

RESEARCH ARTICLE  

Using Citizen Voice to Evaluate Experiments on Politicians: A UK Survey Experiment

Peter John^{1*} , Kristina Kim² and Luis Soto-Tamayo¹ 

¹King's College London, London, UK and ²University of Edinburgh, Edinburgh, UK

*Corresponding author. Email: peter.john@kcl.ac.uk

Abstract

Experiments on the responsiveness of elected officials highlight the tension between the freedom to carry out research and the right of subjects to be treated with respect. Controversy emerges from the power of politicians to block or object to experimental designs using identity deception. One way to resolve this conundrum is to consult citizens who, as constituents of politicians, have an interest in promoting the accountability of elected representatives. Building on the work of Desposato and Naurin and Öhberg, this survey experiment presented research designs to UK citizens for their evaluation. The findings show that citizens strongly approve of experimental research on Members of Parliament (MPs) and are glad to see their representatives participate. There are no differences in support whether designs use identity deception, debriefing, confederates or pre-agreement from MPs. Linked to high interest in politics, more citizens are glad their MPs participate in studies using identity deception than those deploying confederates.

Keywords: Elite experiments; politicians; research ethics; citizen attitudes; survey experiment

The ethical conduct of political science research is increasingly contested, with researchers, professional associations and regulators having to chart a perilous course between respect for the autonomy and dignity of the subject and the need to carry out independent and innovative studies. Experiments on elites, such as elected officials, pose additional challenges. There should be no distinction between the conduct of these experiments and those with other research subjects, but they differ because politicians are powerful. Politicians can communicate their disapproval to funding bodies and universities, potentially blocking or deterring such research. In response, it has been countered that the power of elites should not be a factor in deciding whether to carry out these kinds of studies (McClendon 2012, 17). Researchers should be able to speak truth to power. But it is hard to deny

  This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the [Data Availability Statement](#).

© The Author(s), 2023. Published by Cambridge University Press on behalf of American Political Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

that experiments may be seen as intrusive, especially when they involve identity deception. There is also the need to ensure that elites maintain their trust in research from the academy, which can be undermined by studies that appear to deceive public officials. Partly because of these debates, researchers have moved away from researcher-determined designs to more collaborative forms of endeavour (Butler 2019; Loewen and Rubenson 2022). In such partnerships, politicians and other elites can find out about the advantages of research done on themselves, which can assist in their quest for accountability.

Such issues emerged discernibly over the research project led by Rosie Campbell in 2021 (Campbell and Bolet 2022). This study randomly assigned e-mails from different kinds of fictitious constituents to UK Members of Parliament (MPs), using a method pioneered in the USA by Butler, Broockman and others (Butler and Broockman 2011; Butler and Nickerson 2011; Butler 2014), and widely applied in many other jurisdictions (McClendon 2016; Vries, Dinas, and Solaz 2016; Habel and Birch 2019; Crawford and Ramli 2021). The difference to other studies was the requirement by the university's Research Ethics Committee to carry out a debriefing after the experiment had taken place. Once MPs found out about the research from the debrief, some bitterly complained about being deceived, also saying that it was a waste of their staff's time. There was a media and Twitter storm (see: <https://www.bbc.co.uk/news/uk-politics-56196967>). The Speaker of the House of Commons wrote to the Principal of King's College London, where Campbell is based, and to the executive chair of the Economic and Social Research Council, which funded the study. The question this case poses is how best to achieve the balance between respect for the subject and freedom to carry out research.

As well as politicians, regulators and researchers, there is another important stakeholder whose views might help resolve the conundrum of agreeing ethical designs on elected officials: citizens. Addressing the views of the wider public has become a strong theme in research on human subjects in the biosciences in what is called 'empirical ethics' (see Borry, Schotsmans, and Dierickx 2008 and Appendix A). It also matters what citizens think of experiments designed to improve knowledge about elected representatives, given that MPs are accountable to them. As researchers need to make choices about how to carry out elite experiments, the public's view of the use of identity deception counts, especially when compared to other designs, such as the use of confederates, that is recruiting real constituents who are asked to contact MPs. The public might view the use of deception as not adhering to standards of fairness and fair dealing. On the other hand, the public has become critical of politicians and their motives in recent years (Clarke et al. 2018). Beliefs in 'anti-politics' and distrust of politicians might encourage citizens to approve of stronger review and audit.

Researchers have already started along this path with Desposato's (2018) study of the North American public's view of correspondence audit experiments, especially when they involve deception, and Naurin and Öhberg's (2021) comparison of the views of Swedish voters and officials. Desposato (2018) surveyed US citizens and scholars about their perspectives on the use of deception (i.e. ex ante consent, identity deception and no consent, and identity deception and debrief to gain post-hoc consent) as well according to the target object group (i.e. politicians, business owners or private homeowners) and purpose of research (i.e. assessing

discriminatory behaviour, communications, customer service or constituency service). When asked about the vignette's acceptability, both types of respondents expressed negative reactions to experiments carried out without consent and to all forms of deception. Naurin and Öhberg (2021) performed a similar survey experiment in Sweden, investigating politician and citizen perceptions on research ethics and experimentation. Respondents were asked, 'To what extent would you find the following things to be ethically problematic if you were asked to participate in a survey addressed to you/to you in your capacity as politician?' and asked to rate certain research practices within a hypothetical survey on a scale of 1 – Yes, very problematic, to 7 – no completely unproblematic. Politicians tended to rate each prompt as more ethically problematic than citizens, which challenges the common assumption of politicians being 'less sensitive to experimental designs than ordinary citizens because they are used to being scrutinized by the media, voters, opponents, and others' (Naurin and Öhberg, 2021, pp. 890–891).

The research for this paper extends this line of research to the UK context, evaluating Campbell's design and possible alternatives. We presented randomly assigned scenarios to a representative sample of UK residents, varying the use of identity deception, debriefing and the recruitment of confederates. Following Desposato, we hypothesised that the public would be more approving of scenarios involving the recruitment of real constituents than those where identity deception is used (H1). We also wanted to capture the idea – implied by Desposato's research – that the more overt the deception, the greater the public's disapproval. We thus concluded that the public would find research designs carried out by a researcher from a university more acceptable than an investigation by a journalist using deception (H2). We considered that there was less justification for other forms of deception on politicians than from an independent/official researcher, which have a serious purpose and where safeguards are in place. Also, collaboration with MPs would be seen as preferable to deception, but not approved as much as using confederates (H3). Causing upset among politicians after a debrief would be seen as less acceptable than when there is just identity deception and a debrief (H4). Not debriefing would have less acceptability than deception with debriefing (H5) and the use of confederates with debriefing (H6). There were also exploratory hypotheses based on the anti-politics literature: working class and Conservative voters would be more supportive of measures to hold MPs accountable. The wording of these hypotheses, pre-registered at the Open Science Framework (OSF),¹ is reproduced in Appendix B.

Research methods

We developed seven scenarios to test the hypotheses (see Table 1 for the summary of the scenarios and Appendix C for their full wording). One of these (A3) was designed to replicate Campbell's research design with the debrief and mentioning the furore of the MPs, contrasting with other designs just using identity deception (A1) and one adding the debrief (A2). There were two scenarios with confederates,

¹John, Peter. 2021. 'King's College London Ethical Review of Research. <https://doi.org/10.17605/OSF.IO/X7HT6>. OSF. December 18.

Table 1
Summary of Scenarios

Scenario A1: a researcher sends e-mails to MPs from fictitious constituents
Scenario A2: a researcher sends e-mails to MPs from fictitious constituents with a debrief
Scenario A3: a researcher sends e-mails to MPs from fictitious constituents with a debrief causing upset among MPs
Scenario B1: a researcher recruits real constituents who send e-mails to MPs
Scenario B2: a researcher recruits real constituents who send e-mails to MPs with debrief
Scenario C: MPs pre-agree to a researcher sending e-mails from fictitious constituents
Scenario D: a journalist pretends to be constituent to obtain a story

one without (B1) and the other with the debrief (B2). Scenario C was a collaboration with MPs also with identity deception. To test Hypothesis 2, we included a different kind of scenario: a journalist posing as a constituent to obtain a story from a MP when off-guard. It was loosely based on a real case of the senior Liberal Democrat MP, Vince Cable, who was taped by two journalists posing as constituents (*The Daily Telegraph*, 22 December 2010). The scenario provided a baseline from which to compare the others (without being a placebo).²

We carried out cognitive interviews (Miller et al. 2014) to test the wording of the scenarios and to ensure that respondents understood them (see Appendix D). As a result, the scenarios were redrafted with more ordinary-sounding language.

Outcome variables are the extent to which the research is regarded as acceptable, whether the respondent would like their own MP in the research, and the extent to which the research is seen to be socially valuable. Covariates, which were either supplied by the survey company or came from the survey itself, are gender, age, ethnicity/race, education, religiosity, region, social grade, vote in 2019 and 2021 and income. We asked standard attitudinal questions on political interest and efficacy, which were used as additional independent variables (see codebook in Appendix F).

The survey was carried out by Deltapoll, a quota sample drawn from its panel of just over 750,000 UK adults. The quotas were age, gender, region, past vote and EU referendum vote. The survey launched on 13 December 2021, yielding 8,040 respondents (for data and supporting files, see John et al 2022). A summary table of the demographic variables is contained in Appendix G. The sample is representative on the main demographic variables but is slightly short of Conservative-supporting voters. There is a weight that can deal with the lack of filling the quota, but the tables in the paper are unweighted (key weighted tables are produced in Appendix H). Appendix I reports balance tests, which show equivalence across the scenarios and no more significant terms than would have occurred by chance.

Appendix J reports the manipulation checks. Respondents were able to detect differences between the scenarios, with the percentage correct within each treated group ranging from 34.0 to 47.9%. Some respondents wrongly attributed aspects of

²This scenario did not specify whether the person was working for a 'serious' broadsheet newspaper as an investigative journalist or for a more sensationalist 'tabloid' press which could have affected responses if one of these formats were in respondents' minds, an example of the information equivalence problem (Dafoe, Zhang, and Caughey 2018). Appendix E, however, shows that most respondents did not make a hard and fast distinction between types of print media outlet believing the example was of investigative journalism done inappropriately.

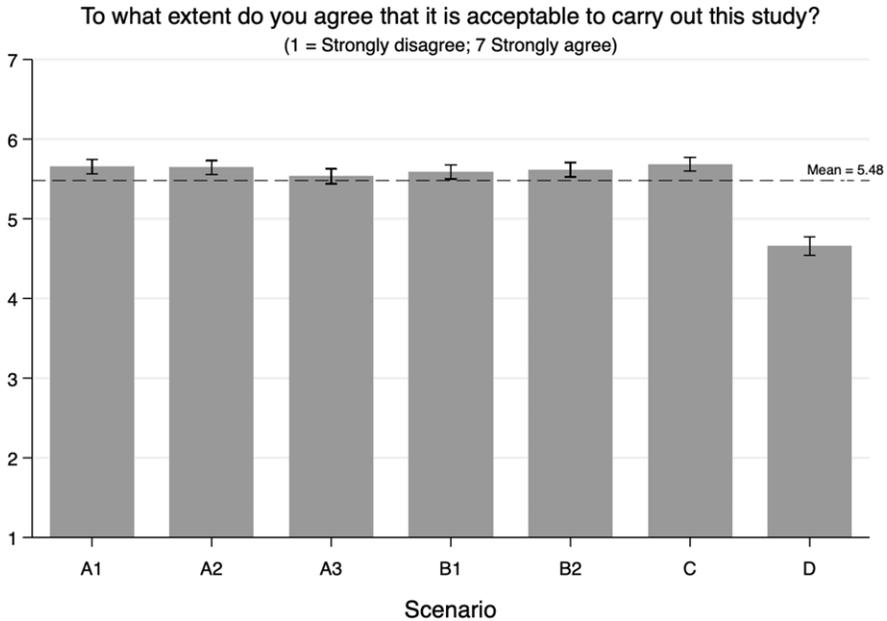


Figure 1
Citizen Views of the Acceptability of Research on Politicians.

these scenarios even when they were not presented with these features, varying from 7.8 to 18.3%. In contrast, 61.9% recognised the journalist scenario, with 13.3% of the rest of the sample who were not shown this scenario believing the journalist was part of theirs.³

Responses to the qualitative analysis (see Appendix K) give added support to the internal validity of the experiment. Respondents were asked: ‘In more than 10 words tell us how you feel about this study?’ and were given a textbox to write their answers. After developing a code-frame, we coded a random sample of 2,000 responses into 15 categories, which showed respondents engaging with the scenarios. For example, 18.3% were coded to the category ‘Honest response from MP/Holding MPs accountable’, 6.1% to ‘Negative response to deception’ and 4.0% to ‘Interested in potential results of study’. The categories that showed no engagement only took up a small proportion of responses: 3.2% ‘Don’t understand/confused’ and 5.8% ‘Other’ (see Table K2).

Results

Figure 1 displays the levels of approval for the different scenarios, with 95% confidence intervals.

³Appendix N, Table N2 presents models with only those respondents who passed these manipulation checks, which do not affect conclusions.

Table 2
Feelings About Their MP Being Included in the Study

	Scenario shown							Total
	(A1)	(A2)	(A3)	(B1)	(B2)	(C)	(D)	
I would be glad my MP was in the study	710	721	685	659	690	706	532	4703
	61.90	62.80	59.72	57.45	59.95	61.13	46.46	58.50
I would rather my MP was not in the study	122	114	121	96	114	129	185	881
	10.64	9.93	10.55	8.37	9.90	11.17	16.16	10.96
I would not care either way	261	264	269	318	282	256	340	1990
	22.76	23.00	23.45	27.72	24.50	22.16	29.69	24.75
Don't know	54	49	72	74	65	64	88	466
	4.71	4.27	6.28	6.45	5.65	5.54	7.69	5.80
Total	1147	1148	1147	1147	1151	1155	1145	8040
	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

The first row presents frequencies. The second row contains column percentages.

With responses to the scenarios averaging at 5.5 on the seven-point scale, we infer that the UK public approves of studies on MPs. There are no significant differences of views of the acceptability between the experimental scenarios, showing that the public neither distinguish between those based on identity deception, confederates and collaboration nor between studies that use debriefing and those that do not. The only scenario showing a strong difference is D, the journalist story. Here, the public were less approving, with a mean of 4.66 on the seven-point scale, 0.97 points ($p < 0.0001$) below the combined mean of 5.62 of the other groups (see Appendix L). We do not confirm any of the hypotheses except H2. Regression analysis shows that support for these measures is driven by political interest and efficacy, not demographics (see Appendix M).

Table 2 shows the results for the question, 'Suppose you learned that a study like the one described before had been carried out in your community with your MP. Which of the following best describes how you would feel about the MP being included in the study?'

Between 46.4 and 62.8% of respondents would be glad their MP was included in these scenarios, which is quite high given ethical concerns about deception. Between 8.4 and 16.1% would rather their MP did not participate, with between 22.2 and 29.7% not caring either way. Differences between most of the experimental research scenarios are not significant (between A1 and A2, A2 and A3, B1 and B2, and between C and the others). As before, the big difference is between Scenario D and the rest, with 46.4% glad their MP is in the study, compared to 60.4 average for the other scenarios ($p > 0.00001$). There is, however, a statistically significant 2.7 percentage point difference ($p > 0.05$) between respondents being glad their

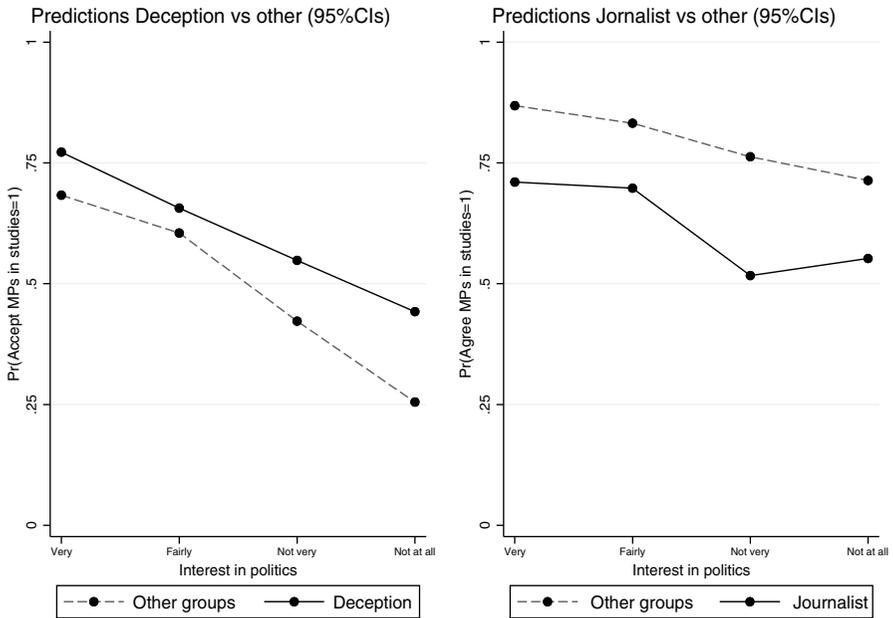


Figure 2
Impact of Political Interest on Feelings About Their MP Being Included in the Study.

MP is in scenarios that use identity deception (A1, A2, A3 and C) and those that do not, the confederate studies (B1, B2).⁴ Rather than confirming our expectation that the public would display greater disapproval of scenarios using deception, the relationship is in the opposite direction.

In Appendix N, we explore why respondents respond more positively to the deception conditions. Although we did not propose this estimation in the pre-analysis plan, it is in the spirit of our exploratory Hypothesis 7: lower educated groups will show greater approval of Scenario A1 and Scenario A2 than other groups, with Conservative voters being more willing to appreciate interventions using deception. Overall, the interaction analysis does not provide evidence that an anti-politics mentality among citizens might be a cause of liking more robust measures of holding politicians to account, which include identity deception. Rather, those who are engaged with politics are more likely to be content their MPs are in these studies. The visual presentations of the marginal plots of the interaction between political interest and the use of deception remain the same with no interaction at high levels of interest and interaction at low levels: see Figure 2 (left panel). There is a similar outcome in seeking to evaluate the difference between

⁴In this paper, we do not consider correction for multiple comparisons since we only test one outcome variable (acceptability) and one treatment (deception vs other groups) across six scenarios that theoretically make sense to compare (instead of comparing all treatment groups with scenarios that include all covariates).

Scenario D and the rest of the treatments on the main outcome variable of acceptability. Figure 2 (right panel) shows that the journalists' scenario is less approved than others, and that this approval is higher within those with more interest in politics (see also regression in Appendix O). In addition, being a co-partisan with the respondent's MP might have caused respondents to be more protective of their MPs being in these studies, but the term is not significant and has the opposite sign (Appendix N, Table N3).

The final outcome measure is derived from the question, 'To what extent do you think that this study is worthwhile to carry out?', with responses coded from 7 highly worthwhile to not at all worthwhile 1. The mean response is 3.4, which shows respondents are at the mid-point. As before, there were no significant differences between the scenarios, except between D and the others, which scored a mean of 3.82 ($p < 0.0001$). We also asked, 'To what extent do you think the study shows how a MP would be likely to answer real e-mails from local residents?' with responses coded from 7 very likely to not at all likely 1, recoding a mean of 4.2 indicating likelihood. There are no differences in responses to this question, even for D, which scores 3.35 ($p > 0.1$).

Discussion

We add to knowledge on public attitudes to field experiments on elected officials, being the third successive survey experiment on this topic. Because of different locations, timings and designs, we cannot make systematic comparisons of results across the three studies. But we can offer interpretations where they share common features in their vignettes. Also, they deploy a similar dependent variable on the acceptability of the study with a seven-point agree–disagree scale.

Naurin and Öhberg present a scenario of an elite experiment using identity deception, similar to Scenario A1, scoring 5.08, which is lower than our 5.64 result, but is still in the same ballpark of overall public acceptability. Desposato's study design is also comparable to our study as we adapted its question wording, even though our vignettes changed because of the cognitive interviewing. Desposato finds that the public give a similar scenario to A1 an average of 4.7, just under a point below the UK results, even if still finding public acceptability overall. Unlike us, he shows that consent improves the score to 5.4, closer to our estimates for Scenario B1 and Scenario B2 (the closest comparison). Like us, he finds that the offer of a debriefing makes no impact on public attitudes. Overall, the UK study shows higher public approval of correspondence experiments on politicians than has been found in other jurisdictions.

Timing might be a factor explaining support for audits in the UK as the survey took place in the middle of a crisis in Prime Minister Johnson's government. Revealed in a series of scandals in newspapers, drinks parties were frequent occurrences in 10 Downing Street and elsewhere in Whitehall, showing the double standard of politicians and officials who disobeyed the very laws and regulations they had enforced on the public. This could have increased the willingness of respondents to subject politicians to greater scrutiny.

Conclusion

We show strong public support for experimental designs on elected officials. In the quest for balance between research aiming for knowledge and freedom of researchers to decide this and respect for consent, this study finds support for the independence of research and approval of more audit and accountability of politicians. UK citizens do not discriminate whether studies use identity deception, have a debriefing, deploy confederates or get pre-agreement from MPs: the high level of approval is the same. Citizens are also glad that their MPs participate, with greater willingness in studies that use identity deception than those that deploy confederates. It is also positive that higher political interest and efficacy are correlated with these assessments, which show that liking for research evaluation on accountability is shared by those who are engaged with politics, seeing it as a benefit. Such attitudes are not part of a backlash from those who have low interest in politics. With these findings, researchers can now be more confident using experimental studies on politicians whilst still wanting to observe the highest standards of ethical conduct.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/XPS.2023.3>

Data availability statement. Support for this research was provided by King's College London. The data, code and any additional materials required to replicate all analyses in this article are available at the *Journal of Experimental Political Science* Dataverse within the Harvard Dataverse Network, at: <https://doi.org/10.7910/DVN/UDFLIA>.

Acknowledgements. We thank King's College London for funding. We are grateful to members of the King's College London Ethical Review of Research – Irene Birrell, Paul Mould, Christina Boswell and Chris Skidmore MP – who commissioned the study (Peter John was also a member). We thank Bobby Duffy, Roger Mortimore, Kirstie Hewlett, Paul Stoneman and Scott Desposato for advice on the study design, and Florian Foos and Rosie Campbell for their comments on an early version of the manuscript. We appreciate those who heard the paper at the panel, 'Experiments with politicians: ethics, power, and the boundaries of political science', which was held at the UK Political Studies Association annual conference, University of York, 13–15 April 2022. We are grateful to Joe Twyman from Deltapoll for advice on the questionnaire and survey, especially for supplying a new variable on MPs' constituencies long after the survey had been fielded. We greatly appreciate LSE students Adam Keyworth and Johannes Rosenbusch for research assistance. Finally, we thank the citizens who gave their time for the cognitive interviews.

Conflicts of interest. The authors declare no competing interests.

Ethics statement. The cognitive interviews were approved by King's College London Research Ethics Committee, 17 November 2021 (MRA-21/22-26957). The survey experiment design was approved by King's College London Research Ethics Committee on 11 December 2021 (MRA-21/22-26956). We affirm that research adheres to APSA's Principles and Guidance for Human Subjects Research. A checklist of the reporting standards recommended by the APSA Organized Section on Experimental Research is presented in Appendix P.

References

- Borry, Pascal, Paul Schotsmans and Kris Dierickx.** 2008. The Origin and Emergence of Empirical Ethics. In *Empirical Ethics in Psychiatry*, eds. Widdershoven G., McMillan J., Hope T. and van der Scheer L. Oxford: Oxford University Press.

- Butler, Daniel M.** 2014. *Representing the Advantaged: How Politicians Reinforce Inequality*. Cambridge: Cambridge University Press.
- Butler, Daniel M.** 2019. Facilitating Field Experiments at the Subnational Level. *The Journal of Politics* 81(1): 371–76. <https://doi.org/10.1086/700725>.
- 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.
- Butler, Daniel M. and David W. Nickerson.** 2011. Can Learning Constituency Opinion Affect How Legislators Vote? Results from a Field Experiment. *Quarterly Journal of Political Science* 6(1): 55–83.
- Campbell, Rosie and Diane Bolet.** 2022. Measuring MPs’ Responsiveness: How to Do It and Stay Out of Trouble. *Political Studies Review*, 20(2): 175–83.
- Clarke, Nick, Will Jennings, Jonathan Moss and Gerry Stoker.** 2018. *The Good Politician: Folk Theories, Political Interaction, and the Rise of Anti-Politics*. Cambridge: Cambridge University Press.
- Crawford, Lee and Ukasha Ramli.** 2021. Discrimination by Politicians against Religious Minorities: Experimental Evidence from the UK. *Party Politics* 28(5): 826–33. June, 13540688211021052. <https://doi.org/10.1177/13540688211021052>.
- Dafoe, Allan, Baobao Zhang and Devin Caughey.** 2018. Information Equivalence in Survey Experiments. *Political Analysis* 26(4): 399–416.
- Desposato, Scott.** 2018. Subjects and Scholars’ Views on the Ethics of Political Science Field Experiments. *Perspectives on Politics* 16(3): 739–50.
- Habel, Philip, and Sarah Birch.** 2019. A Field Experiment on the Effects of Ethnicity and Socioeconomic Status on the Quality of Representation. *Legislative Studies Quarterly* 44(3): 389–420.
- John, Peter, Kristina Kim and Luis Soto Tamayo.** 2022. ‘Replication Data for: “Using citizen voice to evaluate experiments on politicians: a UK survey experiment”’, <https://doi.org/10.7910/DVN/UDFLIA>, Harvard Dataverse, V1, UNF:6:0KjUP5zqeIjrj6QilAnhpTg= = [fileUNF]
- Loewen, Peter John, and Daniel Rubenson.** 2022. Value-Added and Transparent Experiments. *Political Studies Review* 20(2): 243–49.
- McClendon, Gwyneth.** 2012. Ethics of Using Public Officials As Field Experiment Subjects. *Newsletter of the APSA Experimental Section* 3(1): 13–20.
- McClendon, Gwyneth H.** 2016. Race and Responsiveness: An Experiment with South African Politicians. *Journal of Experimental Political Science* 3(1): 60–74.
- Miller, Kristen, Valerie Chepp, Stephanie Willson, and Jose-Luis Padilla.** 2014. *Cognitive Interviewing Methodology*. New York: John Wiley & Sons.
- Naurin, Elin, and Patrik Öhberg.** 2021. Ethics in Elite Experiments: A Perspective of Officials and Voters. *British Journal of Political Science* 51(2): 890–98.
- Vries, Catherine de, Elias Dinas, and Hector Solaz.** 2016. You Have Got Mail! How Intrinsic and Extrinsic Motivations Shape Constituency Service in the European Parliament. IHS Political Science Series No. 140, May 2016. Policy Paper. May 2016. <http://aei.pitt.edu/86057/>.

Cite this article: John P, Kim K, and Soto-Tamayo L (2024). Using Citizen Voice to Evaluate Experiments on Politicians: A UK Survey Experiment. *Journal of Experimental Political Science* 11, 42–51. <https://doi.org/10.1017/XPS.2023.3>