Political Science Research and Methods (2022), **10**, 353–371
doi:10.1017/psrm.2020.34

ORIGINAL ARTICLE

The grapevine effect in sensitive data collection: examining response patterns in support for violent extremism

John McCauley[1]* , Steven Finkel[2], Michael Neureiter[2] and Christopher Belasco[3]

[1]Governent and Politics, University of Maryland, Collge Park, College Park, USA, [2]Political Science, University of Pittsburgh, Pittsburgh, USA and [3]Graduate School of Public and International Affairs, University of Pittsburgh, Pittsburgh, USA
*Corresponding author. Email: mccauley@umd.edu

(Received 19 November 2018; revised 8 January 2020; accepted 16 April 2020; first published online 25 September 2020)

Abstract

This study presents a pattern overlooked in previous research on measuring sensitive political outcomes: over the course of data collection, responses tend to shift in the direction of support for the local incumbent power. We suggest that, whereas earlier responses are largely devoid of this social desirability bias, word of the research spreads across enumeration areas, and individuals interviewed later in the process alter their responses out of fear of retribution for inappropriate answers. We document the pattern using original data from two surveys on support for violent extremism conducted in three different countries in the Sahel region of Africa. We rule out a host of alternative explanations and further confirm that the pattern can arise not just with overt survey measures but even with covert, experimental ones. We then demonstrate the same pattern using out-of-sample data from a separate well-known study. The findings offer a cautionary note to both conventional and experimental approaches to measuring sensitive attitudes.

Key words: Data collection; survey methodology

Studying sensitive topics in political science has become increasingly sophisticated in recent years, as researchers seek methods to gauge true attitudes while minimizing social desirability and other response biases. Scholars now use natural randomization processes to predict judicial bias (Grossman *et al.*, 2016); implicit association tests to evaluate religious tensions (McCauley, 2014); functional brain imaging to explain leftist versus rightist political processing (Schreiber *et al.*, 2013); and conjoint analyses to gauge opinions on immigrants (Hainmueller and Hopkins, 2015). Research on violent extremism similarly confronts the concern that survey respondents may mask their true levels of support; scholars have thus employed endorsement experiments (Bullock *et al.*, 2011; Lyall *et al.*, 2020), list experiments (Blair *et al.*, 2014), and other subtle means to evaluate attitudes about support for terrorist groups or combatants that respondents may be unwilling or unable to express accurately through conventional survey methods.

These advanced and experimental methods help to overcome a number of shortcomings that bedevil conventional studies, and they take us much closer to generating causal inferences regarding attitude formation on sensitive topics. Yet, despite important progress in the methods employed to study sensitive topics, little attention has been accorded thus far to the possibility that the collection of sensitive data itself—whether overt or disguised by experimental methods—can alter the study environment, and thus the outcomes we seek to explain.

This paper explores how the study of a sensitive political topic—in particular, support for violent extremism—using population-based methods on the ground can alter the environment in which the study takes place, and thereby the attitudes of respondents. We suggest that over

© The Author(s), 2020. Published by Cambridge University Press on behalf of the European Political Science Association.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

the course of data collection, study participants who learn of the execution of research on the sensitive topic in their locality and then become subjects themselves may provide systematically different answers as the survey progresses, likely out of fear of retribution from the incumbent power. Social desirability bias typically reflects the concern that individuals alter their responses to survey questions based on information they glean immediately as questions are posed. Here, we argue that social desirability bias can emerge collectively "through the grapevine", as previous respondents share the contours of sensitive enumeration topics with other community members. As a result, respondents' answers to measures of extremist support gradually tend to side with incumbent power holders in the areas where they live.

To test these claims, we rely on two original surveys conducted in the Sahel region of Africa, where terrorist groups have become increasingly disruptive. Using evidence collected over a six-week period from over 7700 respondents in Burkina Faso, Chad, and Niger in 2013, and again over an eight-week period among approximately 7900 respondents in the same context in 2017, we show that expressed support for violent extremism changes systematically over the course of the data collection periods. More specifically, individuals interviewed at later times *within* a given commune—the smallest geographic administrative unit in the Sahel and the basis for the study's sampling frame—are significantly less likely to express support for extremist violence than are individuals interviewed earlier in that same commune. Whereas at the outset of data collection, 55 percent of respondents were neutral or positive on at least one of our measures of extremist support, by the end of the data collection period within respondents' communes this number was 18 percent, a decrease of 37-percentage points over an approximately one-month period. Even responses to an endorsement experiment meant to overcome conventional concerns of social desirability bias prove susceptible to this grapevine effect, though the effects are less consistent than observed with overt measures. We suggest that respondents' answers become increasingly pro-government as the data collection on extremism in the region advances, as respondents fear reprisal from those in power in the enumeration areas.

We account for a number of possible alternative explanations, including selection effects in the type of respondents who take part earlier and later and changes in enumerator behavior over time, and we include commune-level dummy variables in the models in order to control for between-commune confounding in the estimation of within-commune time-of-interview effects. To further evaluate our claims, we study patterns in support for extremist groups using out-of-sample data from a separate well-known study, Corstange's (2016) analysis of support for Hizballah in Lebanon. The results lend credibility to the argument: support for violent extremist groups changes systematically over the course of data collection in the direction of support for the local incumbent power.

The study is novel in several respects. First, it goes beyond average effects to explore systematic changes in responses over time as a function of the research enterprise itself. Second, it recognizes the endogenous relationship between survey outcomes and the local political powers who sanction those data collection activities. When extensive data collection on sensitive topics takes place in communities, local incumbent powers are implicated more extensively than we typically acknowledge, and community members are not naïve to their place in the research. We suggest that those factors bear importantly on the outcomes we study. Third, we exploit original data from an understudied region of the world that is increasingly central to the fight against violent extremism, and we go a step further in illustrating the argument using an additional data source from a different region.

The implications of this methodological cautionary note are clear. Especially with the use of overt measures of sensitive attitudes, but even under subtle, experimental conditions, social desirability bias can still creep into population-based data as information regarding data collection spreads. Researchers studying sensitive topics, especially in the developing world, must be acutely aware of the relationships between their work on the ground, the subjects of their study, and the local powers who sanction that work. Further, the results suggest that more intense data collection

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

over shorter periods of time may better capture the unadulterated attitudes of respondents than population-based research carried out over a longer time period.

1. Explaining changes in response patterns over time

Understanding how response patterns might change systematically over the course of data collection requires explication of the role of the local power holder, the perceptions and self-awareness of those studied, and the content of the study. We explore each of these components.

1.1 Local power holders

We begin by noting that studies involving extensive data collection at the individual level almost always require the authorization of national and local power holders. In our own research outlined here, permissions were obtained at the ministry level and from local authorities in each enumeration area. In turn, the enumeration teams were issued certification letters, and the governments helped to publicize the survey and encourage participation via radio announcements in various languages. As Fujii (2012) notes, it is both implausible and reckless to collect data from local contexts in the developing world without authorization, as local power holders will certainly know of the presence of researchers and maintain a right to sanction the research activities, particularly those undertaken by outsiders. Importantly, the local incumbent power may not always be the formally recognized government; Atran *et al.*, (2017) stress that in the context of studying terrorism, official governments can in some cases be too weak to provide consent, having abdicated that role to rebel opposition or terrorist groups in certain areas. This is a point to which we return later in the paper. Irrespective of who holds de facto local power, however, we suggest that the process of obtaining and publicizing local authorization creates a broader awareness in enumeration areas that the sanctioning power is in some sense implicated in the research.

In the study areas of Burkina Faso, Chad, and Niger exploited for this study, the formally recognized governments constitute the local power holders. Burkina Faso and Niger each score above 5 on the Polity IV democracy index, suggesting relatively democratic institutions and governments that exercise control with at least moderate support and legitimacy. Chad is less democratic but exerts fairly heavy-handed military control over citizens, mitigating the potential for rebel or extremist groups to gain de facto control (Marchal, 2016). All three countries take part in the regional G5 Sahel Cross-Border Joint Force aimed at undercutting terrorist activity,[1] and, while extremists carry out insurgent attacks in the region, no rebel or extremist groups hold territorial control in the study area. As of our data collection, the extremist groups that are active in Burkina Faso, Chad, and Niger by and large have not operated openly; instead, they infiltrate communities and operate in the shadows or with one-off attacks (Antwi-Boateng, 2017).[2] For the purposes of hypothesis testing below, we thus treat the formally recognized governments as the local incumbent power holders.

1.2 Individuals under study

From the perspective of individuals under study, we note three tendencies relevant to response patterns over time. First, especially in rural or developing contexts, community members are typically well aware of the population-based surveys conducted around them. Hershfield *et al.*, (1983) note that the presence of enumerators tends to be more conspicuous than researchers

[1]The G5 operates with support from France, the United Nations, and others. For more details, see France Diplomatie, "G5 Sahel Joint Force and the Sahel Alliance," https://www.diplomatie.gouv.fr/en/french-foreign-policy/defence-security/crisis-and-conflicts/g5-sahel-joint-force-and-the-sahel-alliance/.

[2]This has begun to change somewhat in Burkina Faso as the security situation has deteriorated in the period following our data collection. See "Kalashnikovs and no-go zones: east Burkina Faso falls to militants." The Guardian, 22 April 2019.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

assume, such that word of the research activities often precedes their arrival at households. That public opinion surveys like the Afrobarometer typically record the presence of community members listening in on ostensibly private survey interviews further provides an indication of the extent to which individual data collection is in some sense a community-wide activity. As knowledge of the activities spreads, later respondents provide attitudinal data from a different informational baseline than their counterparts reached earlier in the enumeration period.

Second, participants in public opinion surveys in the developing world often maintain fairly deep skepticism regarding the provenance of survey questionnaires and the confidentiality of their answers. According to Afrobarometer data, over one-third of respondents from Round 6 (and over half from Round 5) believe that enumerators for this non-partisan, non-governmental enterprise were sent by the government, despite a clear introductory statement to the contrary.[3] Longstanding skepticism about the secrecy of votes in elections—sometimes justified—can similarly undermine community members' confidence in other ostensibly confidential activities such as survey enumeration (Kalinin, 2016). In fact, Singer *et al.*, (1992) note that more elaborate assurances of confidentiality in the collection of individual-level data can be counterproductive, increasing participants' concerns that their responses may be revealed. In short, participants in population-based research often assume that their responses may be known to local authorities.

Third, we assume that community members understand the capacity of local power holders to punish. Trust in police forces in our region of study remains low, owing in part to what is perceived as excessive use of force and arbitrary applications of punishment without due process (Goldsmith, 2005). Contexts controlled by rebel or terrorist groups can be much worse: the Taliban notoriously applied rigid punishment against women for minor dress code violations (Goodson, 2001), and as Boko Haram expanded in northern Nigeria, community members who challenged the group's authority were summarily killed (Mohammed, 2014). In response, up to 1.7 million persons are estimated to have moved internally (UNHCR 2017), indicating a clear understanding of the incumbent power's ability to punish.

To summarize the individual-level perspective, knowledge of population-based data collection tends to spread as those activities advance; meanwhile, participants often assume that their answers may be known to local authorities who possess the capacity to punish community members deemed to be deviant or uncooperative. In this context, we suspect that fear—or at least a desire to acquiesce—drives participants to increasingly offer attitudes consistent with the perspective of local incumbent power holders as word of the data collection spreads. This would be especially true regarding sensitive topics that implicate the local power holders, including security matters, religious divisions, and violence.

1.3 The content and context of study

Researchers have long understood the importance of contextual factors for survey outcomes. As social interactions, face-to-face data collection can be affected by respondents' experiences, perceptions regarding interviewers, and local environmental factors (Tourangeau and Yan, 2007; Krumpal, 2013), and all of those pressures are amplified when survey topics are perceived as sensitive or taboo (Berinsky, 2004). For example, researchers have demonstrated changes in response patterns as a function of interview language (Lee and Pérez, 2014), interviewer ethnicity (Adida *et al.*, 2016), interviewer race (Davis and Silver 2003, Finkel *et al.*, 1991), and interviewer gender (Johnson and DeLamater, 1976). Blaydes and Gillum (2013) find that when Egyptian women are surveyed regarding the sensitive issue of personal piety, their responses shift toward the devout when enumerators wear Islamic headscarves.

Broader political factors can also shape response patterns and the willingness or capacity of respondents to discuss sensitive topics truthfully. Robinson and Tannenberg (2019), for example,

[3]See the merged Round 6 data at http://afrobarometer.org/data/merged-round-6-data-36-countries-2016.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

find that in autocratic countries, respondents practice pro-regime self-censorship on items relating to trust. Corstange (2014, 2016) shows that survey refusals increase when sponsors are identified as foreign embassies as opposed to independent research institutions.

All of these contextual effects can be summarized as forms of social desirability bias, whereby respondents provide altered, more socially appropriate answers as a strategic or subconscious reaction to either the enumerator or the broader political context (Holtgraves, 2004). In response to these challenges, scholars have increasingly adopted subtle, experimental means to access sensitive attitudes (see, e.g., Rosenfeld *et al.*, 2016). Importantly, Blair *et al.* (2014) find that responses to experimental measures approximate true attitudes in a way that standard survey questions cannot, adding confidence to the assumption that participants do not simply game the surveys. This is not to suggest, however, that answers will remain stable across respondents as information on the data collection spreads.

We argue that, over the course of population-based data collection on sensitive topics, a grapevine effect takes place. Grapevine effects have been cited in previous research to describe the legal transfer of responsibility for defamation as claims spread (Douglas, 2015) and the informal transfer of purchase information among consumers (TARP 1981). Here, we suggest that survey research on sensitive topics inspires community conversation, both when the topic is raised overtly and even when the subtlety of experimental measures masks the intent behind specific questions.[4] Respondents learn, not from their own experience with a survey question but from word-of-mouth in their communities, that the data collection in some way involves sensitive matters (such as, in this case, support for violent extremism). As a result, subsequent respondents are likely to be more attuned to the consequences of their responses than earlier ones, as fear of reprisal for inappropriate answers spreads.

To summarize the study's hypothesis, we expect that respondents who take part later in the collection of sensitive survey data on violent extremism will give answers systematically different from earlier respondents in their communities, moving in the direction of support for the local incumbent power. We would not expect this pattern to emerge in every survey—those with widely dispersed respondents, for example—but where respondents are queried in fairly close proximity, the word may travel more quickly than the enumeration. In the context of this study, we thus anticipate that expressed attitudes will tilt in the favor of government interests and against violent extremist groups as data collection proceeds. We surmise that this occurs because participants who become aware of the ongoing sensitive survey fear retribution for inappropriate answers from those in power.

The process by which this grapevine effect unfolds can be captured in qualitative reports. Our enumeration teams cited three common behaviors in the households they visited. First, despite requests for private interviews and efforts to situate the respondent away from onlookers, an enumeration supervisor reported that anywhere from two to several bystanders observed many of the interviews.[5] Children and the head of household were often present, but neighbors also sometimes followed proceedings before leaving the compound. Second, in the rural communes and tightknit urban neighborhoods in which the data collection took place, the local well or water pump serves as a center for social engagement, informal networking, and gossip. "If you raised the name of Al Qaeda in a survey with me, I promise you I would discuss it the next day at the water pump," the enumeration supervisor noted in a follow-up interview.[6] Third, those who gain

[4]Of course, experimental measures of sensitive topics in population-based research are designed precisely to prevent participants from deciphering the true nature of the inquisition. Yet, even in the context of experimental treatments embedded in surveys, half of the sample (conventionally) is still exposed to questions that can raise awareness of a sensitive issue and thus trigger the impulse to share information with other community members.

[5]Recorded data indicates that 52 percent and 35 percent of interviews were attended by onlookers during the two respective surveys.

[6]Interview, Malik Traoré, 18 August 2019. The practice has been noted elsewhere, such as in Rathgeber (1996).

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

Table 1. Summary of data collection

Country	Wave	Time Period	Zones	No. of respondents
Chad	Survey 1	September—October 2013	30	2855
Chad	Survey 2	April 2017	30	2846
Niger	Survey 1	November 2013	30	2710
Niger	Survey 2	April 2017	30	2855
Burkina Faso	Survey 1	September—October 2013	23	2155
Burkina Faso	Survey 2	March—April 2017	23	2185
Total	Survey 1			7720
	Survey 2			7886

knowledge of the survey and its general contours, either from observing an earlier interview, hearing about it at the water pump, or learning through other channels, would be likely to share the news with other members of their households. Enumerators were often greeted with "we've been waiting for you to come by our house,"[7] suggesting that awareness of the surveys indeed spread through communities, and references to extremist groups were a likely reason why. We test for evidence of such a grapevine effect below.

2. Data and analyses

We test the hypothesis using original data from two surveys conducted in 2013 and 2017, respectively, within the southern half of Niger, northern Burkina Faso, and the middle portion of Chad (see Figure A1 in the Online Appendix). The geographic region was selected based on its increasing susceptibility to terrorist attacks, and the increasing likelihood of extremist group recruitment among local communities (Alexander, 2012). Data collection in the three countries took place between September and November 2013 (Survey 1) and between March and April 2017 (Survey 2). Primary sampling units are either communes—the lowest geographic subdivision in rural areas, typically representing a small village—or arrondissements (small neighborhoods) within larger cities.

Respondents were selected using a multistage, clustered random sampling procedure with stratification by gender. Each first-level sub-national administrative unit in the study area, of which there are 19 across the three countries, was divided into a maximum of eight sub-areas, depending on size. Those 83 sub-areas were in turn divided into potential primary sampling units (PSUs), containing an average of approximately 200 households. Next, one PSU, or commune, was randomly selected from each sub-area. Within each PSU, enumerators identified households using a fixed-interval procedure and randomly drew a respondent between the ages of 15 and 73 from within that household. Table 1 summarizes the data collection timeline, the number of sampled zones, and the number of interviews conducted by country. Table A1 in the Appendix describes the sampling procedure with spatial details to justify the practical plausibility of a grapevine effect within communes.

In total, data were drawn from 7720 respondents across 83 communes in the three different countries for Survey 1, and from 7888 (different) respondents across the same 83 communes for Survey 2, all using systematic random sampling procedures.

With respect to a potential grapevine effect, data collection protocols remained consistent from commune to commune, but the length of time to complete surveys in each enumeration area varied. For Survey 1, data collection lasted from two days (in multiple communes) to 29 days (in Tchintabaraden, Niger), with a mean of 5.8 days. For Survey 2, data collection took from two days to 39 days (in Ouahigouya, Burkina Faso), with a mean of 8.4 days. In Ouahigouya and a few other communes, data collection was extended as enumerators took breaks to account for extreme heat in March and April.

[7]Interview, Malik Traoré, 10 September 2018.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

Table 2. Overt support for violent extremism (OLS)

	Survey 1	Survey 2
Days since first interview in commune	−0.016**	−0.006**
	(0.008)	(0.002)
Constant	1.450***	1.184***
	(0.016)	(0.010)
Observations	7553	7784

Note: Standard errors clustered by commune in parentheses; commune dummies included in models but not reported; *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

2.1 Overt measures of support for violent extremism

To test our claims regarding a grapevine effect, we first evaluate overt expressions of support for violent extremism, where the effects should be most evident. To measure the outcome of support for violent extremism, we create a composite index from four survey items,[8] each coded on a 3-point scale with higher values indicating greater support:

- a) Do you agree or disagree with the following statement: Al Qaeda's violent actions are permitted under Islamic law?
- b) Do you agree or disagree with the following statement: Violence in the name of Islam can be justified?
- c) When do you think that violence is an effective method to solve problems: often, sometimes, or never?
- d) Do you personally feel that using arms and violence against civilians in defense of your religion can be often justified, sometimes justified, or never justified?

The key explanatory variable is the number of days since the first survey interview within the respondent's commune. Focusing on the passage of time within communes provides a measure appropriate to the geographically constrained logic of a grapevine effect, whereby word-of-mouth would likely travel from one household to another within a community and the process would start anew in other surveyed communities. We use OLS regression with standard errors clustered by commune (our PSUs), along with commune dummy variables to account for different start dates and other between-commune factors that may confound the estimation of within-commune grapevine effects.

As Table 2 shows, in bivariate analyses, the number of days since the first interview in a commune has a statistically significant, negative effect on overt support for violent extremism in both surveys. That is, respondents who are surveyed later in the data collection period within a commune are systematically less likely to express overt support for violent extremism, a pattern in keeping with the preferences of their local incumbent powers (the governments fighting terrorist groups in the region). The results cannot be easily attributed to social or political changes that might covary with the period of data collection: the surveys took place over several weeks and months and with different start times in each commune, and yet a within-commune pattern of decreasing support for violent extremism persists. Additionally, by including commune-level fixed effects, we account for any potential time-invariant factors that could confound the results, such as proximity to terrorist bases or exposure to previous attacks.

To address other potential alternative explanations, we add a series of controls. Most importantly, variation in support for violent extremism may be a function of who lives where within each commune. Previous research suggests that wealthier and better-educated individuals are more supportive

[8]Composite indices are commonly used to measure violent extremism, given its multifaceted nature. See, for example, Schils and Pauwels (2014) and Wray (2012). We tested the effects of the four indicators discretely; three of the four have effects consistent with the composite index (see Appendix Table A2).

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

Table 3. Overt support for violent extremism—alternative explanations (OLS)

	Survey 1	Survey 2
Days since first interview in commune	−0.017*	−0.006***
	(0.009)	(0.002)
Wealth	−0.007	0.010***
	(0.007)	(0.003)
Education	0.008	0.002
	(0.005)	(0.002)
Employed	0.007	−0.005
	(0.016)	(0.009)
Constant	1.459***	1.150***
	(0.031)	(0.014)
Observations	3762[9]	7782

Note: Standard errors clustered by commune in parentheses; commune dummies included in models but not reported; *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

of violent extremist groups (Blair *et al.*, 2013); if those better off individuals live in clustered areas and are more likely to be surveyed earlier in the data collection period, our explanation would be spurious. Similarly, employment status may covary with proximity to local institutions and services and may also affect attitudes regarding the disruption caused by terrorism. Thus, if the household identification process led enumerators closer to or further from those centers over subsequent days, what may appear to be a consequence of word-of-mouth learning as the days progress could instead be driven by differences in employment status.

Table 3 adds controls for wealth, education, and employment. We measure wealth using an additive index, *Wealth*, denoting how many out of 13 possible household items such as a refrigerator, TV, and radio respondents have in their households. *Education* is measured on a 10-point scale ranging from no formal education to a postgraduate degree. *Employed* takes the value of 1 if a respondent reports being employed and 0 otherwise. If the pattern we presented above is simply a function of the type of respondent surveyed earlier and later in the data collection process, we should find that the number of days since the start of data collection loses its significance at the expense of those variables. As Table 3 shows, however, the coefficient on the variable for survey timing remains negative and significant, indicating that support for extremism continues to decrease as data collection time elapses, even with those controls.[10]

The indexed measure of support for violent extremism does not easily lend itself to substantive interpretation, so we consider responses to the discrete components. Fifty-five percent of respondents at the outset of Survey 1 provided a neutral or positive response to at least one of the four measures of extremism and violence.[11] By the end of data collection within communes, however, that number had fallen to 18 percent, a decrease of 37 percentage points over approximately a one-month period or less within communes. Support for violent extremism in Survey 2 similarly declined by 30 percentage points.

[9]The Survey 1 results also hold when omitting employment status, which had substantial amounts of missing data. In a model with only wealth and education as controls ($n = 7551$), the coefficient on Days since First Interview in Commune is −0.16 ($p < 0.05$).

[10]Without randomizing the order in which respondents are interviewed, which is rarely the case in field surveys, one cannot be certain of avoiding location-based confounders, especially from the center to peripheries of communes. We are confident that the fixed-interval method with randomized daily start points led to full and varied coverage of the enumerated areas and that the included controls account for the principal factors that could lead to spurious findings. We thank an anonymous reviewer for the insight.

[11]The remaining 45 percent disagreed with or responded "never" to all four propositions.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

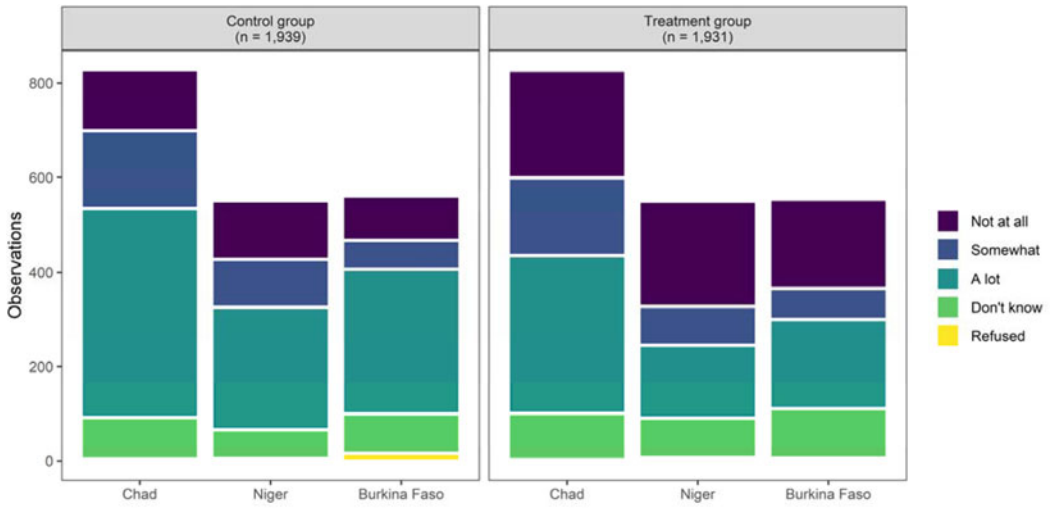


Figure 1. Descriptive Plot of the Endorsement Experiment (Survey 1).
Note: "A Lot" indicates strong support for the polio vaccination program, which we interpret as a rejection of AQIM (since the treatment reported that AQIM is opposed to the initiative). "Not at all" indicates a rejection of the vaccination program, or support for AQIM.

2.2 Experimental measures of support for violent extremism

While respondents may be particularly likely to alter their stated answers in response to overt survey questions, we are also interested in evaluating whether the grapevine effect may influence even those data collection methods that aim to circumvent strategic responses. To measure support for violent extremism in a subtle, unobtrusive manner, we apply an endorsement experiment to gauge individuals' support for violent Islamist groups without asking them to directly reveal their attitudes about this group. The use of specific extremist group labels as indicators of support for violent extremism more generally has been established elsewhere (see Fair *et al.*, 2018). The endorsement experiment was designed as follows. First, respondents were randomly assigned to control and treatment groups, with one half of the sample assigned to each. Respondents in the control group were asked the following question:

> *The World Health Organization recently announced a plan to introduce universal Polio vaccination across {Country}. To what extent do you approve of such a plan?*

(1) Not at all.
(2) Somewhat.
(3) A lot.
(4) I don't know.
(5) I refuse to answer.

The treatment group heard the same question but with the inclusion of an extremist group endorsement; respondents were additionally told that "*It is likely that Al-Qaeda in the Islamic Maghreb (AQIM), an Islamist group, will oppose this program.*"[12] We measured support for AQIM as the difference across the treatment and control groups in the proportions of individuals who did "not at all" approve of the polio vaccination.

Figure 1 below shows a descriptive plot of the endorsement experiment for the first survey.[13] It can be seen that across all three countries, substantially more individuals in the treatment group

[12]While Boko Haram was more active in Niger, AQIM was active and widely known across the three countries.
[13]A similar descriptive plot for the second survey is provided in Appendix Figure A2.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

Table 4. Covert support for AQIM (OLS)

	Survey 1		Survey 2	
Treatment	0.179***	0.177***	0.296***	0.295***
	(0.024)	(0.024)	(0.033)	(0.033)
Days since first interview in commune	−0.013	−0.013	−0.004	−0.003
	(0.009)	(0.009)	(0.005)	(0.005)
Treatment * Days	0.003	0.003	−0.006**	−0.006**
	(0.004)	(0.004)	(0.003)	(0.003)
Wealth		−0.013**		0.003
		(0.006)		(0.004)
Education		−0.011*		−0.007
		(0.006)		(0.005)
Employed		0.009		−0.025
		(0.022)		(0.017)
Constant	0.232***	0.292***	0.284***	0.307***
	(0.020)	(0.035)	(0.031)	(0.039)
Observations	3277	3251	3685	3684

Note: Standard errors clustered by commune in parentheses; commune dummies included in models but not reported; *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$. The experiments were administered to a random subset of respondents.

were "not at all" in support of the policy, which we infer as covert support for AQIM. On average, approximately 18 percent of our sample supported AQIM covertly, with these figures relatively consistent across the three-country contexts.

Next, we test whether covert support for violent Islamist organizations is a function of the timing of interviews. We measure time as the number of days that passed since the first interview was conducted in the respondent's commune (thus, 1–28 for Survey 1 and 1–38 for Survey 2). We reverse-code the treatment responses to transform the outcome variable from "support for the polio vaccination program" to "support for the violent Islamist organization." We then dichotomize the measure so that "not at all" supporting the vaccination program is coded as 1 (support for the Islamist organization) and moderate or strong support for the vaccination program is coded as 0 (meaning little or no support for the Islamist organization).

In addition to variables for the treatment (the Islamist endorsement) and the number of survey days elapsed, we include an interaction term that interacts survey timing with the endorsement treatment. This represents the additional impact of time on the vaccination policy response among the experimental group versus the control group, and thus serves as a test of whether respondents are more or less likely to support the extremist group as the survey progresses in their commune.

Table 4 shows regression results measuring support for the Islamist group in each of the two surveys, with controls included. Following Freedman (2008), we use Ordinary Least Squares regressions both because randomization alone does not justify the logit linking function and because of the ease of interpretability of OLS results, though we also include logistic analyses in Appendix Table A3. As noted, standard errors are clustered by commune. A negatively signed coefficient indicates a decrease in respondents' level of support for the Islamist group.

The results suggest that not even covert, experimental measures are immune to grapevine effects. The treatment variable measuring support for AQIM in the presence of the endorsement is positive and significant, as expected. Further, while the interaction of treatment and days is not significant in Survey 1, it is negatively signed and significant in Survey 2, indicating that those receiving the AQIM endorsement increasingly sided with the government and not the terrorists as data collection in their communes progressed, despite the disguised means of inquiring about their allegiances.

For interpretative purposes, Figure 2 illustrates the linear impact of the treatment on individuals' likelihood of supporting AQIM in Survey 2. While in the initial days of the data collection,

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

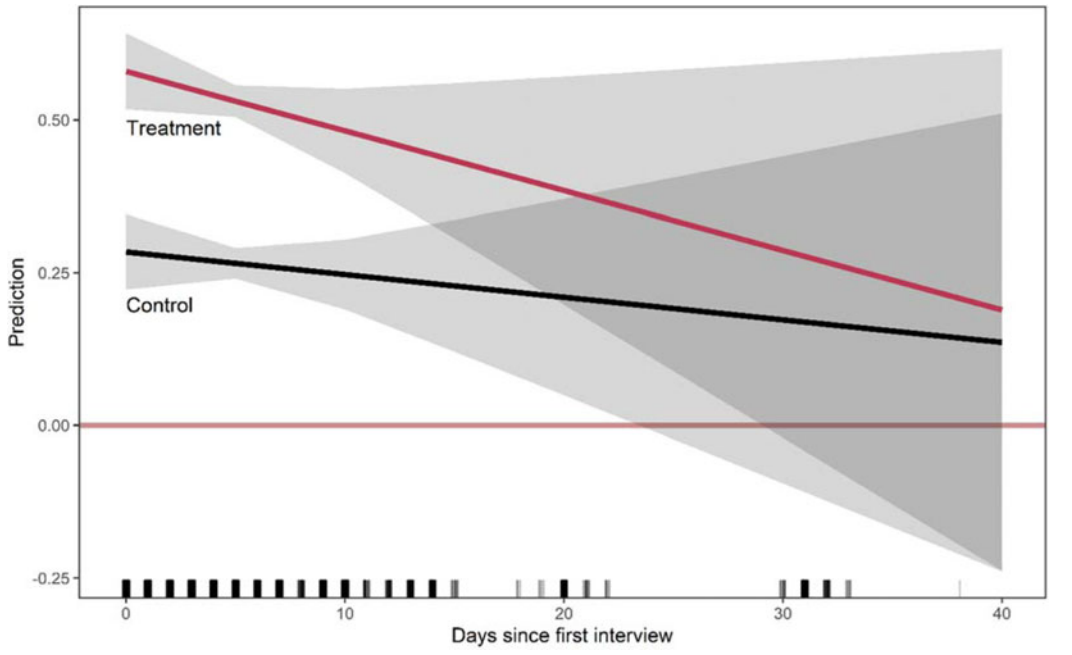


Figure 2. Support for AQIM in Survey 2 (95 percent CIs).

the treatment increases support for the Islamist organizations (i.e. respondents who hear the Islamist group's position are more likely to "not at all" support the vaccination program), that effect steadily declines over the course of the data collection. By the end of the period of data collection in each commune, receiving the treatment has no statistically significant bearing on respondents' likelihood of expressing support for the extremist groups.

Importantly, time appears to affect individuals' responses in the treatment group but less so those in the control group, a finding consistent with the hypothesis. The overall results suggest that support for Islamist groups shifts away from the Islamist groups and thus in the direction of the anti-extremist governments, representing the local incumbent power holders. We argue that this pattern should be anticipated, as individual respondents become increasingly fearful about countering the local power holder during what many perceive as non-confidential inquiries.

While the commune-level fixed effects employed in our models account for any unobserved confounders in the broader context and in time-invariant differences between communes, we also add calendar day fixed effects to the models as a robustness check. The results remain consistent with our expectations. See Appendix A7 for details.

3. Placebo tests

An important remaining concern may be that better trained and more experienced supervisors are responsible for monitoring the earliest interviews in each locality, which might suggest that they are better equipped to elicit responses untainted by respondent fear or skepticism. Alternatively, enumerators may unwittingly pose survey questions in a systematically different manner over the course of data collection, thereby generating response patterns that gradually reveal stronger support for the local incumbent power. To address potential enumerator effects, we begin by testing whether certain enumerators conducted earlier interviews while others conducted interviews during the later stages of data collection in each commune. The results do not

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

support this alternative hypothesis: the identity of the supervisor is not correlated with our main independent variable, i.e., time since the first interview, and payment receipts along with enumerator codes for individual enumerators working in teams under those supervisors likewise reveal that enumerators worked consistently from beginning to end of the data collection within their enumeration areas (receipts reviewed but not shown). We also ran analyses including enumerator fixed effects and the coefficient on days since First Interview in Commune remains nearly identical, so the patterns observed are unlikely to be explained by enumerator effects (see Appendix Table A4).[14]

To determine whether enumerators may have systematically changed the way they posed the sensitive survey questions as enumeration days elapsed, we evaluate patterns in responses to separate questions unrelated to support for violent extremism. An appropriate placebo test in this context should use as an outcome variable an issue over which the local incumbent power has an understood preference, but which would not be so sensitive a matter as to elicit fears of reprisal among respondents who wish to respond in a manner counter to that preference.

We note, first, that the control category in the endorsement experiment described above—those who receive information on the polio vaccination program without reference to AQIM—itself serves as a placebo. Governments in all three countries collaborate with the World Health Organization, the United Nations, and local partners on existing polio vaccination programs,[15] which suggests an understood preference though not one that would merit reprisal for dissenters. If enumerators coaxed more suitable answers out of respondents over time in response to any question for which the local incumbent power has an understood preference, we should also expect increasing support for the program among this group. In this sense, control categories in the context of experimental manipulation can be useful beyond the comparison to treatment groups. That control group responses remain largely stable within communes over the course of data collection, however, suggests that enumerators are not playing such a role.

As an additional placebo test, we use responses to a question regarding support for inter-ethnic marriage. The question similarly connotes an attitude by local residents that could reflect poorly on incumbent governments, who may reasonably wish to present an image of ethnic tolerance. Thus, if enumerators "learn" over the course of data collection, either consciously or otherwise, to pose survey questions in such a way as to elicit responses supportive of the local power holder, we would expect the same pattern to emerge in responses to the inter-ethnic marriage question that we demonstrate regarding the measure of support of Islamists. Yet it does not: as Figure 3 illustrates, the trends move modestly in different directions for the two surveys, but the confidence intervals overlap with the zero lines in both cases, indicating that responses to the inter-ethnic marriage question remain statistically stable as the number of survey days elapses. To the extent that a rejection of inter-ethnic marriage may be perceived as inappropriate to express in an interview and counter to the wishes of the incumbent power, the placebo test suggests that enumerators do not refine the manner in which they pose questions to elicit pro-government responses. Instead, we argue that the pattern over time in responses to our survey questions regarding Islamist group support shifts in the direction of support for the local incumbent power because of the topic's unique sensitivity and the threat of reprisal that respondents perceive for counter-conventional views.[16]

[14]We have data on enumerators only for Survey 2; therefore, for Survey 1 we use sampling point as a proxy for enumerator.

[15]See https://www.who.int. We also thank an anonymous reviewer for providing additional information on the polio vaccination programs and their potential effects on control and treatment recipients. Insofar as control responses remain stable over time within communes, we do not attribute the treatment effects to an increase over time in the vaccination program's popularity.

[16]We acknowledge the possibility that enumerators might not only learn to pose sensitive questions in a manner that elicits socially desirable answers, but that they may also discern which among the sensitive questions might elicit potential reprisal from local incumbent power holders and then *only* pose those questions (but not other sensitive ones) in a different manner over time within each commune. We view the grapevine effect as a more plausible explanation, given the totality of the evidence.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

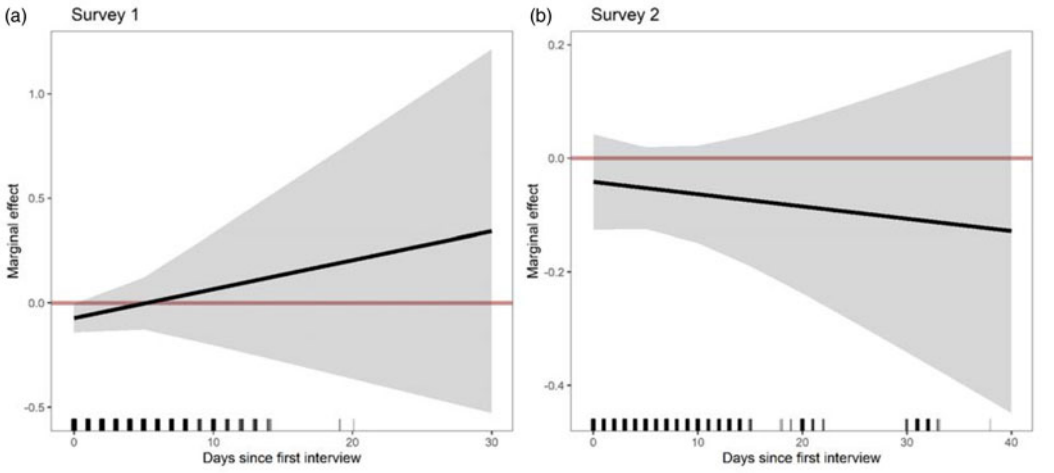


Figure 3. Placebo Test: Support for Inter-Ethnic Marriage (95 percent CIs).

To further address any concerns regarding enumerator effects or enumerator learning over time, we take two additional steps. First, we replicate our results including the number of days since the first interview conducted by each enumerator as an additional covariate to control for potential changes in enumerator behavior over time. The results remain robust to this change.[17] In addition, we add an enumerator variable to the model with calendar day fixed effects, and again the results hold up. Results are reported in Appendix Table A9.

Finally, to address the possibility of objective changes in Islamist group popularity during the course of data collection, we note that the effects generally hold across two separate surveys conducted at different time periods across three different countries and that the actual start date of the interviews varied by enumeration area for each of the two surveys. The results that we present here are thus inconsistent with the Islamist groups in question becoming objectively less popular during a particular time window across all the surveyed areas. Instead, the pattern persists within individual communes during the data collection there.

4. Out-of-sample evidence

We have demonstrated a pattern in a multi-country context, using two different surveys: over the course of data collection within communes, survey responses to sensitive research questions regarding support for violent extremism tend to shift in the direction of support for the local incumbent power. In our sample, the local powers are the formally recognized governments and security forces as opposed to rebel or terrorist groups, but we noted at the outset that it is possible for such non-state actors to hold de facto local control. In this section, we expand the analysis to a different context using data from a well-known study on support for the rebel Shi'a group Hizballah in Lebanon. In doing so, we introduce variation in the type of local incumbent power and demonstrate that the argument remains robust to this change.

4.1 Support for Hizballah's disarmament among Beirut residents: Corstange (2016)

In a 2016 study, Daniel Corstange discusses the results of a survey experiment of 2481 households in the greater Beirut metropolitan area, in which he manipulates the putative sponsor of the

[17]If anything, enumerators may induce responses in the opposite direction that would have run counter to a grapevine effect. See Appendix Table A8. Since we do not have data on enumerators for Survey 1, we conduct this robustness check only for Survey 2.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

Table 5. Overt support for Hizballah (Corstange, 2016)

	Model 1
Days since first interview in neighborhood	−0.057
	(0.067)
Days* Shi'a-majority neighborhood	0.197**
	(0.081)
Education	−0.028
	(0.039)
Income	0.036
	(0.040)
Constant	3.086***
	(0.251)
Observations	1910

Note: OLS regressions with standard errors clustered by neighborhood in parentheses; neighborhood dummies included in models but not reported; *** p < 0.01 ** p < 0.05 * p < 0.10.

survey and asks local residents about their views regarding support for Hizballah. The study includes both overt, direct measures of support for Hizballah and measures of support for Hizballah following a covert manipulation. It was conducted within a month of the formation of a Hizballah-led cabinet, between June 29 and July 14, 2011.[18] One complexity in the study context is that, while Hizballah had formally taken control, non-Shi'a neighborhoods in Beirut remained in a stance of conflict vis-à-vis Hizballah.

We test whether responses changed in the direction of the local power—either pro- or anti-Hizballah depending on the sectarian make-up of the community—as data collection unfolded. While our analyses do not undermine the central findings of the Corstange paper, they do suggest that a grapevine effect may emerge in population-based studies beyond our own. Like the analyses of our own survey data, we find strong support for such an effect using the overt measures and suggestive patterns even with the use of experimental manipulations.

To replicate the overt, direct measure of support for Hizballah in the Corstange (2016) study, we use the same dependent variable, a composite index of four questions regarding Hizballah's disarmament, the Syrian conflict, the neutrality of the international Supreme Tribunal for Lebanon (STL, which has implicated Hizballah members in domestic assassinations), and the desirability of abolishing the STL. All four items are coded on a 5-point scale, with higher values indicating greater support for Hizballah; we use the numerical average for the composite measure. We include available controls similar to our own, the key independent variable of days since the first interview in a locality, and an additional interaction term to account for the effect of days specific to Shi'a-majority neighborhoods.[19]

As Table 5 shows, the average effect of time since the first interview in a neighborhood is insignificant. However, in Shi'a-majority neighborhoods, where Hizballah is the recognized local power, we see a different effect: the interaction of days and Shi'a neighborhood indicates that respondents in those areas increasingly provide responses in support of Hizballah as the data collection in their neighborhood advances.[20]

Figure 4 illustrates this point. The effect of time in non-Shi'a neighborhoods is marginally negative in slope, as might be expected from sectarians who do not accept the authority of

[18]Reuters, June 13, 2011. Lebanon gets Hizballah-led cabinet after five-month lag. https://www.reuters.com/article/us-lebanon-government/lebanon-gets-Hizballah-led-cabinet-after-5-month-lag-idUSTRE75C48K20110613.

[19]We present analyses using all 17 enumeration areas. We recognize that a grapevine effect may not be expected to emerge when data collection within a commune is completed within a single day. In Appendix Table A5, we thus drop the three enumeration areas for which that occurred, and the results remain consistent.

[20]The sampling procedure in Corstange (2016), a multi-stage random sampling from across the city, makes the detection of a grapevine effect more difficult, though respondents were drawn exclusively from neighborhoods within Beirut. We thank an anonymous reviewer for pointing this out.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

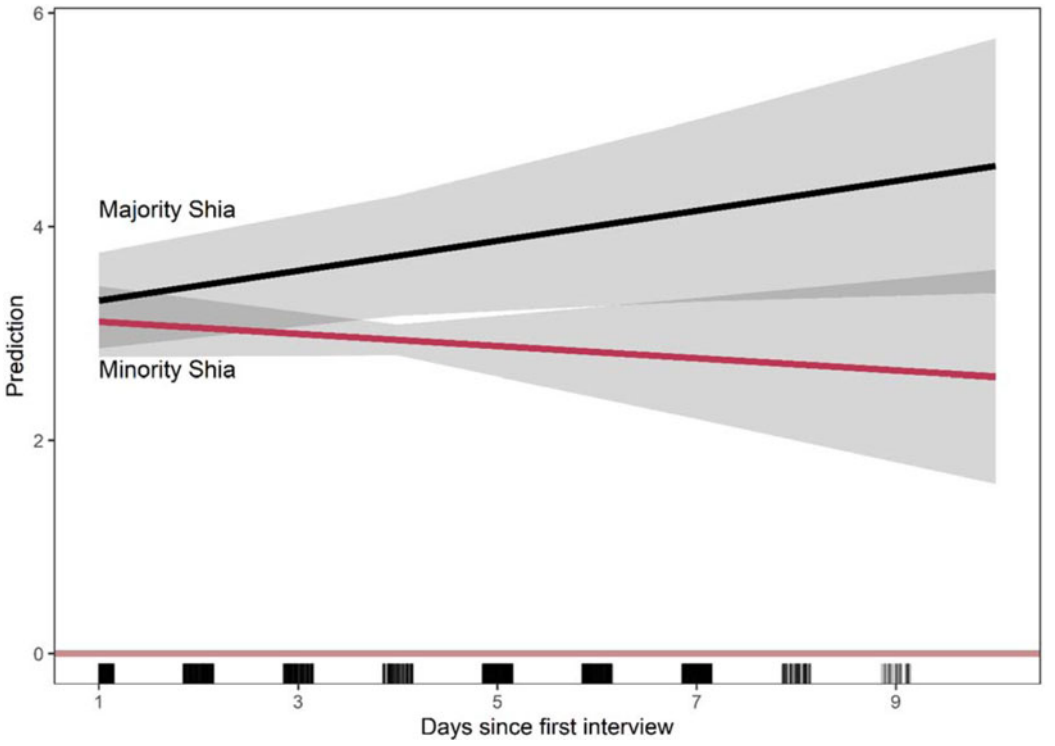


Figure 4. Effect of Days on Hizballah Support in Shi'a versus Non-Shi'a Neighborhoods (Corstange, 2016).

Hizballah. In Shi'a neighborhoods, however, responses move in a positive, pro-Hizballah direction as data collection within each neighborhood progresses.

The Corstange study also includes an experimental treatment aimed at eliciting the effects of a Hizballah antagonist, the United States, in an unobtrusive manner. The treatments—embedded in enumerator introductions—expose respondents to a variety of putative sponsors, including academic and governmental, and the United States and non-U.S. sponsors.[21] Among the 2100 respondents who consented to the survey, 381 individuals were assigned to the control group, while 434 individuals were told that the survey was conducted on behalf of "the American Embassy." Other treatments included "the University of Maryland" and "the Canadian Embassy."[22] Again, the outcome variable is a composite index that captures local residents' support for Hizballah.

We have no theoretical expectations regarding the starting point of responses at the outset of data collection. However, if a grapevine effect exists in the data, we should expect to see that both treated and control responses from Shi'a neighborhoods tilt in the pro-Hizballah direction as data collection progresses (because both respond to the same outcome measure of Hizballah support), and that those exposed to the United States treatment display an even stronger tendency in this direction after having been cued to this antagonist of the incumbent local power.

The limited sample size for the experimental component of the study does not leave us room to detect statistically significant results, but Figure 5 illustrates a pattern in keeping with a potential grapevine effect. In it, we plot OLS regression results for the effects of time on Hizballah support among treated and control respondents living in Shi'a and non-Shi'a neighborhoods.

[21]See Appendix Table A6 for the treatment script.

[22]For the purposes of this paper, we include only the control group and the group treated with "American Embassy" in our analysis.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

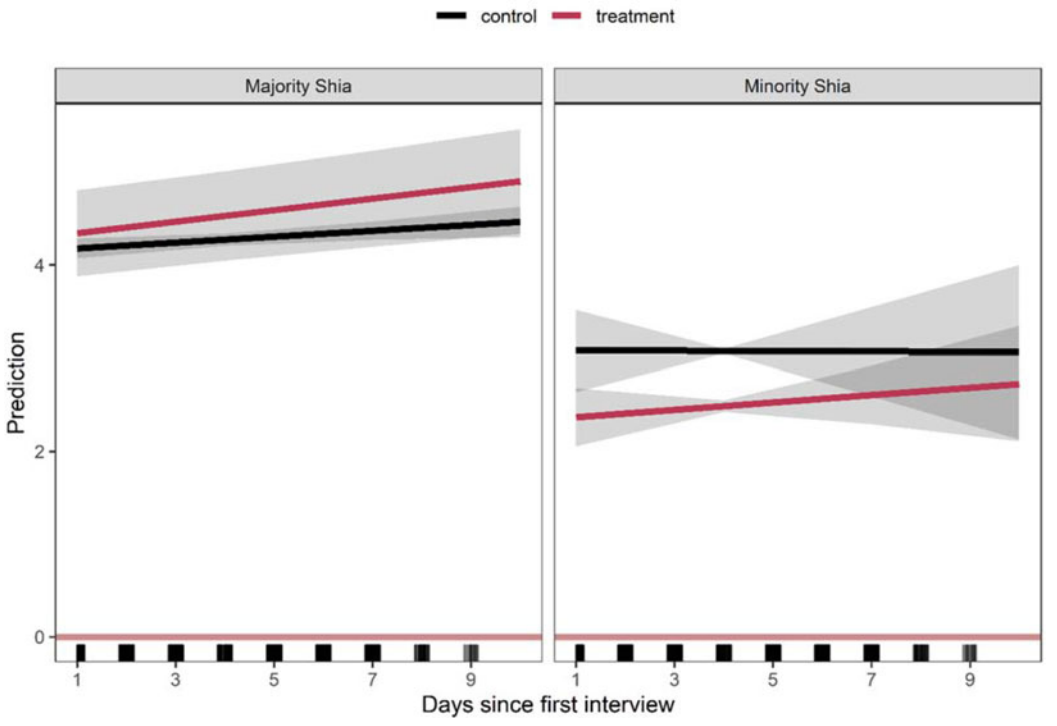


Figure 5. Plotted Experimental Results in Shi'a and Non-Shi'a Neighborhoods (Corstange, 2016).
Note: Standard errors clustered by neighborhood.

The top two lines represent control and treatment responses in the Shi'a-majority areas. They indicate that support for Hizballah gradually rises among the control group and that the responses of treated participants increase even more starkly in this direction. The bottom two lines represent the control and treatment responses in non-Shi'a neighborhoods; interestingly, while the control group respondents remain steady in their (clearly inferior) support for Hizballah, evidence of increasing support for Hizballah over the course of data collection among those treated with the American cue again emerges, perhaps as fear of Hizballah reprisal seeps even into those neighborhoods.

Corstange (2016) shows that, in the context of systematic opting-out of survey participation by some respondents, purported U.S. sponsorship makes people in Beirut more supportive of U.S.-favored policies and less supportive of Hizballah interests. Our grapevine analysis additionally suggests that, while the average treatment effect remains robust, responses in Shi'a neighborhoods gradually tilt in a relative pro-Hizballah direction as data collection progresses. Corstange's conclusion may thus be considered a conservative one, as it persists despite the apparent tendency of some respondents to alter their responses out of fear of the local Shi'a power.

5. Conclusion

The study of sensitive research topics, such as support for violent extremism, is often put at risk by respondent's inability or unwillingness to provide truthful answers. Even when subtle, experimental measures are employed, the research still typically takes place on-the-ground and amongst populations keen to understand the purpose and to respond appropriately.

In this study, we described a pattern previously overlooked in research on sensitive topics: as data collection on violent extremism proceeds, community members who respond later in the

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

data collection process tend to give answers that differ systematically from earlier ones, tilting in the direction of support for the local incumbent power holder. We first document that effect in responses to overt survey questions in two different surveys, each with over 7700 respondents, conducted in Burkina Faso, Chad, and Niger. The results suggest a shift of 30 percentage points or more away from support for extremist violence and thus toward the government's preferred view. We further demonstrate that even a subtle measure of extremist support using an endorsement experiment is susceptible to systematic changes as data collection in the locality progresses, appearing in one of our two surveys. We interpret these findings as indicating steadily increasing expressions of support for the formal government and security forces of these countries as time elapses. We find consistent evidence not just in our original surveys that elicit attitudes regarding AQIM in the Sahel, but also from a separate study based in Lebanon (Corstange, 2016), which differs from our own surveys in terms of the implicated Islamist group, the type of experimental treatment, and the entity with de facto local power.

We suggest that the most plausible mechanism driving these results is a grapevine effect that results in the gradual emergence of a pro-incumbent power bias among respondents that we attribute to a fear of reprisal. Participants in population-based surveys, especially in rural and developing areas, tend to share details of the data collection with neighbors, and the accumulated knowledge likely raises suspicions regarding even subtle and nuanced treatments. Participants may also suspect that their responses are not always kept confidential, so as more community members find out about a sensitive survey and then take part themselves, their responses may tilt in the favor of the de facto local power for fear of retribution for deviant or uncooperative responses. Alternatively, the pattern we have identified could be the result of genuine attitude changes rather than strategically manipulated ones, perhaps as local powers begin to engage in more salutary behavior vis-à-vis community members over the course of data collection in each commune. We view this explanation as less likely,[23] though we leave the testing of mechanisms for other studies.

The study raises other possibilities for future research, as well. First, this study focused on one sensitive topic, support for violent Islamist groups. Other studies might examine changes in response patterns over the course of data collection in studies related to government versus opposition support, attitudes toward foreigners and immigrants, and other sensitive topics. It is worth reiterating, however, that we anticipate grapevine effects principally when fear of reprisal for insubordinate responses exists. Second, with more such studies, researchers will be in a position to evaluate the time it takes for response patterns to change in favor of the local incumbent power. We suspect that it will be consistent with the time it takes for news to travel through the grapevine across a local community. Finally, we do not expect that the grapevine effect would persist beyond the life of one enumeration of surveys, but researchers could explore the temporal legacy of a grapevine effect with follow-on surveys at systematically different intervals following the first.[24]

The findings suggest three important lessons for researchers studying sensitive topics through population-based data collection, using both overt and experimental methods. One is to be mindful of information sharing and participants' perceptions of local power holders. For example, where culturally appropriate, the study may represent a call to more stridently insist on privacy in the conduct of individual interviews with participants. Second, if indeed subtle, experimental measures provide a more precise estimation of true attitudes, this would suggest that earlier

[23]We evaluated patterns in both government and extremist violence over the course of our survey periods, using data from ACLED; we found no pattern that would explain a change in the behavior of authorities. Further, their efforts to change behavior would have to be carried out at different times in different communes, according to the data collection schedule.

[24]Researchers are increasingly attuned to survey fatigue across multiple surveys in heavily enumerated areas. If this phenomenon alters responses on subsequent surveys, so too could fears of reprisal regarding sensitive topics, though the grapevine effect that we note may be supplanted by broader pro-incumbent power support in that case. We thank an anonymous reviewer for this insight.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

responses are capturing attitudes before strategic gaming and the fear of retribution set in and adulterate the very attitudes we seek to measure. Data collection using experimental measures certainly appears to mitigate the challenge to some degree. Finally, more enumerators working over a shorter period of time should constitute an additional advantage over more drawn out survey work on sensitive topics like violent extremism. Researchers face a tradeoff in recruiting larger enumeration teams, which can pose challenges in terms of quality control, but they would have the added confidence that word-of-mouth would not undermine the validity of their findings on sensitive research topics.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psrm.2020.34

Acknowledgements. We thank Trey Billing, Jushua Kalla, and anonymous reviewers for their feedback and input. Data collection was supported by the United States Agency for International Development (USAID/West Africa) Evaluation & Analytical Services (EAS) Project.

References

Adida CL, Ferree KE, Posner DN and Robinson AL (2016) Who's asking? Interviewer coethnicity effects in African Survey Data. *Comparative Political Studie* **49**, 1630–1660.

Alexander Y (2012) Special update report: terrorism in North, West, & Central Africa: from 9/11 to the Arab Spring. *International Center for Terrorism Studies* **3**, 1–52.

Antwi-Boateng O (2017) The rise of pan-islamic terrorism in Africa: a global security challenge. *Politics & Policy* **45**, 253–284.

Atran S, Axelrod R, Davis R and Fischhoff B (2017) Challenges in researching terrorism from the field. *Science (New York, N.Y.)* **355**, 352–354.

Berinsky A (2004) Can we talk? Self-presentation and the survey response. *Political Psychology* **25**, 643–659.

Blair G, Fair CC, Malhotra N and Shapiro JN (2013) Poverty and support for militant politics: evidence from Pakistan. *American Journal of Political Science* **57**, 30–48.

Blair G, Imai K and Lyall J (2014) Comparing and combining list and endorsement experiments: evidence from Afghanistan. *American Journal of Political Science* **58**, 1043–1063.

Blaydes L and Gillum RM (2013) Religiosity-of-interviewer effects: assessing the impact of veiled enumerators on survey response in Egypt. *Politics and Religion* **6**, 459–482.

Bullock W, Imai K and Shapiro JN (2011) Statistical analysis of endorsement experiments: measuring support for militant groups in Pakistan. *Political Analysis* **19**, 363–384.

Corstange D (2014) Foreign-sponsorship effects in developing-world surveys: evidence from a field experiment in Lebanon. *Public Opinion Quarterly* **78**, 474–484.

Corstange D (2016) Anti-American behavior in the Middle East: evidence from a field experiment in Lebanon. *The Journal of Politics* **78**, 311–325.

Douglas M (2015) Their evil lies in the grapevine effect: assessment of damages in defamation by social media. *Media and Arts Law Review* **20**, 367–379.

Davis DW and Silver BD (2003) Stereotype threat and race of interviewer effects in a survey on political knowledge. *American Journal of Political Science* **47**, 33–45.

Fair CC, Littman R and Nugent ER (2018) Conceptions of sharia and support for militancy and democratic values: evidence from Pakistan. *Political Science Research and Methods* **6**, 429–448.

Finkel SE, Guterbock TM and Borg MJ (1991) Race-of-interviewer effects in a preelection poll Virginia 1989. *Public Opinion Quarterly* **55**, 313–330.

Freedman DA (2008) Randomization does not justify logistic regression. *Statistical Science* **23**, 237–249.

Fugii LA (2012) Research ethics 101: dilemmas and responsibilities. *PS: Political Science and Politics* **45**, 717–723.

Goldsmith A (2005) Police reform and the problem of trust. *Theoretical Criminology* **9**, 443–470.

Goodson LP (2001) Perverting Islam: Taliban social policy toward women. *Central Asian Survey* **20**, 415–426.

Grossman G, Gazal-Ayal O, Pimentel SD and Weinstein JM (2016) Descriptive representation and judicial outcomes in multiethnic societies. *American Journal of Political Science* **60**, 44–69.

Hainmueller J and Hopkins DJ (2015) The hidden American immigration consensus: a conjoint analysis of attitudes toward immigrants. *American Journal of Political Science* **59**, 529–548.

Hershfield AF, Rohling NG, Kerr GB and Hursh-Cesar G (1983) Fieldwork in rural areas. In Bulmer M and Warwick DP (eds), *Social Research in Developing Countries: Surveys and Censuses in the Third World*. London: Routledge, pp. 241–252.

Holtgraves T (2004) Social desirability and self-reports: testing models of socially desirable responding. *Personality and Social Psychology Bulletin* **30**, 161–172.

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press

Johnson WT and DeLamater JD (1976) Response effects in sex surveys. *Public Opinion Quarterly* **40**, 165–181.
Kalinin K (2016) The social desirability bias in autocrat's electoral ratings: evidence from the 2012 Russian presidential elections. *Journal of Elections, Public Opinion and Parties* **26**, 191–211.
Krumpal I (2013) Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity* **47**, 2025–2047.
Lee T and Pérez EO (2014) The persistent connection between language-of-interview and Latino political opinion. *Political Behavior* **36**, 401–425.
Lyall J, Zhou Y-Y and Imai K (2020) Can economic assistance shape combatant support in wartime? Experimental evidence from Afghanistan. *American Political Science Review* **114**, 126–143.
Marchal R (2016) An emerging military power in Central Africa? Chad under Idriss Déby. *Sociétés politiques comparées* **40**, 1–20.
McCauley JF (2014) Measuring and reducing religious bias in post-conflict zones: evidence from Côte d'Ivoire. *Political Psychology* **35**, 267–289.
Mohammed K (2014) The message and methods of Boko Haram. In Pérouse de Monclos M-A (ed.), *Boko Haram: Islamism, Politics, Security and the State in Nigeria*. Leiden, Netherlands: African Studies Centre, pp. 9–32.
Rathgeber E (1996) Women, men, and water-resource management in Africa. In Rached E, Rathgeber E, and Brooks DB (eds), *Water Management in Africa and the Middle East: Challenges and Opportunities*. Ottowa: International Development Research Centre, pp. 49–69.
Robinson D and Tannenberg M (2019) Self-censorship of regime support in authoritarian states: evidence from list experiments in China. *Research and Politics* **6**, 1–9.
Rosenfeld B, Imai K and Shapiro J (2016) An empirical validation study of popular survey methodologies for sensitive questions. *American Journal of Political Science* **60**, 783–802.
Schils N and Pauwels L (2014) 2–14. "explaining violent extremism for subgroups by gender and immigrant background, using SAT as a framework.". *Journal of Strategic Security* **7**, 27–47.
Schreiber D, Fonzo G, Simmons AN, Dawes CT, Flagan T, Fowler JH and Paulus MP (2013) Red brain, blue brain: evaluative processes differ in democrats and republicans. *PLoS One* **8**, e52970.
Singer E, Hippler H-J and Schwarz N (1992) Confidentiality assurances in surveys: reassurance or threat? *International Journal of Public Opinion Research* **4**, 256–268.
Technical Assistance Research Programs (TARP) (1981) Measuring the grapevine consumer response and word-of-mouth, US Office of Consumer Affairs, Washington, DC.
Tourangeau R and Yan T (2007) Sensitive questions in surveys. *Psychological Bulletin* **133**, 859–883.
UNHCR (2017) Nigeria Situation: Populations Forcibly Displaced in the Lake Chad Region. 13 January 2017. Available at https://reliefweb.int/sites/reliefweb.int/files/resources/nga_situation_A4L_2017-01-23.pdf.
Wray R (2012) *Fighting Talk: What Goes on Before the Bomb Goes Off? A Study of the Causal Factors that Influence Propensity to Violent Behavior* (Ph.D. disseration), University of London.

Cite this article: McCauley J, Finkel S, Neureiter M, Belasco C (2022). The grapevine effect in sensitive data collection: examining response patterns in support for violent extremism. *Political Science Research and Methods* **10**, 353–371. https://doi.org/10.1017/psrm.2020.34

https://doi.org/10.1017/psrm.2020.34 Published online by Cambridge University Press