Reflections on the Importance of Fieldwork for Survey Experiments on Sensitive Topics

Virginia Oliveros
Tulane University

Fieldwork is critical for the successful implementation of field and survey experiments. Good contextual knowledge is key for any sound empirical study, but even more so in the case of experiments because these are design-based research strategies (Dunning 2012)—most of the work and important decisions need to be done before the implementation phase. Once the experiment is conducted, there is little room to fix mistakes or bad choices. Thorough preliminary fieldwork is therefore critical. In my contribution to the symposium, I focus on one particular type of survey experiment, the list experiment—a technique developed to study sensitive topics. I begin by describing my research on patronage in Argentina, which relied heavily on a series of list experiments. I then discuss three key aspects of this research for which deep knowledge of the case, as well as extensive preliminary fieldwork and being in the field while the pilot and the survey were being conducted, were key.

Public Sector Jobs and Political Services: The Machine at Work

In my book, *Patronage at Work: Public Jobs and Political Services in Argentina* (Oliveros 2021), I study the exchange of public sector jobs for political support, or patronage. Even though patronage is a widespread phenomenon, the difficulty in collecting systematic data about it means that we know very little about how patronage works. The book provides a comprehensive description of what patronage employees in low and mid-level positions do in exchange for their jobs, as well as a novel explanation of why they do it. *Patronage at Work* thus aims to understand the specific mechanisms behind the electoral returns to patronage politics.

While patronage is often perfectly legal, it is particularly difficult to study because it constitutes a "gray area" of acceptable practice (Van de Walle 2007, 52). To measure the types and extent of the political services that employees hired through patronage contracts provide to their patrons, I take an approach that allows me to elicit accurate information from public sector employees by minimizing social desirability bias. I use an original face-to-face survey of 1,200 low and mid-level public employees in three Argentinean municipalities (Salta, Santa Fe, and Tigre) that incorporates two strategies to elicit honest responses. The first, following Scacco (2010), consists of employing a number of techniques to earn respondents' trust by guaranteeing the confidentiality of the most sensitive questions. The second is the use of list experiments, a survey technique that protects the privacy of responses by using indirect questioning. Another research tool, the vignette experiment, allows me to assess why public sector employees comply with their side of the patronage agreement. I also conducted multiple interviews. Some of them were part of my preliminary fieldwork; others were conducted later on to illustrate and provide a thicker description of the main findings.

The Sample

As in many other democracies of the Global South, information on public employment in Argentina is not publicly available, and politicians and bureaucrats are reluctant to share it. The first challenge of the research project was therefore to get access to public employment data in order to be able to draw a representative sample for the survey. Preliminary fieldwork was key to achieving this. To get access to this data, I used personal connections to reach several local political authorities and then met with high-level public officials and politicians to explain the purpose of the study, gain their trust, and eventually obtain lists of public employees and receive authorization to conduct the survey. Because I am Argentinean and went to college there, I had some contacts (both academic and political) that proved a good starting point. But even with this "home" advantage, obtaining public employment data in all three municipalities was still daunting and time consuming.

For example, my initial trip to Salta was unsuccessful. My contact in the administration avoided me for a week, stopped replying to my emails, and scheduled in person or telephone appointments at times when he knew he would not be at the office. He eventually informed me that his bosses had requested that I waited until after the upcoming local election to conduct the survey. This meant withholding data for six months. In Tigre, I similarly struggled to get access to the data and the permission to conduct the survey. My contact at the municipality warned me initially that while he could probably guarantee an interview with a gatekeeper

(someone close to the mayor), he was skeptical that they would share the data because it was too sensitive. Indeed, as this contact anticipated, the director of personnel proved very reluctant to assist me with the study and waited until he had written authorization from the mayor to release the data—put differently, he refused to share the data just with a phone call from a high-level official close to the mayor. Finally, in Santa Fe, there were several failed attempts to get an appointment to discuss my project with the relevant officials. A phone call from a former federal congressperson from the mayor's party facilitated access.[1]

In the end, local authorities in all three municipalities met with me, read the survey instrument, and authorized me to access the data and conduct the survey. Knowing what was sensitive in those questionnaires was key to be able to pass this barrier. That knowledge came from my familiarity with the case and the sensitivity of the issues in the Argentine context. For example, to maximize the chances of getting official approval for the survey, I described the survey to local authorities in broad terms as concerning the relationship of public sector employees with local public life (*la relación de los empleados públicos con la vida pública local*). "Local public life" included politics but also other aspects like participating in community meetings and projects, as well as volunteering. My main interest was, of course, politics, but this broader description sounded less "threatening" to the authorities whose main fear seemed to be that I might find some irregularities in public sector appointments (nepotism or too many partisan affiliates) and share that information with journalists. I also took two other precautions. First, I excluded particularly direct, sensitive questions—especially ones related to the mayor.[2] Second, I designed the survey instrument to be as short as possible to ensure employees would not be kept away from their jobs for long periods of time.

Strategies to Ask Questions on Sensitive Issues

Preliminary fieldwork and good knowledge of the case were also important to find ways to deal with the sensitive topics that did make it to the instrument.[3] In order to conduct the survey, enumerators received a random sample of names of public employees and their work addresses, and directly approached them at their workplaces during work hours. Since the focus was on mid- and low-level positions in the administration, places of work ranged from the city hall and decentralized offices, to cemeteries, construction sites, health centers, parks, and the street. Because the survey was conducted face-to-face at this broad array of locations, getting truthful answers presented a challenge. While high-ranking public officials often have private offices, most public employees in Argentina share their workspaces. The issue was that public employees could be unwilling to reveal sensitive information in front of others. How to obtain truthful answers under these conditions? I implemented two distinct but complementary strategies to elicit honest responses and thus minimize social desirability bias.

First, I designed a series of list experiments—a technique that protects the privacy of responses by using indirect questioning (more on this below). Second, I followed Scacco's (2010) strategy (originally developed to study riot participation in Africa) and split the questionnaire into two parts. The first part included background information about the respondent, the less sensitive questions, and the list experiments. The second one included the more sensitive questions about voting behavior, ideology, and political preferences. Each part of the questionnaire was marked with a different survey identification number, which could only be matched with a document not available to the enumerators. Other than this number, the second part of the questionnaire had no information—such as age, gender, or occupation—that could be used to identify the respondent.

Enumerators administered the first part of the questionnaire, while the sensitive part was read and filled out by the respondents themselves. Other public employees who were present at the time of the survey were therefore able to hear neither the questions nor the answers. This part of the questionnaire was purposely designed to be short and easy to understand and answer, with only closed-ended questions. At the end of the interview, respondents were asked to insert this second part of the questionnaire in a sealed cardboard box similar to a ballot box. Enumerators were instructed to provide a detailed explanation of these procedures before handing the sensitive part of the questionnaire to the respondents and to make sure respondents understood that the survey fully protected the confidentiality of their

1 Note that the information I was requesting (a complete list of public employees) is something that is public information in most advanced democracies. This information is not sensitive in itself and does not put the research subjects at risk.

2 For instance, while the survey included a list experiment question about attending rallies, there was no question about the existence of any sort of pressure from the local authorities to attend those rallies. More generally, there was no question about the mayor's role in getting public employees to perform political services.

3 All the details of the preliminary fieldwork and the interviews were reported in the methodological appendix of the book (Oliveros 2021, 207–22). An alternative would have been to include them in a Pre-Analysis Plan, as suggested by Pérez and Tiscornia in their contribution to this symposium.

responses. Their understanding was critical to ensure the success of the data collection strategy.[4]

The specific details about this design strategy were based on preliminary fieldwork and good contextual knowledge. First, an idea of the types of places where interviews would be taking place and crucially the fact that there would be little to no privacy in most of these settings, was key in coming up with my decision to segment the questionnaire. Since the goal was to interview employees in mid-level secretarial and administrative roles and professionals, as well as employees in low-level positions such as street sweepers, janitors, drivers, maintenance workers and security officers, thinking about the places where interviews would be conducted was important. Second, knowledge of the case was also relevant to the decision to use a cardboard box similar to a ballot box. I knew this would be familiar to respondents because paper ballots and cardboard ballot boxes are used in Argentinean elections. This familiarity made the strategy easier to understand. Third, being confident that respondents would be able to fill the sensitive part of the questionnaire by themselves—literacy rates are high in Argentina—was also vital.[5]

Above all, preliminary fieldwork made it clear that some of the questions in the survey were indeed sensitive and that strategies to protect anonymity were therefore necessary. For instance, take my questions about the political services that public sector employees perform on behalf of their patrons. Along with questions on voting behavior and political preferences, these were the toughest to ask. Employees could be unwilling to reveal that kind of information in front of others, but it was also possible that they would be unwilling to reveal the information in private or, even worse, provide inaccurate responses. For these types of questions, I opted to use list experiments. List experiments (and indirect questioning in general) are typically used to improve measurement of behavior or beliefs the respondents would prefer to hide. In the case of the political services studied here, however, it was possible that some employees would actually want to broadcast their contributions and loyalty to the incumbent. But whether an employee would prefer to broadcast or hide his or her political contributions was not random. For instance, most interviews with low-skilled workers took place in front of others, sometimes including their own bosses. If bosses or coworkers were supporters of the incumbent, one could expect the employee to have an incentive to over-report his or her contributions to political services. But bosses or co-workers could also be employees appointed by the previous administration or via meritocratic processes, in which case employees might prefer to hide their political activities. The advantage of list experiments is that they prevent both underreporting and overreporting.

Considering the provision of favors (one of the political services I studied) a sensitive issue might be counterintuitive.[6] After all, providing favors is a way to help others in the community. Preliminary interviews show that in some cases, employees show pride in being helpful. In other cases, however, the sensitivity of the issue was quite evident. A broker and public sector employee from Greater Buenos Aires that I interviewed during my preliminary fieldwork provides a good example of how someone could get slightly offended by the implication that employees provide favors. After a couple of questions about favors, he replied emphatically: "But politics is not a favor machine! (*una máquina de hacer favores*)."[7] Another Peronist broker and public employee from the province of Buenos Aires wanted to make sure not to give the impression that providing favors was a broker's main role: "Peronism is not just about helping people (*no es solamente asistencia*)…Assisting people is just a small part."[8]

The literature on clientelism tends to assume that this is always a sensitive issue (González-Ocantos et al. 2012). By contrast, in the interviews I conducted with political brokers in Argentina it was clear that this was not always a sensitive issue for them, and that some were willing to discuss openly a lot of things researchers consider sensitive. Of course, the framing of the questions matters and asking bluntly if they "buy votes" may not be a good strategy. But, for the most part, brokers are proud of the work they do. They often cite helping those in need as one of their duties, and in the cases of public employees, they do not hide that their jobs were obtained because they were political brokers who could perform that sort of work. What is more: some consider patronage jobs fair compensation for their political contributions. Brokers emphasize that *on top of* doing their job in the public administration as everyone

4 To test the effectiveness of the strategy, I included an additional question about the upcoming presidential election in the questionnaire fielded in one of the municipalities. Half of the respondents were asked this question directly (in the first part of the questionnaire); the other half found this question at the end of the sensitive part which they completed in private. The results confirm my intuition about the importance of affording respondents higher levels of anonymity. Employees responded differently when asked under the protected scheme (see Oliveros 2021).

5 According to the 2010 Argentinean census, only 1.96 percent of the total population older than 10 years old is illiterate.

6 On the provision of favors, see also Oliveros (2016).

7 Author's interview, La Matanza, August 10, 2009.

8 Author's interview, La Plata, August 5, 2009.

else does, they also perform political work. Knowing that brokers were open to discuss their political work was key for conducting successful in-depth interviews—knowledge that could not have been drawn necessarily from the existing literature.

In sum, whether a topic is sensitive or not is often an empirical question and not something that researchers can assume beforehand.[9] Moreover, the sensitivity of the issue varies by research strategy as well—a sensitive issue in a survey may not be that sensitive in an interview setting where the researcher can establish rapport with the interviewee. One can, of course, choose to err on the side of caution, but strategies to deal with sensitive questions are not without cost. For instance, using the split questionnaire strategy described above meant that the survey took longer to complete because enumerators had to spend time explaining the procedure. In the case of list experiments there is also a well-known trade-off between accuracy and efficiency. List experiments reduce response bias by minimizing the incentives for respondents to lie, but they do so at the cost of efficiency.[10] Moreover, for successful implementation, methods of indirect questioning for sensitive questions, such as list experiments, require larger sample sizes than direct questioning (Corstange 2009; Yadav 2015). For these reasons, strategies to deal with sensitive issues should only be used when the issues are indeed sensitive—a key empirical question that the researcher needs to address during preliminary fieldwork.

The List Experiment

The list experiment technique I used to ask about the provision of political services is straightforward. The sample is randomly split into a treatment and a control group. Each group is read the same question and shown a card with a number of response options.[11] List experiments work by including the item one cares about (the sensitive item) in a list containing other items, usually non-sensitive ones. Cards for the two groups differ only in the number of response categories. Respondents are asked to report the number of items on the list that apply to them, but not which ones. Since respondents are randomly assigned to either the group with the sensitive item (treatment) or the one without it (control), the two groups are, on average, indistinguishable on observable and unobservable characteristics. Differences in the mean number of items, or in my case, activities, reported by the two groups therefore provide a point estimate of the proportion of respondents who performed the sensitive activity.[12]

List experiments are, of course, not the only method of indirect questioning to deal with social desirability bias. Two interesting alternatives are the randomized response technique (e.g., Gingerich 2013) and the crosswise model (e.g., Corbacho et al. 2016). I chose to use list experiments over these alternatives mainly for their simplicity. Instructions are easy to understand, and respondents tend to trust that the anonymity of their responses will be protected (Coutts and Jann 2011). Since respondents were low- and mid-level employees, some with low levels of education, this simplicity was an important advantage.

Although the technique is fairly easy to implement and understand, it is still more demanding than direct questioning. Careful survey implementation is crucial for obtaining accurate responses. Preliminary fieldwork and being in the field at the time of the pilot were therefore key. Two examples from my experience illustrate this point. In both cases, being in the field at the time of the pilot, in permanent contact with the enumerators, and conducting many survey interviews myself made me realize two simple issues with list experiments that, at the time of the survey, were not mentioned in the literature on best practices.[13]

First, during the pilot I uncovered two types of error responses by respondents who did not follow or did not understand the instructions. One type occurred when respondents provided a count of the frequency with which they performed each of the activities on the list, instead of counting the items or activities that applied to them. The second type of error was identifying the item or items by using their numbers on the list, causing confusion about whether they were referring to the number of activities that applied to them or to a specific activity on the list (which was not what I wanted). Because of this discovery during the pilot, I decided to switch the numbers to letters, so the cards listed the items by letter (A, B, C) instead of by number. The use of letters instead of numbers to order the list made confusion with the instructions evident to the enumerators, who were instructed to repeat the instructions if respondents showed any lack of understanding. Because the survey

9 This resonates with Bell-Martin's claim in her article in this symposium that "ethnographic evidence facilitates greater construct and ecological validity of our instruments" (p. 2). While I didn't conduct an ethnography, in-depth interviews served a similar purpose.

10 The standard errors for list experiment estimates are larger than they would have been for a direct question with no response bias (Blair and Imai 2012; Corstange 2009).

11 For this project, the list of responses was not read aloud to increase privacy.

12 For other examples of the use of list experiments to measure clientelism and patronage see, for instance, Frye, Reuter, and Szakonyi (2014), González-Ocantos et al. (2012), González-Ocantos and Oliveros (2019), and Mares and Young (2018; 2019).

13 Note that the survey was implemented in 2010 and 2011. Since then, a lot has been written about how to conduct list experiments.

included four list experiments, enumerators had a chance to explain the procedure again if the reaction to the first experiment had alerted them to a misunderstanding. Opting for letters instead of numbers was a free and easy solution that surely increased the accuracy of the responses.

The second issue had to do with "floor effects." To protect anonymity in list experiments it is crucial to avoid lists that could result in respondents choosing none or all of the items, generating "floor" or "ceiling" effects, respectively (see Kuklinski et al. 1997). If a respondent's truthful answer were "yes" or "no" to all the items in the control list, the list experiment would fail to provide the desired deniability on the sensitive item. In other words, respondents would necessarily have to reveal their participation in the sensitive activity when answering sincerely. To minimize ceiling effects, lists usually include rare activities or activities that one cannot perform concurrently. To minimize the risk of floor effects, high-prevalence activities are often included. In my survey, the strategy to minimize ceiling effects was successful and only around one percent of respondents in the control groups for all list experiments reported all four of the control items. The inclusion of high-prevalence activities to minimize the risk of floor effects was less successful. Although I am not aware of any systematic study of this issue, anecdotal evidence from the survey interviews that I conducted suggests that at least some of those zero responses were indeed "DK/NA." List experiments do not include this response option, so when respondents were in a hurry or did not want to answer for any reason, a "zero" response seemed to be the choice. This implies that even in well-designed list experiments in which high prevalence items or activities are included, a number of zero responses may be unavoidable. Although I discovered this issue while in the field, there was little to do about it. Some public employees were indeed in a hurry and chose the zero response. Knowing this, however, was key to my understanding that there was nothing intrinsically wrong with the design of the list experiment. In the end, because the presence of either ceiling or floor effects leads to the underestimation of the sensitive activity (Blair and Imai 2012), this meant that the list experiment estimates were likely conservative.

Concluding Thoughts

The importance of fieldwork and good contextual knowledge for design-based research strategies (Dunning 2012) such as experiments cannot be overstated. Experiments (both field experiments and survey experiments) require that most of the research effort is done before the implementation phase. Once the experiment is in the field, there is little room to turn back the clock on design choices. When the issues under study are sensitive political phenomena—like clientelism or patronage—preliminary fieldwork is even more critical, for both practical and ethical reasons.

From a practical standpoint, failing to acknowledge the sensitivity, or lack thereof, of a particular issue could mean ending up with poor data. In a case in which the researcher does not realize how sensitive an issue is, this could mean that responses are biased, inaccurate, or just plain refusals. But if the researcher choses one of the available techniques to deal with social desirability bias—such as the list experiment or the segmented questionnaire describe above—in a context in which this is not necessary, estimates may end up being less efficient or data more costly to gather. Research strategies designed to deal with sensitive issues should therefore only be used when the issues are indeed sensitive. However, whether an issue is sensitive or not in a particular context is an empirical question. The way to avoid both of these potential problems is to conduct thorough preliminary fieldwork.

From an ethical standpoint, preliminary fieldwork is also vital to assess the sensitivity of the issue and the potential risks for research subjects. Obtaining inaccurate responses due to misreporting or non-response bias is not the worst outcome of a poorly designed research strategy; putting subjects at risk—even if minimal—is. Of course, this is more relevant for subject areas that are more sensitive than patronage. In the end, Argentina is a well-functioning democracy and the "risk" of others finding out about the political preferences or activities of coworkers in the public administration is not serious. Still, it could lead to uncomfortable situations that need to be avoided. Strategies like the ones described above, such as not reading the questions aloud or keeping separate the responses to sensitive questions from the ones that could lead to the identification of an employee, are good examples of effective strategies to protect respondents. And protecting respondents should always be our primary goal.

References

Blair, Graeme, and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20, no. 1 (Winter): 47–77. https://doi.org/10.1093/pan/mpr048

Corbacho, Ana, Daniel W. Gingerich, Virginia Oliveros, and Mauricio Ruiz-Vega. 2016. "Corruption as a Self-Fulfilling Prophecy: Evidence from a Survey Experiment in Costa Rica." *American Journal of Political Science* 60, no. 4 (October): 1077–92. https://doi.org/10.1111/ajps.12244

Corstange, Daniel. 2009. "Sensitive Questions, Truthful Answers? Modeling the List Experiment with LISTIT." *Political Analysis* 17, no. 1 (Winter): 45–63. https://doi.org/10.1093/pan/mpn013

Coutts, Elisabeth, and Ben Jann. 2011. "Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT)." *Sociological Methods & Research* 40, no. 1 (February): 169–93. https://doi.org/10.1177/0049124110390768

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach*. New York: Cambridge University Press.

Frye, Timothy, Ora John Reuter, and David Szakonyi. 2014. "Political Machines at Work Voter Mobilization and Electoral Subversion in the Workplace." *World Politics* 66, no. 2 (April): 195–228. https://doi.org/10.1017/S004388711400001X

Gingerich, Daniel W. 2013. *Political Institutions and Party-Directed Corruption in South America: Stealing for the Team*. New York: Cambridge University Press.

González-Ocantos, Ezequiel, Chad Kiewiet de Jonge, Carlos Meléndez, Javier Osorio, and David W. Nickerson. 2012. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." *American Journal of Political Science* 56, no. 1 (January): 202–17. https://doi.org/10.1111/j.1540-5907.2011.00540.x

González-Ocantos, Ezequiel, and Virginia Oliveros. 2019. "Clientelism in Latin American Politics." In *The Encyclopedia of Latin American Politics*, edited by Gary Prevost and Harry Vanden. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.1677.

Kuklinski, James H., Paul M. Sniderman, Kathleen Knight, Thomas Piazza, Philip E. Tetlock, Gordon R. Lawrence, and Barbara Mellers. 1997. "Racial Prejudice and Attitudes toward Affirmative Action." *American Journal of Political Science* 41, no. 2 (April): 402–19. https://doi.org/10.2307/2111770

Mares, Isabela, and Lauren E. Young. 2018. "The Core Voter's Curse: Clientelistic Threats and Promises in Hungarian Elections." *Comparative Political Studies* 51, no. 11 (September): 1441–71. https://doi.org/10.1177/0010414018758754

———. 2019. *Conditionality & Coercion: Electoral Clientelism in Eastern Europe*. Oxford: Oxford University Press.

Oliveros, Virginia. 2016. "Making It Personal: Clientelism, Favors, and the Personalization of Public Administration in Argentina." *Comparative Politics* 48, no. 3 (April): 373–91. https://doi.org/10.5129/001041516818254437

———. 2021. *Patronage at Work: Public Jobs and Political Services in Argentina*. New York: Cambridge University Press.

Scacco, Alexandra. 2010. "Who Riots? Explaining Individual Participation in Ethnic Violence." PhD diss., Columbia University.

Van de Walle, Nicolas. 2007. "Meet the New Boss, Same as the Old Boss? The Evolution of Political Clientelism in Africa." In *Patrons, Clients and Policies: Patterns of Democratic Accountability and Political Competition*, edited by Herbert Kitschelt and Steven I. Wilkinson, 50–67. New York: Cambridge University Press.

Yadav, Vineeta. 2015. "Lobbying, Corruption, and Non-Responses in Small Samples." *Comparative Politics Newsletter* 25, no. 1 (Spring): 32–37.