Ethics Audits in Cross-National Research: Experiences from Correspondence Study Field Experiments with National Politicians in Four European Democracies

Article details

Pedersen, H. H., Louwerse, T., & Zittel, T. (2022). Ethics Audits in Cross-National Research: Experiences from Correspondence Study Field Experiments with National Politicians in Four European Democracies. Political Studies Review, 20(2), 184–191. https://doi.org/10.1177/14789299211033449

Symposia and New Ideas

Political Studies Association

PSR POLITICAL STUDIES REVIEW

Political Studies Review
2022, Vol. 20(2) 184–191

DOI: 10.1177/14789299211033449
journals.sagepub.com/home/psrev

Ethics Audits in Cross-National Research: Experiences from Correspondence Study Field Experiments with National Politicians in Four European Democracies

Helene Helboe Pedersen[1], Tom Louwerse[2] and Thomas Zittel[3]

Abstract

This essay contemplates experiences from four national ethics audits designed to facilitate correspondence study field experiments with national politicians in Germany, the Netherlands, Denmark and the United Kingdom. The experimental study aims to reveal possible biases in legislators' responsiveness to distinct types of constituents such as non-partisans, lower-class constituents, ethnic minorities, and women, and to unveil possible unsubstantiated fears or misperceptions in this regard. The national research teams proposed the same experimental design but received three different ethical evaluations. Specifically, the relevant Institutional Review Boards (IRBs) in the UK and Denmark asked for two different de-briefing procedures. In the Danish case, this led to withdrawal of the experiment due to severe costs with regard to research quality. In the UK case, it led to increased risk of backlash. Our experiences imply a need for more consistent ethics regimes in the European research community designed to facilitate comparative social science research.

Keywords

correspondence study field experiments, comparative research, experiments with elites, ethics audit

Accepted: 30 June 2021

[1]Department of Political Science, Aarhus University, Aarhus, Denmark
[2]Department of Political Science, Leiden University, Leiden, The Netherlands
[3]Department of Social Sciences, Goethe-University Frankfurt, Frankfurt am Main, Germany

Corresponding author:
Helene Helboe Pedersen, Department of Political Science, Aarhus University, Bartholins Allé 7, 8000 Aarhus C, Denmark.
Email: helene@ps.au.dk

Introduction: The Ethical Implications of Field Experiments with Politicians

Field experiments with politicians have become a more common method in political science, because of the important scientific advantages they offer. Due to the randomized assignment of treatment conditions, field experiments allow for causal identification compared to traditional approaches such as survey or observational research. Compared to laboratory or survey experiments, they can be conducted with high external validity because they do not suffer from low response rates or self-selection bias as non-responses also count as an observation, and artificiality is less of a problem as they take place in a natural context. As such, field experiments provide important insights into the input legitimacy of democracy, such as racial discrimination (Broockman, 2013; Dinesen et al., 2021; Gell-Redman et al., 2018), gender bias (Magni and de Leon, 2020; Thomsen and Sanders, 2020; Wiener, 2020), the electoral prerequisites of geographic representation (Bol et al., 2021; Breunig et al., 2021; Giger et al., 2020), and responsiveness to public-policy preferences (Butler and Nickerson, 2011; Butler et al., 2012). In this vein, field experiments help us to critically assess the quality of democracy and identify needs for reform as well as unsubstantiated fears.

Despite the many benefits, field experiments with politicians entail important costs (Desposato, 2016), not only in terms of deceiving and burdening experimental subjects but also third-party effects such as distorting the political process and provoking backlashes to the profession (Zittel et al., 2021). Field experiments thus require careful ethical consideration.

In our field experiment, we aimed to go beyond the dominant mode of single-country experiments and test legislators' responsiveness to different types of constituents across four political systems, the United Kingdom, Germany, the Netherlands and Denmark. This allows us to test for the robustness of the envisioned social biases in legislators' responsiveness and to understand how institutional contexts structure interactions between citizens and politicians. To achieve these goals, we aimed to keep our experimental design constant across the four countries. Reconciling this goal with a realistic experimental design was demanding, because we had to consider country specificities. For instance, in the UK and Germany, credible citizen inquiries originate from district inhabitants expecting their representative to respond to any concern they might have. This stands in contrast to the Netherlands and Denmark, where citizens are likely to contact the policy specialist relevant to their concern. We anticipated such challenges and worked carefully to design a comparative as well as realistic experiment. But we did not anticipate the second major challenge that we discuss in this paper. In our experiment, we sought local ethics review, that is, approval of the same experimental design by four different institutional review boards (IRB). This challenge severely constrained our research and resulted in drop-out of the Danish case and costly accommodation of the design in the UK case.

In this essay, we critically reflect upon our experiences with obtaining local ethical approval within our comparative design. We delineate the procedural differences we faced across our four cases and identify the issues that resulted in substantially different rulings. With this, we wish to stress the need for more consistent ethics regimes in European contexts that accommodate the need for comparative political science research.

The Context for Ethics Audits of Field Experiments with Politicians in Europe

There is no legal framework in the European context, neither at the national nor at the supra-national level, that requires auditing of field experiments with politicians with regard to their ethics. However, the obvious issues raised by this form of research result in related expectations within different organizational spheres. This concerns, for example, universities that might require ethics audits for any research that involves humans to promote professional standards and prevent backlash. Likewise, and for similar reasons, funding agencies might require ethical statements and approval, for instance the European Union Horizon research programme (European Commission, 2020). Finally, in many scientific journals, ethical review (e.g. American Political Science Review, 2021) is a prerequisite for publishing research with humans.

This special attention to issues of research ethics in general, and the ethics of field experiments in particular, is ill matched by a tapestry of standards that lack authority and consistency. Many national political science associations have drafted ethic guidelines to inform their members. However, they neither constrain the practice of ethics audits, nor do they apply across European countries to facilitate comparative research in political science. Ethical principles have been formulated most clearly and globally for medical research driven by professional medical associations (World Medical Association, 2008), but they do not exist in the social sciences (Israel, 2014).

The American context elucidates the problems with the European ethics regime in the social sciences. In the US, any research involving humans is subject to ethical approval under the National Research Act of 1974, the ensuing Belmont Report issued in 1976, and the resulting Title 45, Part 46 of the Code of Federal Regulations: Protection of Human Subjects. This legal framework forms the authoritative basis for all IRBs and provides a legal context for further specifications issued by professional associations that aim to take issues specific to distinct disciplines into account. For example, the Council of the American Political Science Association (APSA) approved a set of principles for human subject research in the social sciences (APSA, 2020) that carefully balance different considerations in social science research. An important consideration, and particularly important for our study, is informed consent:

> There are some cases in which it might be appropriate for researchers to alter or forgo the consent process. Relevant considerations include when the research is minimal risk, when seeking consent increases the risks for participants, when the research design requires the use of deception or misrepresentation, or when researchers are studying powerful actors and institutions. In each case, researchers should use their best judgment, and explain and justify their decisions in publications and presentations (APSA, 2020: 6).

In this statement, the APSA Council generally advises researchers to seek informed consent among experimental subjects. However, it ties back to the federal legal frame in the US by stressing existing leeway for weighting costs against benefits. It shifts to a discipline-specific perspective and highlights the special role of political science research in auditing institutions and actors of power and resulting needs to forego informed consent to achieve this goal.

Experiences from Local Ethics Audits in Four European Democracies

In our four experiments, we aimed to send two waves of e-mails to members of the UK House of Commons, the Dutch *Tweede Kamer*, the Danish *Folketing* and the German *Bundestag*. The e-mails were designed to originate from fictitious senders varying in terms of ethnicity (by name), gender (by name), class status (by occupation) and partisanship (indicating to be party supporter or not). Their content expressed concerns about the societal consequences of Covid-19 and asked about strategies of the legislator and their party to navigate the country through the crisis. The pre-analysis plan elaborates the research design in greater detail (Baumann et al., 2020).

We took specific note of three ethical concerns. First, we aimed to not distort the representative process by providing misleading or false information and therefore addressed a salient issue in individual and non-positional ways, expressing worries and asking questions rather than stating distinct positions on Covid-19. Second, we aimed to minimize potential costs for participants asking about a general statement on a salient issue, which should be relatively easy to answer compared to a casework request or a request for detailed information on more arcane issues. Third, we aimed to maximize the scientific value of our study, which is crucial for justifying this kind of experimental research (Zittel et al., 2021). To achieve this, we decided to not seek informed consent and thus to engage in activity deception. Seeking consent prior to the experiment would introduce social desirability biases, i.e. only responsive legislators would have been likely to participate. Seeking consent after the experiment would most likely again introduce social desirability biases but also put additional burdens on participants having to confirm their participation and increase the risk of backlash against the profession.

After drafting the design of the planned study and its ethical justifications, the national PIs submitted a request for an ethics audit to the relevant IRBs. Substantially, this request involved the same experiment with the exact same ethical considerations. However, the application formats, the ensuing processes, and the results of these processes varied substantially across the four countries.

Table 1 provides an overview of the ethical approval processes and their results in the four countries.[1] One source of difference concerns the organizational level of the IRB. In Germany and the Netherlands, ethical review was performed at the faculty/school level by IRBs involving only social-science sub-disciplines. In Denmark and the UK, the review was conducted at the university level by IRBs that consisted of representatives of multiple disciplines. In the UK, the IRB is divided into sub-boards including one on social sciences, humanities and law, which was responsible for processing and deciding on the British application. The sub-committee consists of 10 members, three of which are from social science (war studies, education and business, law and political economy).[2] In Denmark, there are no sub-committees. The IRB consists of 10 members and a chair, three of which are from social science (law, business management, political science). The political scientist was disqualified from participating in discussions on projects proposed by political scientists, so in practice no discipline experts were involved.

A second source of difference concerns the timing of the ethical audit. In Germany and the Netherlands, PIs applied for ethical approval prior to the official reward of funding, as required by the funding agency. In Denmark and the UK, ethical approval was applied for immediately prior to the fieldwork. This means that IRBs in Germany and the Netherlands reviewed the overall research design of sending 2–3 waves of fictitious e-mails to national

Table 1. Overview of Ethical Review Processes.

	Germany	Netherlands	UK	Denmark
Level of IRB	Faculty level	Faculty level	University level	University level
Timing of IRB audit	Prior to official reward of funding	Prior to official reward of funding	Prior to fieldwork	Prior to fieldwork
Outcome of first application	Approved after additional questions re possible legal transgressions were answered	Approved after additional questions re anonymity of respondents in the data policy were answered	Disapproved. Refer to principles of informed consent and deception and demand a debrief with opt-out option	Disapproved. Refer to principles of informed consent and demand pre- or debriefing
Outcome of second application	Not relevant	Not relevant	Approved, ask for debriefing letter to reference GDPR-rules and more details on data management	Disapproved. Not satisfied with debriefing including opt-out opportunity. Requires debriefing with opt-in requirement and no access to data prior to opt-in
Field phase of experimental study	Nov 2 to Dec 11 2020	Nov 2 to Dec 11 2020	Nov 2 to Dec 18 2020	Not conducted
Backlash	None	None	Concern among MPs regarding use of their time expressed in the news	Not relevant

IRB: Institutional Review Board.
In the Netherlands, three institutes in the Faculty of Social and Behavioural Sciences are linked to the Ethics Review Committee Social Sciences (Political Science, Cultural Anthropology and the Centre for Science and Technology Studies). In Germany, the institute of political science and the institute of sociology are linked to the faculty review committee.

MPs, whereas IRBs in Denmark and the UK also reviewed the exact stimulus material. IRBs in the UK and Denmark did not question the exact stimuli but only the overall research design and mainly the issue of informed consent. Timing and details, thus, do not seem to be decisive for the result of the audit processes.

As Table 1 shows, the outcomes of the audits differed substantively. The German and Dutch applications were approved after one review round where PIs in the Netherlands were asked to respond to written queries on data protection and the PIs in Germany to queries on legal implications. The queries were answered in writing and to the satisfaction of the IRBs. In contrast, the Danish and British applications were not approved in the first round. In both cases, IRBs required the PIs to seek consent among experimental subjects either prior to or after the experiment. The Danish IRB found that debriefing with the possibility to withdraw from the study was an invariable principle for any study including deception. National PIs decided to accommodate the concerns of the IRBs by offering a de-briefing with a drop-out option. This increased the risk of social desirability biases and backlash and caused unwanted variation in design across countries, but it seemed imperative to get ethical approval in the Danish and UK cases and to realize the envisioned comparative study.

In the second round, the Danish and British PIs submitted the exact same de-briefing letters, which included an explanation of the purpose of the research, references to the PI with contact details, an invitation to ask questions, and an option to drop out of the study. With this adjustment, the British application was approved with minor comments, including that the debriefing letter should reference GDPR-rules, and the PI should prepare a formal data-sharing agreement. In contrast, the second Danish application was not approved. The IRB states that it was willing to review the project again if the debrief included an opt-in rather than an opt-out option, and that the PI committed to not process any data prior to obtaining consent. The IRB did not refer to Danish law or internal guidelines in their answer. Based on this requirement, the Danish PI decided not to run the experiment in Denmark. Compliance with the requirements would significantly lower the scientific value of the research to the point that it would no longer justify the intervention, because the opt-in requirement would exacerbate social desirability biases and self-selection biases. Even worse, without the possibility to access data to conduct dropout analyses, it would be impossible to explore and understand such biases, making them even more damaging.

The three remaining experiments were fielded between November 2020 and December 2020. In the UK, the required de-briefing round led to public and academic debates (see Campbell and Bolet in this issue). Importantly, King's College London offered the exposed PI significant and unequivocal support throughout the process.

Reflections on Our Experiences

Our experiences with local ethics audits in four European democracies illustrate the challenges to comparative field experiments with politicians caused by highly decentralized regimes and ensuing inconsistent decisions. The decentral nature of the process that we faced particularly affected our comparative study in two ways. First, it led to unwanted variance in design across the national experiments that negatively affects the comparability of the country specific studies. Second, it diminished the number of our cases and thus compromised our case-selection strategy.

We consider this outcome most unfortunate since cross-national comparative research is crucial to uncover how institutional context mediates interactions between politicians and citizens. Cross-national comparative field experiments also help to test the robustness of

distinct patterns in the interactions between politicians and citizens and thus to evaluate the generalizability of theoretical claims and case-specific evidence regarding elite behaviour.

One solution to the problem is that PIs apply to a single national IRB to carry out field experiments across countries. In our case, this was not possible due to funding structures. But even if it were possible, this solution is potentially problematic, as the IRB may not be sufficiently aware of relevant national-level considerations. We therefore ask how ethic regimes could be altered to promote comparative research and conclude with two considerations in this regard.

The first consideration concerns the set-up of local IRBs and their disciplinary homogeneity. Our experiences are too slim to make firm conclusions as to why the same experimental study was evaluated differently across IRBs. However, it appears that the closer the IRB is to the discipline of the proposed study, the more likely it is that the study is approved. This is obviously not necessarily a desirable outcome. Disciplinary homogeneity might contradict serious and impartial review. However, it also allows for reviews that can take the disciplinary relevance of the proposed research into account and are familiar with the existing practices in the field. The principles invoked by the cited APSA statement point to the importance of considering such disciplinary issues. In political science, this concerns a professional ethos to provide a critical corrective to institutions of power, i.e. government. This ethos needs to be taken into account when we weigh the costs and benefits of field experiments with politicians involving deception. We therefore encourage discipline-specific audit processes.

The second consideration concerns the legal context in which local IRBs operate. The lack of a legal framework similar to what we find in the US case provides substantial leeway for local IRBs to conduct ethics audits as they see fit. This creates unnecessary and unfortunate consequences for comparative research. We therefore support cooperation between national political science associations, efforts in international associations (e.g. ECPR) and particularly in political realms (EU) to continue working towards a consistent European framework in research ethics.[3]

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by grants from the Deutsche Forschungsgemeinschaft (Grant Number: ZI 608/8-1) and the Nederlandse Organisatie voor Wetenschappelijk (Grant Number: 464.18.108).

ORCID iDs

Helene Helboe Pedersen https://orcid.org/0000-0002-2630-4518
Tom Louwerse https://orcid.org/0000-0003-4131-2724

Notes

1. We refer to countries when describing the differences, but the actual reviews were made locally and not by any national boards. Therefore practices may vary across institutions within the four countries
2. Members are not listed by names. The committee includes a lay member and a student representative with unknown disciplines.
3. European countries not members of the EU such as Switzerland or the UK could be incorporated via bilateral agreements.

References

American Political Science Review (2021) Notes from the Editors. *American Political Science Review* 115 (1): v–viii. Available at: https://www.cambridge.org/core/journals/american-political-science-review/article/notes-from-the-editors/9B735E23BC3D2ADC08027A2CCADA46F3 (accessed May 19 2021).

APSA (2020) Principles and Guidance for Human Subject Research. Available at: https://www.apsanet.org/Portals/54/diversity%20and%20inclusion%20prgms/Ethics/Final_Principles%20with%20Guidance%20with%20intro.pdf?ver=2020-04-20-211740-153 (accessed 19 May 2021).

Baumann M, Bolet D, Campbell R, et al. (2020) The Responsiveness of MPs to Citizen-Initiated Policy Queries: Pre-analysis Plan. Available at: https://osf.io/b7rz9

Bol D, Gschwend T, Zittel T, et al. (2021) The Importance of Personal Vote Intentions for the Responsiveness of Legislators: A Field Experiment. *European Journal of Political Research* 60 (2): 455–473.

Breunig C, Grossman E and Hanni M (2021) Responsiveness and Democratic Accountability: Observational Evidence from an Experiment in a Mixed-Member Proportional System. Epub ahead of print 30 December. *Legislative Studies Quarterly*. DOI: 10.1111/lsq.12326.

Broockman DE (2013) Black Politicians Are More Intrinsically Motivated to Advance Blacks' Interests: A Field Experiment Manipulating Political Incentives. *American Journal of Political Science* 57 (3): 521–536.

Butler DM and Nickerson DW (2011) Can Learning Constituency Opinion Affect How Legislators Vote? Results from a Field Experiment. *Quarterly Journal of Political Science* 6 (1): 55–83.

Butler DM, Karpowitz C and Pope JC (2012) Field Experiment on Legislators' Home Style: Service versus Policy. *Journal of Politics* 74 (2): 474–486.

Desposato S (2016) *Ethics and Experiments. Problems and Solutions for Social Scientists and Policy Professionals*. New York; London: Routledge.

Dinesen PT, Dahl M and Schiøler M (2021) When Are Legislators Responsive to Ethnic Minorities? Testing the Role of Electoral Incentives and Candidate Selection for Mitigating Ethnocentric Responsiveness. *American Political Science Review* 115 (2): 450–466.

European Commission (2020) Horizon 2020 Online Manual. Available at: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/ethics_en.htm (accessed 19 May 2021).

Gell-Redman M, Visalvanich N, Crabtree C, et al. (2018) It's All About Race: How State Legislators Respond to Immigrant Constituents. *Political Research Quarterly* 71 (3): 517–531.

Giger N, Lanz S and De Vries C (2020) The Motivational Basis of Constituency Work: How Intrinsic and Extrinsic Motivations Interact. *Political Science Research and Methods* 8 (3): 493–508.

Israel M (2014) *Research Ethics for Social Scientists*. London: SAGE Publishing.

Magni G and de Leon ZP (2020) Women Want an Answer! Field Experiments on Elected Officials and Gender Bias. *Journal of Experimental Political Science*. Epub ahead of print 10 August. DOI: 10.1017/XPS.2020.23.

Thomsen DM and Sanders BK (2020) Gender Differences in Legislator Responsiveness. *Perspectives on Politics* 18: 1017–1030.

Wiener E (2020) Getting a High Heel in the Door: An Experiment on State Legislator Responsiveness to Women's Issue Lobbying. *Political Research Quarterly*. Epub ahead of print 9 July. DOI: 10.1177/1065912920939186.

World Medical Association (2008) *Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects*. Ferney-Voltaire: World Medical Association.

Zittel T, Louwerse T, Pedersen HH, et al. (2021) Should We Conduct Correspondence Study Field Experiments with Political Elites? *International Political Science Review*. Epub ahead of print 15 July. DOI: 10.1177/01925121211026489

Author Biographies

Helene Helboe Pedersen is professor at the Department of Political Science, Aarhus University (Denmark). Her research interests include comparative and Danish politics, political parties and interest groups, and parliaments and government.

Tom Louwerse is associate professor at the Institute of Political Science, Leiden University (The Netherlands). His research interests include political representation, parliamentary politics and elections.

Thomas Zittel is professor for comparative politics at the Department of Social Sciences, Goethe-University Frankfurt (Germany). His research focuses on issues of political representation with special emphases on legislative politics, electoral systems, campaign behavior, and personalization.