

Lessons on Running a Laboratory Experiment (in Graduate School)

Daniel J. Mallinson, *Penn State Harrisburg*

ABSTRACT Experimentation has taken on a new life in political science. As the use of experimental methods proliferates, it is important for researchers to share their experiences and best practices, particularly with early-career researchers. This article provides reflections from practical experience in the laboratory, particularly geared toward graduate students and early-career researchers who are conducting their first laboratory experiment. These lessons do not apply only to first-time experimenters. Experiences are presented regarding time management, using confederates and deception, incentivizing participation, and keeping a laboratory notebook. Finally, early-career researchers are encouraged to “go for it” if the methods are appropriate to their research question.

Experimentation has taken on a new life in political science, particularly within the study of American institutions and behavior (Butler 2014; Butler and Broockman 2011). By no means are experiments new to the field (Gosnell 1926), but they represent a frequently accessed research tool (Druckman et al. 2011; Morton and Williams 2010), particularly with increased attention to causal inference (Keele 2015). Although field experiments offer researchers less control over the experimental setting, they have the distinct advantage of increasing the external validity of our findings. Conversely, laboratory experiments leverage greater control of the experimental conditions for the purpose of isolating a treatment effect, particularly for psychological mechanisms that are expected to extend beyond the sample used in the experiment (Carlson and Settle 2016; Levitan and Verhulst 2016). Finally, there is increased use of population-based survey experiments for the purpose of generalizing laboratory results (Mutz 2011).

The purpose of this article is not to rehash the benefits and drawbacks of laboratory- versus field- versus population-based experiments, and neither is it to advocate for one method over another. As other scholars have noted, the choice of method depends on the experimental context and the research question (Druckman et al. 2011). Instead, this article provides reflections from practical experience in the laboratory and is particularly geared toward graduate students and early-career researchers who are conducting their first experiment. However, these lessons apply not only to first-timers; many are broadly applicable to anyone conducting laboratory experiments. Because this methodological approach remains an important part of an experimentalist’s toolbox, it also is necessary for practitioners to share experiences—particularly in the beginning, when the learning curve is the steepest. Considering other

early-career researchers, this article presents a series of lessons drawn from a first laboratory experiment. Before describing the lessons, however, it is important to set the context for this learning by explaining the experiment.

THE EXPERIMENT

Conducted in the summer and fall of 2013, the experiment examined the effects of conformity pressure on individual behavior in an extended-discussion setting. It used confederates who were trained to take an opposing position to a single participant on a high-salience and identity-laden issue of debate on the campus of a large research university. Deception was employed so that participants did not know that they were the sole subjects in a given round of discussion. Participants in the control condition completed an online survey that gathered demographic and psychological data, as well as their opinions on a series of issues facing students on campus. Participants in the treatment condition completed this survey but also were brought to a conference room to discuss a single issue with the group of trained confederates. We observed how the participants behaved when confronted with unified opposition by a group of peers and whether they changed their opinion overtly or covertly by the end of the discussion session (more details are in Mallinson and Hatemi 2013). A third-year graduate student, I was the principal investigator. This experiment emerged from a course in political psychology and ultimately became a chapter of my dissertation. Thus, although the following lessons drawn are not only for graduate students, some of the advice pertains particularly to this group of early-career researchers.

LESSON 1. EVERYTHING TAKES LONGER THAN YOU THINK AND HOPE—AND IT SHOULD

I was fairly naive going into this experience. I believed that I would be able to get the study through the Institutional Review Board

Daniel J. Mallinson is assistant professor of public policy and administration at Penn State Harrisburg. He can be reached at mallinson@psu.edu.

(IRB), recruit students, collect data, and present results at the APSA Annual Meeting only nine months away. However, the design stage alone took the remainder of the semester, and then IRB took a few months. IRB wait times, of course, vary greatly by institution. I waited until the study was approved to move forward with recruiting confederates and lining up details (time management is discussed in Lesson 2). The study began haltingly in May and data collection ended in December; this is not an unreasonable timeline.

Instead, this article provides reflections from practical experience in the laboratory and is particularly geared toward graduate students and early-career researchers who are conducting their first experiment.

The design stage *should* take time because a good design is important for preventing threats to inference and producing reliable results. A rushed design can result in the researcher neglecting to measure important variables, creating flawed treatments, omitting necessary controls, and more. All decisions ultimately will need to be justified for publication, and any errors undermine the success of the study. Furthermore, it is important to recognize that not all research questions can be answered by a single experiment. More often than not, follow-up experiments are necessary to adequately answer major questions. It also is helpful to conduct a small pilot study to identify potential problems up front. Finally, the logistical challenges of laboratory experiments are costly in terms of time. This is especially true if the principal investigator is a graduate student taking classes and comprehensive exams or writing a dissertation. Therefore, give an experiment the time it deserves to avoid later regrets.

Of course, even with these seemingly knowable extensions of a project's timeline, inevitable bumps in the road will set the project back: confederates leave the program; there is a snag in the protocol; other work consumes your time for conducting the experiment. Additionally, the logistical details of participant recruitment and coordination of the experiment increase the time to completion. Because we did not have a subject pool, we distributed advertisement flyers on campus and made announcements in the classrooms of supportive faculty. Even if a program has a subject pool, subjects must be prodded to sign up and subsequently reminded of their appointments. Even with several reminders, subjects failed to show up for their time slots. Coordinating schedules among the investigator, the confederates, the subject, and an available conference room proved to be our most substantial logistical challenge.

Thus, it is important to have appropriate expectations for the amount of time needed for the project. This advice is perhaps the most crucial for a PhD candidate considering an experiment with human subjects. Do NOT underestimate how long the experiment can—and should—take. Consider your expectations in terms of a timeline and then add at least three to six months. This is not meant to be discouraging but rather realistic. If you are pushing to finish graduate school quickly and are concerned about the time it will take to plan and execute an effective experiment, it is wise to reconsider. If an early-career researcher has the time, an experiment can enrich a dissertation that otherwise would

rely solely on observational data, and it can expand their methodological toolbox.

LESSON 2. USE TIME WISELY

Whereas some delays in a project are due to unexpected events, others are caused by the mismanagement of time. One of the most substantial periods of wasted time for this study was during IRB review. For some reason, I felt paralyzed by waiting for approval. Instead of lining up conference rooms or recruiting and training

confederates, I waited. This was not a good use of time. Granted, everything should not be done during the review because the IRB may require changes to the study. For example, printing advertisements would have been unfortunate because the IRB required additional information in them. It also is important to recognize that the IRB response time varies greatly among schools. However, the main point is to assess what can be accomplished during the approval process so that the study can commence more quickly once approval is received.

LESSON 3. IT IS OKAY TO USE DECEPTION

Presenting and discussing this research with fellow political scientists revealed a nontrivial amount of discomfort with deception. Deception, however, is common in psychological studies and may be vital to ensuring appropriate experimental conditions (McDermott 2013). For example, this experiment would not have worked if participants were informed of the study's true purpose—that is, to examine whether they would change their opinion in a room that is stacked against them. If participants had known they would be discussing an issue with three people who disagreed with them, they most certainly would have behaved differently.

Deception is not always necessary but, when it is, it must be respectful of the participants. Especially in a laboratory setting, they should be debriefed at the end of the session and informed of the true purpose of the study. This serves two purposes. First, it brings participants more fully into the process of the research being conducted in a way that respects their important role and respects them as people. Whereas it may be necessary to deceive them for a time, it is beneficial for them to understand what the researcher actually is trying to study and why the deception was needed. In fact, this can be an important learning experience for participants. Many of our participants appreciated learning about their important role in our study. Although pollution of the participant pool is possible with immediate debriefing, we did not find this to be true. In our case, however, we did not have an established subject pool and thus recruited students widely from a large campus. Additionally, there were no other experiments in progress at the time that could have recruited the same students, which is not the case in every program. Some schools have more experimental research being conducted, even enough to have an established subject pool. If this is the case, the researcher must consider whether and how subjects may communicate with one another, thereby polluting the pool. If there are other laboratory

experiments occurring, coordination among investigators will prevent problems across multiple studies.

The second purpose for debriefing is to assess whether the deception is working. When informed of the deception, participants' candid reactions reveal that they were truly surprised. Likewise, a consistent expression of suspicion among participants is an important sign that adjustments in confederate behavior are nec-

incentives should be fixed (i.e., the psychological tradition) or tied to performance (i.e., the economic tradition) is useful (Dickson 2011); however, his does not establish which type of incentives to implement within each category and how much. Previous experiments can be a helpful guide. In cases in which a study is time-intensive or physically intrusive, direct payments or other tangible benefits that otherwise would have a significant cost

Given the import of deception for psychological studies and the growth of psychological research within political science, researchers should not hesitate to employ deception if it is useful for their studies.

essary. In our case, the confederates worked so well together that it made the environment appear artificial to some participants. Knowing this prompted us to make adjustments to maintain consistent experimental conditions throughout the study.

If we want to pursue certain interesting questions, political science must become more comfortable with the use of deception in experimental research. It is necessary, however, to be aware of variations in discipline-specific norms regarding its use. For example, social psychology may be accepting but deception generally is forbidden in the field of economics and its journals. Political science appears to be somewhere in the middle. Given the import of deception for psychological studies and the growth of psychological research within political science, researchers should not hesitate to employ deception if it is useful for their studies. However, ensure that participants are treated with the respect that they deserve. More often than not, they will find it to be a positive experience.

LESSON 4. CHOOSE CONFEDERATES WISELY

Choosing confederates wisely may appear to be common sense, but it is important for both the sanity of researchers and the internal validity of their studies. I was fortunate to have excellent graduate and undergraduate confederates. Depending on the size of a graduate program, recruiting good confederates may not be easy; however, graduate and undergraduate students who have gained respect among the faculty or staff are a good place to start. Regardless of the difficulty, be discerning; this is not a place to cut corners simply to get the project done. Also, make it clear to the confederates how long the experiment is expected to take (see Lesson 1) and ask them to be honest about whether they will still be in the program or at least still in town during the entire span of the study. Confederate dropout weakens the study design—but it does not have to derail its completion.

In addition to choosing good confederates, treat them well. Incentivizing participants is important (see Lesson 5) but, if possible, also set aside funds to pay them. In our case, paying confederates was more important than directly paying each participant. The payment was not large but expressed our appreciation and provided incentive for them to remain committed to the study. Moreover, consider not paying them until completion of the study. This avoids any awkwardness of asking for a return of funds if a confederate must drop out.

LESSON 5. BE CREATIVE IN INCENTIVIZING PARTICIPATION

Incentivizing participation is an important topic but not one on which there is clear guidance. A categorical guideline for whether

(e.g., results of diagnostic scans or a picture of their brain) may be necessary to convince people to participate.

If the funds for direct payments are not available, a raffle may be a successful extrinsic motivator. Our study presented a limited time commitment (i.e., about one hour), but it also offered nothing tangible. In fact, participants rarely signed up solely because they were interested in the purported aim of the study. We lacked sufficient funds to pay confederates and provide meaningful payments to each participant, so a raffle of eight \$25 Amazon gift cards was offered as incentive to participate. An alternative approach is to raffle a single higher-value item (e.g., an iPad), but we chose to offer a higher probability of receiving a lower-value prize. Of course, even a higher probability of payoff is not always enough. One participant arrived for the study and promptly left when he realized it was a raffle and not a direct payment (which was clearly indicated in the advertisement).

There are other creative ways to incentivize participants, particularly students. Extra credit, for instance, could be offered in the researcher's own classes or those of supportive faculty. We found that this did not bias our subject pool toward either high- or low-performing students. When offering extra credit, however, it is important to have an alternative assignment for students who would like extra credit but who do not want to participate in the study; not doing so is coercive. Students should not feel forced to participate in an experiment, especially given the importance of informed and voluntary consent. Of course, extra credit is not helpful if the protocol requires nonstudent participants. In that case, direct payment or a raffle may be necessary.

When designing an incentive, carefully consider the target population and its extrinsic and intrinsic motivations. The presence of incentives alone can induce participant-reaction bias. For instance, some participants may seek to please the investigator, trying to figure out the purpose of the experiment and then giving what is wanted. Other participants may do the opposite in an effort to assert their independence, which also is known as "participant reactance." Both reactions threaten the internal validity of the study. Careful consideration is necessary to balance the need for participants, resources available for incentives, and participant motivations. A pilot study and a thorough and honest debriefing can reveal some of these issues.

LESSON 6. KEEP A JOURNAL

Laboratory notebooks are as important in political science experiments as in the physical sciences. The notes should not be solely about details in how the protocol changed during the study or

the process of data cleaning and analysis. A laboratory notebook is also an opportunity for researchers to record thoughts and interesting observations as they watch the experiment unfold. Note participant behavior and responses, jotting down those “aha moments” that occur during laboratory sessions. Reviewing them later helps to piece together interesting patterns of behavior.

There are other creative ways to incentivize participants, particularly students. Extra credit, for instance, could be offered in the researcher’s own classes or those of supportive faculty.

They also provide a richer description of what happened during the experiment and the meaning of that behavior. We have a large set of wonderful quotes and notations of peculiar behavior about our participants. Some are highlighted in the resulting manuscript but others are kept simply for us—reminders of the experience of conducting that particular study.

LESSON 7. GO FOR IT!

Laboratory experiments are time consuming in design and execution, especially during a first attempt. If this time can be invested and the tool is meaningful for a dissertation project, I advise going for it. First, experiments enrich observational and associational work that can only assume cause and effect. Psychological experiments, in particular, tie broadly observed human behaviors to specific psychological mechanisms. Second, there is increasing use of experimental methods, support for experimentation, and publication of experiments in political science (McDermott 2002). This is evidenced by the new *Journal of Experimental Political Science*; mini-conferences on field and laboratory experiments; the growth of Evidence in Governance and Politics and study preregistration; graduate courses in political science (Hamenstadt 2012); and short-term training programs for experimental methods, including the Inter-University Consortium for Political and Social Research (ICPSR) summer program. Even if graduate students are uncertain whether they will continue experimental research post-dissertation, running their own study provides the tools to do so, and many of the skills that are developed are useful for any type of research. Furthermore, personal experience improves the ability to consume and critique the growing body of experimental political science research. Ultimately, the best way to learn how to conduct an experiment is simply by doing it. It is one thing to learn how to do something in a classroom and another to gain mastery by engaging in the skill.

There is no perfect time to conduct a laboratory experiment in graduate school—or after, for that matter. In our case, summer should have been the best time for quickly completing data collection, but it was a nightmare for coordinating schedules. During the regular semester, coordination remained difficult due to additional commitments including classes, comprehensive exams, and graduate-assistantship responsibilities. However, at least our confederates were in town for extended periods. Thus, persistence was the key to keeping the study moving—and that persistence is necessary regardless of when the study is conducted during the calendar year and graduate studies.

In addition to persistence, “patience is a virtue” in conducting a first laboratory experiment. The experiment is exciting at first, but challenges and delays will drain motivation. Stay committed and remind yourself why you are doing the study. Data collection will conclude and the paper(s) will be written. In the meantime, be patient and flexible.

Finally, although the use of experimental methods may be growing in political science, experience in using them is not uniformly distributed across all universities and colleges. Resources such as ICPSR and methodological texts (Druckman et al. 2011; Kittel, Luhan, and Morton 2012; Morton and Williams 2010) are useful if there is not a deep well of experimental research experience at a particular institution. Reaching out to other experimentalists also is recommended—some are more willing than others, but it is a generally supportive group of scholars. In fact, feel free to reach out to me. ■

REFERENCES

- Butler, Daniel M. 2014. *Representing the Advantaged: How Politicians Reinforce Inequality*. New York: Cambridge University Press.
- Butler, Daniel M., and David E. Broockman. 2011. “Do Politicians Racially Discriminate against Constituents? A Field Experiment on State Legislators.” *American Journal of Political Science* 55 (3): 463–77.
- Carlson, Taylor N., and Jaime E. Settle. 2016. “Political Chameleons: An Exploration of Conformity in Political Discussions.” *Political Behavior* 38 (4): 817–59.
- Dickson, Eric S. 2011. “Economics Vs. Psychology Experiments: Stylization, Incentives, and Deception.” In *Cambridge Handbook of Experimental Political Science*, ed. James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, 58–72. New York: Cambridge University Press.
- Druckman, James N., Donald P. Green, James H. Kuklinski, and Arthur Lupia (eds.). 2011. *Cambridge Handbook of Experimental Political Science*. New York: Cambridge University Press.
- Gosnell, Harold F. 1926. “An Experiment in the Stimulation of Voting.” *American Political Science Review* 20 (4): 869–74.
- Hamenstadt, Ulrich. 2012. “Teaching Experimental Political Science: Experiences from a Seminar on Methods.” *European Political Science* 11 (1): 114–27.
- Keele, Luke. 2015. “The Discipline of Identification.” *PS: Political Science & Politics* 48 (1): 102–106.
- Kittel, Bernhard, Wolfgang J. Luhan, and Rebecca B. Morton (eds.). 2012. *Experimental Political Science: Principles and Practices*. New York: Palgrave MacMillan.
- Levitan, Lindsey C., and Brad Verhulst. 2016. “Conformity in Groups: The Effects of Others’ Views on Expressed Attitudes and Attitude Change.” *Political Behavior* 38 (2): 277–315.
- Mallinson, Daniel J., and Peter K. Hatemi. 2013. “How Political Decisions Are Made: Social Conformity and Political Opinions.” Chicago: Paper presented at the American Political Science Association Annual Meeting.
- McDermott, Rose. 2002. “Experimental Methods in Political Science.” *Annual Review of Political Science* 5 (1): 31–61.
- . 2013. “The Ten Commandments of Experiments.” *PS: Political Science & Politics* 46 (3): 605–10.
- 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.
- Mutz, Diane C. 2011. *Population-Based Survey Experiments*. Princeton, NJ: Princeton University Press.