveness to constituency and thus electoral interests (except for the PAC treatments in Chin et al. 2000 and Chin 2005). Future field experimental work should consider treatments that go beyond constituency-oriented responsiveness. Particularly, the collective nature of decision making in congressional committees, parties, and the legislature as a whole should be examined. Most of the experiments reviewed here examine decisions and outcomes in which an individual legislator faces few constraints beyond the constituency treatment. It is possible that the typical state of the world (the control condition in most of these studies) is one in which legislators are often influenced by factors unrelated to the constituency. When a treatment induces constituency influence, though, this typical state of the world disappears and legislators demonstrate accountability.

It is also possible that the classic observational work has underestimated the role of constituency interests while placing a higher premium on additional explanations for roll-call voting and policy decisions. Whereas I find this unpersuasive, more field experimentation is needed to determine how influential constituency explanations are in shaping legislative decisions. On the basis of the experimental evidence thus far, they appear to take a more prominent position than many observational studies would suggest.

CONCLUSION: THE FUTURE OF FIELD EXPERIMENTAL STUDIES OF POLITICAL INSTITUTIONS

As this review has argued, political-institutions scholars were slower to embrace the experimental revolution, and field experiments in particular, than were scholars of political behavior. The institutions in which field experiments have been most commonly employed are legislatures, though some work has also been done on the courts and bureaucracies. To conclude, I offer some suggestions for future research:

- Field experimental work on legislative representation and responsiveness has tested claims made in the normative theoretical literature (Pitkin 1967, Mansbridge 1999). Future research should consider predictions from formal theoretical work as well and test implications of these models of institutions using field experiments.
- Principal-agent relationships can be tested using field experiments, particularly by using bureaucrats as subjects. Decades ago, Simon (1965, p. 34) stated that “field experiments have not been an important procedure for learning about organizational decision making.” Today, this claim could still be made even though there is much that could be learned by employing field experiments.

For example, competing predictions exist in the theoretical literature regarding the effect of increased monitoring by principals on shirking behavior of bureaucratic agents. In the case of bureaucrats, exerting less effort could result in policy outcomes that are not implemented as the principal (an agency head) may prefer. Some posit that increased monitoring by principals should lead to less shirking, given that direct monitoring reduces the ability for the agent to do what she or he pleases without frequent oversight (Grossman & Hart 1983). Others claim that increased monitoring can lead to a reduction in effort by bureaucratic agents in certain circumstances (Cowen & Glazer 1996).

A field experiment could be employed to test these contrary predictions. A researcher would simply need to convince an executive branch official or agency head to randomly assign workers to either a monitoring treatment group or a control group where there is not a significant amount of monitoring. Then measures of productivity and work effort would be taken for all employees, and the monitoring treatment group would be assessed relative to the control group.

- The studies of race in legislatures and courts have found that political elites exhibit discrimination and bias when making decisions. One additional avenue of inquiry in the study of legislatures is to determine whether these biases in responsiveness are correlated with policy decisions related to race, ethnicity, and civil rights. Is explicit or implicit bias by legislators against racial minority groups correlated with policy decisions that can harm those groups?
- Field experimental research on political institutions looks almost exclusively at decisions made by elites in isolation of decisions made by other elites in the same institution. Future work should consider individual decision making conditional on the collective institutions in which the actors are embedded (such as committees). In addition, spillover effects of

treatments within formal and informal networks of elites could also be studied (Bowers et al. 2013).

- Are elected officials behaviorally different from citizens? One promising area of inquiry in which work is already underway (Butler & Kousser 2013, Enemark et al. 2013) is whether elected officials differ from ordinary people when subjected to laboratory experiments of decision making. Assumptions of rational, goal-oriented behavior underlie many theories of individual choice within political institutions. However, if laboratory experiments conducted with average citizens yield empirical results that do not confirm the predictions of such models, does that mean the predictions of the models can be rejected in the context of decision making with elected officials? Or is it possible that ordinary citizens are behaviorally different from elected and public officials?

In conclusion, there is much work to be done leveraging field experiments to study political institutions. The experimental revolution in political science has begun to reshape how we study individuals within those institutions. Theoretical, experimental, and observational work will be improved by the use of field experiments to study political institutions.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

- Alvarez RM, Hopkins A, Sinclair B. 2010. Mobilizing Pasadena Democrats: measuring the effects of partisan campaign contacts. *J. Polit.* 72:31–44
- Ashworth S. 2012. Electoral accountability: recent theoretical and empirical work. *Annu. Rev. Polit. Sci.* 15:183–201
- Avellaneda CN. 2013. Mayoral decision-making: issue salience, decision context, and choice constraint? An experimental study with 120 Latin American mayors. *J. Public Adm. Res. Theory* 23:631–61
- Bailey MA. 2007. Comparable preference estimates across time and institutions for the court, Congress, and the presidency. *Am. J. Polit. Sci.* 51:433–48
- Barnes J. 2007. Bringing the courts back in: interbranch perspectives on the role of courts in American politics and policy making. *Annu. Rev. Polit. Sci.* 10:25–43
- Bergan DE. 2009. Does grassroots lobbying work? A field experiment measuring the effects of an e-mail lobbying campaign on legislative behavior. *Am. Polit. Res.* 37:327–52
- Bertelli AM, Grose CR. 2009. Secretaries of pork? A new theory of distributive public policy. *J. Polit.* 71:926–45
- Bertelli AM, Grose CR. 2011. The lengthened shadow of another institution? Ideal point estimates for the executive branch and Congress. *Am. J. Polit. Sci.* 55:767–81
- Bertrand M, Mullainathan S. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *Am. Econ. Rev.* 94:991–1013
- Bianco WT. 1994. *Trust: Representatives and Constituents*. Ann Arbor: Univ. Michigan Press
- Boudreau C, McCubbins MD, Rodriguez DB, Weller N. 2010. Making talk cheap (and problems easy): how legal and political institutions can facilitate consensus. *J. Empir. Leg. Stud.* 7:868–85
- Bowers J, Fredrickson MM, Panagopoulos C. 2013. Reasoning about interference between units: a general framework. *Polit. Anal.* 21:97–24
- Broockman DE. 2013a. Black politicians are more intrinsically motivated to advance blacks' interests: a field experiment manipulating political incentives. *Am. J. Polit. Sci.* 57:521–36
- Broockman DE. 2013b. Distorted communication, unequal representation: Constituents communicate less to representatives not of their race. *Am. J. Polit. Sci.* In press. doi: 10.1111/ajps.12068
- Burden B. 2007. *The Personal Roots of Representation*. Princeton, NJ: Princeton Univ. Press

- Butler DM, Broockman DE. 2011. Do politicians racially discriminate against constituents? A field experiment on state legislators. *Am. J. Polit. Sci.* 55:463–77
- Butler DM, Karpowitz CF, Pope JC. 2012. A field experiment on legislators' home styles: service versus policy. *J. Polit.* 74:474–86
- Butler DM, Kousser T. 2013. *How do public goods providers play public goods games?* Presented at State Polit. Policy Conf. May 23–25, Iowa City, IA. <http://www.uiowa.edu/~stpols13/papers/Butler%20How%20Do%20Public%20Goods%20Providers%20Play%20Public%20Goods%20Games%20SPPC%202013.pdf>
- Butler DM, Nickerson DW. 2011. Can learning constituency opinion affect how legislators vote? Results from a field experiment. *Q. J. Polit. Sci.* 6:55–83
- Canon DT. 1999. *Race, Redistricting, and Representation*. Chicago: Univ. Chicago Press
- Chin ML. 2005. Constituents versus fat cats: testing assumptions about congressional access decisions. *Am. Polit. Res.* 33:751–86
- Chin ML, Bond JR, Geva N. 2000. A foot in the door: an experimental study of PAC and constituency effects on access. *J. Polit.* 62:534–49
- Clinton JD. 2012. Using roll call estimates to test models of politics. *Annu. Rev. Polit. Sci.* 15:79–99
- Clinton JD, Bertelli A, Grose CR, Lewis DE, Nixon DC. 2012. Separated powers in the United States: the ideology of agencies, presidents, and Congress. *Am. J. Polit. Sci.* 56:341–54
- Cover AD, Brumberg BS. 1982. Baby books and ballots: the impact of congressional mail on constituent opinion. *Am. Polit. Sci. Rev.* 76:347–59
- Cowen T, Glazer A. 1996. More monitoring can induce less effort. *J. Econ. Behav. Organ.* 30:113–23
- de Rooij EA, Green DP, Gerber AS. 2009. Field experiments on political behavior and collective action. *Annu. Rev. Polit. Sci.* 12:389–95
- Deaton A. 2010. Instruments, randomization, and learning about development. *J. Econ. Lit.* 48:424–55
- Dropp K, Peskowitz Z. 2012. Electoral security and the provision of constituent service. *J. Polit.* 74:220–34
- Druckman JN, Green DP, Kuklinski JH, Lupia A. 2006. The growth and development of experimental research in political science. *Am. Polit. Sci. Rev.* 100:627–35
- Druckman JN, Green DP, Kuklinski JH, Lupia A, eds. 2011. *Cambridge Handbook of Experimental Political Science*. New York: Cambridge Univ. Press
- Enemark DP, Gibson CC, McCubbins MD, Zimmerman BA. 2013. *The effect of holding office on the behavior of politicians*. Work. Pap., Dept. Polit. Sci., Univ. Calif. San Diego. http://polisci2.ucsd.edu/denemark/papers/Enemark_et_al_PoliticianExperiments.pdf
- Epstein L, Knight J. 2013. Reconsidering judicial preferences. *Annu. Rev. Polit. Sci.* 16:11–31
- Epstein L, Martin AD, Segal JA, Westerland C. 2007. The judicial common space. *J. Law Econ. Organ.* 23:303–25
- Esterling KM, Neblo MA, Lazer DMJ. 2011. Means, motive, and opportunity in becoming informed about politics: a deliberative field experiment with members of Congress and their constituents. *Public Opin. Q.* 75:483–503
- Fenno RF. 1978. *Home Style: House Members in Their Districts*. Boston: Little, Brown
- Fiorina M, Plott C. 1978. Committee decisions under majority rule. *Am. Polit. Sci. Rev.* 72:575–98
- Gailmard S, Patty JW. 2012. Formal models of bureaucracy. *Annu. Rev. Polit. Sci.* 15:353–77
- Gerber AS. 2011. Field experiments in political science. See Druckman et al. 2011, pp. 115–40
- Gerber AS, Gimpel JG, Green DP, Shaw DR. 2011. How large and long-lasting are the persuasive effects of televised ads? Results from a randomized field experiment. *Am. Polit. Sci. Rev.* 105:135–50
- Gerber AS, Green DP. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: Norton
- Green DP, Gerber AS. 2002. The downstream benefits of experimentation. *Polit. Anal.* 10:394–402
- Griffin JD. 2014. When and why minority legislators matter. *Annu. Rev. Polit. Sci.* 17:327–36
- Griffin JD, Flavin P. 2007. Racial differences in information, expectations, and accountability. *J. Polit.* 69:220–36
- Grose CR. 2005. Disentangling constituency and legislator effects in legislative representation: black legislators or black districts? *Soc. Sci. Q.* 86:427–43

- Grosec CR. 2010. *Priming rationality: a theory and field experiment of participation in legislatures*. Presented at New York Univ–Coop. Congr. Elect. Study Exp. Polit. Sci. Conf., Feb. 5–6, New York. <http://cess.nyu.edu/conferences/2-2010/papers/GrosecNYU2010.pdf>
- Grosec CR. 2011. *Congress in Black and White: Race and Representation in Washington and at Home*. New York: Cambridge Univ. Press
- Grosec CR. 2013. *Field experiments on elected and public officials: ethical obligations and requirements*. Presented at Ethics Comp. Polit. Exp. Conf., May 1–2, La Jolla, CA
- Grosec CR, Malhotra N, Van Houweling RP. 2013. *Explaining explanations: how legislators explain their policy positions and how citizens react*. Presented at New York Univ–Coop. Congr. Elect. Study Exp. Polit. Sci. Conf., Feb. 15, New York. http://cess.nyu.edu/policon2013/wp-content/uploads/2013/02/grosec_malhotra_vanhouweling_nyu_cess.pdf
- Grosec CR, Mangum M, Martin C. 2007. Race, political empowerment, and constituency service: descriptive representation and the hiring of African-American congressional staff. *Polity* 39:449–78
- Grossman SJ, Hart O. 1983. An analysis of the principal-agent problem. *Econometrica* 20:231–59
- Harden JJ. 2013. Multidimensional responsiveness: the determinants of legislators’ representational priorities. *Legis. Stud. Q.* 38:155–84
- Haynie KL. 2001. *African American legislators in the American States*. New York: Columbia Univ. Press
- Heckman JJ. 2005. Rejoinder: response to Sobel. *Sociol. Methodol.* 35:135–50
- Heckman JJ, Smith JA. 1995. Assessing the case for social experiments. *J. Econ. Perspect.* 9:85–110
- Huber J. 2013. Is theory getting lost in the “identification revolution”? *Monkey Cage*, June 14. <http://themonkeycage.org/2013/06/14/is-theory-getting-lost-in-the-identification-revolution>
- Humphreys M, Weinstein J. 2009. Field experiments and the political economy of development. *Annu. Rev. Polit. Sci.* 12:367–78
- Kingdon JW. 1989. *Congressmen’s Voting Decisions*. Ann Arbor: Univ. Michigan Press
- Lublin DI. 1997. *The Paradox of Representation: Racial Gerrymandering and Minority Interests*. Princeton, NJ: Princeton Univ. Press
- Malesky E. 2013. *Manipulating elites*. Presented at Ethics Comp. Polit. Exp. Conf., May 1–2, La Jolla, CA
- Malesky E, Schuler P, Tran A. 2012. The adverse effects of sunshine: a field experiment on legislative transparency in an authoritarian assembly. *Am. Polit. Sci. Rev.* 106:762–86
- Mansbridge J. 1999. Should blacks represent blacks and women represent women? A contingent yes. *J. Polit.* 61:628–57
- Mansbridge J. 2003. Rethinking representation. *Am. Polit. Sci. Rev.* 97:515–28
- Martin AM, Quinn KM. 2002. Dynamic ideal point estimation via Markov chain Monte Carlo for the U.S. Supreme Court, 1953–1999. *Polit. Anal.* 10:134–53
- Mason DP. 2013. Putting charity to the test: a case for field experiments on giving time and money in the nonprofit sector. *Nonprofit Volunt. Sect. Q.* 42:193–202
- Mayhew DR. 1974. *Congress: The Electoral Connection*. New Haven, CT: Yale Univ. Press
- McCarty NM, Poole KT. 1995. Veto power and legislation: an empirical analysis of executive and legislative bargaining from 1961 to 1986. *J. Law Econ. Organ.* 11:282–312
- McClendon G. 2012. Ethics of using public officials as field experiment subjects. *Exp. Polit. Sci.* 3:13–20
- McClendon G. 2013. *Race responsiveness, and electoral strategy: a field experiment with South African politicians*. Work. Pap., Dept. Polit. Sci., Harvard Univ.
- McDermott R. 2002. Experimental methods in political science. *Annu. Rev. Polit. Sci.* 5:31–61
- Michelson MR, Nickerson DW. 2011. Voter mobilization. See Druckman et al. 2011, pp. 228–42
- Miller G. 2005. The political evolution of principal-agent models. *Annu. Rev. Polit. Sci.* 8:203–25
- Miller WE, Stokes DE. 1963. Constituency influence in Congress. *Am. Polit. Sci. Rev.* 57:45–56
- Minta M. 2011. *Oversight: Representing the Interests of Blacks and Latinos in Congress*. Princeton, NJ: Princeton Univ. Press
- Moore S. 2011. We’ve become a nation of takers, not makers. *Wall Street Journal*, Apr. 1
- Morton RB, Williams KC. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. New York: Cambridge Univ. Press
- Palfrey TR. 2009. Laboratory experiments in political economy. *Annu. Rev. Polit. Sci.* 12:379–88

- Pérez EO. 2010. Explicit evidence on the import of implicit attitudes: the IAT and immigration policy judgments. *Polit. Behav.* 32:517–45
- Pitkin HF. 1967. *The Concept of Representation*. Berkeley: Univ. California Press
- Poole KT, Rosenthal H. 1997. *Congress: A Political-Economic History*. New York: Oxford Univ. Press
- Rachlinski JJ, Johnson SL, Wistrich AJ, Guthrie C. 2009. Does unconscious racial bias affect trial judges? *Notre Dame Law Rev.* 84:1195–246
- Richardson L, John P. 2012. Who listens to the grass roots? A field experiment on information lobbying in the UK. *Br. J. Polit. Int. Relat.* 14:595–612
- Rohde DW. 1991. *Parties and Leaders in the Postreform House*. Chicago: Univ. Chicago Press
- Salisbury RH, Shepsle KA. 1981. U.S. Congressman as Enterprise. *Legis. Stud. Q.* 6:559–76
- Simon HA. 1965. Administrative decision making. *Public Adm. Rev.* 25:31–37
- Swain CM. 1993. *Black Faces, Black Interests: The Representation of African Americans in Congress*. Cambridge: Harvard Univ. Press
- Tate K. 2003. *Black Faces in the Mirror: African Americans and Their Representatives in the U.S. Congress*. Princeton, NJ: Princeton Univ. Press
- Wantchekon L. 2003. Clientelism and voting behavior: evidence from a field experiment in Benin. *World Polit.* 55:399–422
- Whitby KJ. 1997. *The Color of Representation*. Ann Arbor: Univ. Michigan Press
- Wilson RK. 1986. Forward and backward agenda procedures: committee experiments on structurally induced equilibrium. *J. Polit.* 5:390–409