Article

Political Research Quarterly
2023, Vol. 76(1) 173–185

DOI: 10.1177/10659129211073910
journals.sagepub.com/home/prq
SAGE

Does Familiarity Breed Esteem? A Field Experiment on Emergent Attitudes Toward Members of Congress

Kevin M. Esterling[1], William Minozzi[2] and Michael A. Neblo[2]

Abstract

Canonical theories of democratic representation envision legislators cultivating familiarity to enhance esteem among their constituents. Some scholars, however, argue that familiarity breeds contempt, which if true would undermine incentives for effective representation. Survey respondents who are unfamiliar with their legislator tend not to provide substantive answers to attitude questions, and so we are missing key evidence necessary to adjudicate this important debate. We solve this problem with a randomized field experiment that gave some constituents an opportunity to gain familiarity with their Member of Congress through an online Deliberative Town Hall. Relative to controls, respondents who interacted with their member reported higher esteem *as a result of enhanced familiarity*, a mediation effect supporting canonical theories of representation. This effect is statistically significant among constituents who are the same political party as the member but not among those of the opposite party, although in neither case did familiarity breed contempt.

Keywords

representation, familiarity, survey research, causal mediation

In modern democracies, representation and accountability depend on two crucial—yet potentially conflicting—features of constituent attitudes: familiarity and esteem. Constituents must be familiar with officeholders to hold them accountable for their actions (Arnold, 1990; Grant and Rudolph, 2004). Simultaneously, the drive to be esteemed—that is, to be held in higher regard, reflected in higher trust, approval and warmer feelings among constituents—is essential to incentivize faithful representation (Fenno, 1978; Madison, 1961). Some scholars worry these two elements are incompatible—that familiarity, as the saying goes, breeds contempt (Brady and Theriault 2001; Hibbing 2002; see Mondak et al. 2007, 35). If that were the case, representatives might be reluctant to become more familiar to wide swaths of their constituents, contributing to the persistent dearth of knowledge about elected officials, and damaging canonical accounts of democratic representation (Mansbridge, 2003; Pitkin, 1967).

The idea that familiarity breeds contempt is not uncontested in the scholarly literature, however. Prominent studies of the U.S. Congress suggest that, as constituents learn more about their member of Congress (MC), they also like them more (Alvarez, 1997; Fenno, 1978; Fiorina et al., 1987). Indeed, even though Americans routinely report low approval of Congress as an institution, they also typically report much higher approval of their own MC—a well-established finding known as the Paradox of Congressional Support (Lammers et al., 2021). Thus, it could be that individual MCs do not face the tradeoff between familiarity and esteem suggested by other empirical accounts.

Scholars have had trouble resolving this debate because unfamiliarity presents a difficult measurement problem. Esteem is typically measured with survey items such as approval of and trust in the member. Constituents who are unfamiliar with their representative are more likely to skip such questions or select a "Don't Know" (*DK*) response. Because of this, our understanding of whether Americans typically love or typically despise their MC is based on incomplete information. Drawing

[1]Public Policy and Political Science, UC – Riverside, Riverside, CA, USA
[2]Political Science, Ohio State University, Columbus, OH, USA

Corresponding Author:
Kevin M. Esterling, Public Policy and Political Science, UC – Riverside, 900 University Ave, Riverside, CA 92521, USA.
Email: kevin.esterling@ucr.edu

conclusions based on survey responses depends crucially on how one handles the missing evidence of these reported attitudes, as well as one's assumptions about any confounding in the statistical relationship between survey respondents' self-reported familiarity and esteem.

But nonresponse due to unfamiliarity evokes a deeper, more foundational problem. If "attitudes" just are the things measured by survey responses, then nonresponse does not merely imply missing evidence. Instead, a *DK* response signifies the absence of any attitude at all (Norpoth and Lodge, 1985). In this case, the relationship between familiarity and esteem verges on nonsensical—those who are unfamiliar have no attitudes, so the functional correspondence between esteem for the familiar and for the unfamiliar is simply undefined (Pierce and Rose, 1974). Put another way, familiarity could not cause a change toward either contempt or esteem as a theoretical matter, because the outcome under conditions of unfamiliarity does not exist.

We propose a resolution to this problem, and use our solution to provide new evidence on the relationship between familiarity and esteem. We conceptualize constituents' attitudes toward their members as *latent affect*, rather than the survey responses themselves. Our theoretical approach to non-attitudes countenances the notion that respondents possess a latent affect even when they are unfamiliar with a particular elected official. We use a construct measurement approach (Bollen, 1989) to uncover constituents' latent affect and its relationship with familiarity within the constituent–legislator relationship.

To fill the evidentiary gap relating familiarity and esteem, we rely on evidence gathered from a field experiment. In the experiment, we rely on representative samples of constituents in 12 congressional districts who were randomly assigned invitations to interact directly with their currently sitting MC via an online Deliberative Town Hall (Neblo et al., 2018). Because it creates an exogenous opportunity to *become* familiar with one's MC, this experiment gives us the means to estimate the causal relationship between familiarity and latent affect—provided that we have plausible measures of each. The experiment provides evidence for the causal relationship between familiarity and affect driven by the kinds of outreach efforts Fenno (1978) describes as routine parts of representational activity, which increasingly are hosted in online settings. Other endogenously induced causes of familiarity—for example, online attack ads—are likely to yield different results of course. We discuss this distinction in the conclusion.

Our analysis proceeds in two steps. First, we construct a measure of familiarity with one's MC, and in so doing, warrant the claim that constituent unfamiliarity is an important driver of *DK* responses on attitude questions about them (Krosnick and Milburn 1990). Theoretically, a *DK* response is the result of many factors (Berinsky 1999, 1214; Berinsky and Tucker 2006, 74). We estimate a measurement model to evaluate the construct validity of the familiarity measure that is recovered from *DK* responses making use of a pretest survey, and show that *DK* responses on attitude questions about MCs is specifically driven by a systematic lack of familiarity rather than omitted factors.

Second, we estimate the effect on latent effect of attending an online Deliberative Town Hall. Importantly, we treat familiarity as a *mediator* between exposure to the member in the town hall and the respondent's level of esteem for the member (c.f., Lammers et al., 2021, 8). Here, we measure latent affect using responses to items relevant to the respondent's esteem of the member, including approval, trust, vote intention, and a feeling thermometer (Wilcox et al., 1989; Winter and Berinsky, 1999). To identify causal effects we implement a causal mediation model within the generalized endogenous treatment (GET) causal framework (Esterling et al., 2011a). The GET model serves two purposes: it identifies causal effects in the presence of noncompliance and nonresponse in our field experimental data, and it enables a parametric solution to the well-known problem of sequential ignorability in causal mediation analysis (Imai et al., 2011).

We draw several conclusions from this analysis. Attendance at online Deliberative Town Halls dramatically enhanced constituents' familiarity, and thus their propensity to report an attitude on the esteem measures. This increase in familiarity translated into increases in latent affect, indicating that familiarity tends to breed more positive attitudes on the esteem items. Moreover, the total change in latent affect was due almost entirely to enhanced familiarity with the member. Finally, while these findings were driven primarily by changes in latent affect for constituents who identify with the political party of their MC, we found no evidence that familiarity diminished latent affect for constituents who identify with the opposite party (similar to Broockman and Butler, 2017). Together, our findings suggest that members of Congress need not fear that deeper familiarity comes at a cost, and that these twin pillars of representation—familiarity and esteem—can be mutually reinforcing not only in theory but also in the practice of democracy.

Familiarity, Esteem, and Overgeneralization

Political scientists are divided in their expectations about the effect of familiarity on attitudes toward members of Congress (MCs). Some assert that the Americans who hold considerations in mind regarding their MC tend to hold the MC in low regard. For example, Hibbing and

Theiss-Morse (1995) and Hibbing (2002) report findings from surveys and focus groups showing that many U.S. citizens hold negative affect toward members of Congress because of their beliefs that elected officials are self-serving and disconnected from the interests of ordinary people. If these beliefs are true, then confirming them with increased familiarity will lead to cynicism, distrust, and overall colder attitudes toward members of Congress.

In his landmark book *Homestyle*, however, Fenno (1978) demonstrates that many MCs at least believe that constituents will hold warmer attitudes toward them as they gain familiarity, particularly as the member herself engages in outreach and engagement among constituents. Box-Steffensmeier et al. (2003), Cover and Brumberg (1982), Fiorina et al. (1987), King (1991) and Lammers et al. (2021) also find evidence that constituent casework and outreach foster positive affect and incumbent advantages. Similarly, Alvarez (1997) shows that constituents dislike the uncertainty that stems from unfamiliarity, leading to less favorable attitudes, and so increasing familiarity should curtail such displeasure.

Or, perhaps neither side is exactly right; increasing familiarity might simply foster ambivalence in constituents (Turgeon 2009, 354; Visser et al. 2007, 135), in which case familiarity might have nuanced effects on attitudes.

Part of the reason the issue remains unsettled is that different authors mean different things by the term "familiarity." Whereas Fenno (1978) uses the term to denote a personalized rapport with constituents in the district (as does Fiorina et al., 1987; Grimmer et al., 2012), others use it to mean factual and non-relational knowledge often rooted in the representative's activity in Washington, D.C. (Delli Carpini and Keeter, 1996). While these definitions center on different aspects of the concept of representation, at a deeper level they have an important and irreducible intersection. Both meanings of familiarity require that given the specific attention and engagement of a constituent, she would be able to form and retrieve attitudes that she can report about her representative; without such attention, she would not be able to form such attitudes. We therefore take this intersection as the working definition of familiarity: having either relational or non-relational knowledge of the representative enhances the formation of attitudes, while having both kinds of knowledge enhances attitude formation the most.

Even with this definition in hand, a resolution to the dissensus still faces a significant methodological and conceptual hurdle: that in general, the lack of familiarity will systematically yield missing data. Respondents with sufficiently low familiarity will simply have no attitude regarding the survey measure. As such, it is generically incoherent to discuss differences in potential outcomes based on differences in familiarity, at least insofar as the outcome is supposed to be measured in a survey. In practical terms, we would—by definition—be able to measure an outcome only when the respondent is familiar; when she is sufficiently unfamiliar, no such measurement will be possible. Hence, the functional relationship between these potential outcomes will be similarly undefined.

We resolve this problem by arguing that when constituents lack specific knowledge about their MC, they overgeneralize their judgment of the member (Firestone and Scholl, 2015), and thereby develop negative latent affect for them. As Mondak et al. (2007) show, constituents who are poorly informed about Congress base their appraisals of the institution on considerations that are peripheral to Congress but happen to be cognitively accessible. In the case of individual members, constituents may take unfamiliarity with the member herself or with her actions in D.C. to demonstrate that the personal relationship implied in the concept of representation is broken (Fenno, 1978; Fiorina et al., 1987), and fill in the gap with negative existing stereotypes about politicians (Lammers et al., 2021; Neblo et al., 2010).

Constituents need not be consciously aware of this overgeneralization in order for unfamiliarity to lead to negative affect, and the presence of familiarity to positive affect. Instead, two general cognitive mechanisms can account for this pattern, whether or not the respondent makes the connection explicitly. First, uncertainty aversion creates negative affect, separate from preferences (Alvarez, 1997; Lee, 2001). Second, the "mere exposure" phenomenon from social psychology suggests that increases in familiarity will typically result in more positive attitudes (Winkielman and Cacioppo, 2001). The mere exposure effect holds that benign exposure to an object induces positive latent affect toward that object (Harmon-Jones and Allen, 2001; Zajonc, 2001). The mere exposure effect suggests that, upon a constituent's overt exposure to her MC, the result is likely to be an improvement in esteem (Grimmer et al., 2012).

Given the pervasive role of party identification in American politics, we expect to see differences in these effects of familiarity based on whether a constituent identifies as a member of the same party as her MC (Grant and Rudolph, 2004). Building on our conjecture that constituents overgeneralize with unfamiliar MCs, the effect of familiarity should depend on party cues. But we should observe familiarity increase esteem even among constituents who identify with a different party than the member. In contrast, if exposure only triggers partisan evaluations, then upon gaining familiarity with one's MC, same-party constituents would have more positive affect while opposite-party constituents would have more negative affect—that is, familiarity would have asymmetric effects.

Of course, the mere exposure and uncertainty effects might not apply well to the political context. Constituents may become even more dissatisfied with politicians the more they become familiar with them (Hibbing, 2002). On this view we should expect to see a negative relationship between familiarity and latent affect among all respondents.

Study Design, Measurement, and Causal Mediation

Our goal is to estimate the causal relationship between familiarity and esteem. To do so, we leverage evidence from a randomized field experiment, justify our measurement strategy for familiarity, and apply a causal mediation model. This study relies on randomly assigning some respondents an invitation (i.e., an encouragement) to interact with their own MC at an online Deliberative Town Hall. Attendance, of course, should increase familiarity with the MC. But our main interest is *not* the direct effect of attendance on attitudes (for which, see Minozzi et al., 2015). Rather, we seek to resolve the measurement dilemma given that most citizens are unfamiliar with their MC *ex ante* (Delli Carpini and Keeter, 1996). To solve this problem, we must estimate the *mediated effect* of attendance—through its increase in familiarity—on attitudes toward the member.

Study Design

The data come from a large field experiment conducted in the summer of 2006, which held a series of 20 online Deliberative Town Halls in 12 congressional districts. The 12 members of Congress (MCs) were all running for reelection; 5 were Republicans, 7 Democrats; the districts were drawn from each region of the country. Table 1 lists the members and congressional districts that were included in the study. Below we discuss issues with temporal validity of using data from 2006, given increasing polarization in the electorate[1] and given the changing relationship of the electorate to communication technology (Settle 2018). In short, there is no evidence that the magnitudes of the effects of deliberation interventions such as the one we examine have diminished in recent years (e.g., Abernathy, 2018), and the platform used was not vulnerable to pathologies recently observed in social media platforms (as in Lazer, 2015).

The Deliberative Town Halls involved moderated discussion between the MC and the group of constituents. Each group had between 8 and 40 constituents who typed questions and comments into a text box, which were in turn posted to a queue only visible to the study team. A moderator posted questions and comments from the queue sequentially into the town hall in a separate text box visible to the member as well as all participants, and the member responded orally. The moderator filtered questions if they were redundant or off topic, and to ensure that everyone had an opportunity to post a question or comment before getting a second question posted. The moderator did not filter on substance. Each session lasted 35 min. At the conclusion, the member and her staff left the event and the constituents were given 25 additional minutes to discuss the session with each other via a text-based chat room.

Within each district in the study, the online polling firm Knowledge Networks (KN, now GfK) recruited a representative sample of constituents from each district and randomized the constituents to one of three conditions.[2] In

Table 1. Field experimental sample.

Member	District	Total	In sample	Invited	Attended
James Clyburn	SC-06	28	18	14	8
Zoe Lofgren	CA-16	116	28	24	15
Jim Matheson	UT-02	55	29	22	12
Donald Manzullo	IL-16	75	43	25	15
Michael Capuano	MA-08	195	50	40	22
David Price	NC-04	187	60	49	32
George Radanovich	CA-19	246	61	31	10
Anna Eshoo	CA-14	109	72	56	36
Jack Kingston	GA-01	221	81	59	27
Dave Weldon	FL-15	231	100	76	40
Mike Conaway	TX-11	275	104	75	28
Earl Blumenauer	OR-03	334	114	38	17
Total		2072	760	509	262

"Total" refers to all constituents assigned to one of the three experimental conditions. "In Sample" includes only respondents who completed the background materials survey. "Invited" indicates number randomly assigned to be invited to event. "Attended" indicates number who complied with invitation.

the deliberative group (DG) condition, constituents were first administered a pretest survey and subsequently were provided reading material on the topic of the session (U.S. immigration policy) prior to the session, invited to attend the deliberative online town hall, and then administered a post-test survey about one week after the town hall. In the information only (IO) group, constituents were administered a survey before the town hall, provided the reading material, and then surveyed after the town hall, but were not invited to attend the town hall itself. In the true control condition, constituents were asked to respond to the pretest and post-test surveys, but received no reading material and were not invited to attend the town hall.

The Deliberative Town Hall sessions are described qualitatively in (Neblo et al., 2018, chapters 5–6). The authors report that each of the sessions centered on extensive, substantive discussions of U.S. immigration policy and met many ideals for deliberative engagement. For example, out of over 1000 questions and comments from constituents, not a single one needed to be screened for vulgar or abusive language. Members in each session were observed to openly disagree and counterargue with constituents and to explain their positions. The authors found that the sessions had ratings on the Discourse Quality Index (Steenbergen et al., 2003) that exceeded those of European parliaments. They found equality of participation across gender, education, age and race, in both the number of questions submitted and the members' responsiveness. And on the follow up survey, nearly all constituents rated the sessions as informative and valuable for democracy.

To enhance the deliberative design of the town hall, each session with the member was followed by a post-session, text-based chat among the constituents present. Because the chat was combined with the session it is not possible to identify the causal effect of each of these components on familiarity and esteem; instead, the treatment must be understood as the two components combined. The chat transcripts show that constituents were generally positive about their experience; a simple count of positive and negative statements showed that 14% were negative and 86% positive, and virtually, all of the negative statements were that the sessions were too short. In addition, constituents used the opportunity to discuss the pros and cons of the member's substantive remarks, and so the post-session chat helped constituents strengthen and deepen their beliefs about, and so familiarity with, the member.

The appendix describes the recruitment, assignment, survey administration, variables and statistical models. In total 2072 constituents were assigned to the three experimental conditions. Throughout this paper, we focus on the subsample of 760 respondents who were assigned to either the DG or the IO condition *and* who responded to a survey indicating they read the background materials. We make this restriction because those who select into this subsample are more comparable to each other, in that they all were willing to read material and comply with that component of the study.[3]

And beyond that, given our study design, including respondents who did not read the material in the analysis would introduce a significant confound. Our interest is in identifying the effect of exposure to a member in a town hall on the respondent's familiarity and esteem, not the effect of reading the background materials that were unrelated to the member. Since every respondent who attended a town hall also read the background materials, comparing those who read the material and attended the town hall to those who did neither does not identify the counterfactual of interest. Instead, among those who read the material, comparing those who those who did to those who did not attend a town hall does identify the counterfactual of interest, given the randomization and the causal statistical model we describe below.

Using "Don't Knows" to Measure Familiarity

To evaluate the effects of familiarity, we first need to measure it. Some of our survey items contain "Don't Know" (*DK*) as a response option, and so the participant can choose to respond *DK*, or give a substantive response that we indicate with *DK* meaning "Not DK." To construct our measure of familiarity, we assume that unfamiliarity is the primary driver of *DK* responses on attitude items regarding the member, and hence, items containing *DK* response categories can serve as item response opportunities to measure familiarity. The less knowledge one has regarding the attitude object, the more cognitive effort is required to issue a response (Krosnick, 2002, 89; Krosnick and Milburn, 1990). Unfamiliar respondents must exert greater cognitive effort to render summary judgment on attitude questions—retrieving relevant considerations about the member's personal qualities, performance in D.C., or both, from memory and integrating these considerations into responses (Berinsky and Tucker, 2006, 76; Krosnick, 2002, 94; Turgeon, 2009, 354).

Hence, a lack of familiarity is a primary reason for a *DK* response (Delli Carpini and Keeter, 1996; Visser et al., 2007, 128); that is, we assume "Don't Know" simply means "Don't Know" (Jessee, 2017; Luskin and Bullock, 2011).

Thus, for purposes of this study, we take the empirical propensity to provide attitude responses on a battery of items about the member, when the response sets include *DK* as an option, as indicators of familiarity. While an epistemic lack of familiarity is a primary driver of the *DK* response, it need not be the only driver and indeed there may be confounding factors that reduce the construct validity of this measure (Atir et al., 2015; Mondak, 1999). In Supplemental Appendix A.4 we report an investigation into the construct validity of our measure of member familiarity by comparing this measure with similar measures of familiarity toward two additional objects, policy preferences and policy knowledge. Our construct validity model uses *DK* responses toward each of these three objects—regarding policy preferences, regarding policy knowledge, and regarding the member—to construct a measure of familiarity toward each, which we designate as η_1, η_2, and η_3, respectively, and decomposes the three measures into a specific component and a systematic component, which we designate as ξ_1. We find, among other results, that 96% of the member familiarity measure (η_3) is attributable to the systematic component (ξ_1), indicating the absence of confounding or extraneous causes of *DK* response toward the member other than familiarity. In addition, we find few covariates that predict familiarity, other than interest in political news reporting, which further underscores the lack of confounds lurking in our familiarity measure.

Causal Mediation Model

Given this experimental setting and our measurement strategy for familiarity, we use a mediation model and the identification strategy recommended in Imai et al. (2011) to identify causal effects. For estimation we use a structural equation approach (Bollen, 1989) because familiarity and esteem both are endogenous latent variables that require measurement.[4]

Figure 1 presents the model. Observed variables are indicated with rectangles; latent variables with ovals. Arrows show variables assigned to equations. In this figure, the intervention is exposure to a Deliberative Town Hall with the MC under the randomized invitation to attend; $D = 1$ if the respondent attended and $D = 0$ if she did not attend. We measure post-treatment *familiarity* with the member as η_4, using each constituent's propensity to offer a *DK* response on a set of member attitude items rather than *DK* (skips are set to missing for each item); we measure the respondents' post-treatment latent *esteem* for the member with η_5 using their responses to a battery of items measuring their attitudes toward the member such as trust and approval; and ξ_2 is a measure of the respondent's propensity for *compliance* with experimental activities.[5] The **X**'s indicate that each latent variable is conditioned on the covariates **X** listed in footnote 28 and described in Supplemental Appendix A.2. Full details, including code, appear in Supplemental Appendix A.5.

The main estimand of interest is the average causal mediation effect (ACME), or the marginal effect of treatment exposure on esteem that occurs because of change in familiarity. Imai et al. (2010) show the ACME estimand is

$$\mathrm{ACME}_d = \mathrm{E}_S[\eta_5(d,\eta_4(1)) - \eta_5(d,\eta_4(0))] \quad \text{for} \quad \mathrm{d} = 0,1 \tag{1}$$

where the expectation is taken over the sample S and subscripts are omitted. In equation 1, parentheses indicate the *counterfactual* values of each variable. To unpack equation (1), the variable $\eta_4(1)$ represents the counterfactual level of familiarity, measured by η_4, when the respondent is exposed to the Deliberative Town Hall, indicated by 1 within the parentheses. Likewise $\eta_4(0)$ is the level of familiarity when not exposed. Next we note that the *ACME* estimand is defined for two cases, one case where the respondent is exposed to the town hall, $d = 1$, and one where the respondent is not exposed, $d = 0$. Next, we define $\eta_5(1,\eta_4(1))$ as the respondent's level of esteem having been exposed to the town hall and having her familiarity set to the counterfactual under exposure to the town hall, and $\eta_5(1,\eta_4(0))$ as the respondent's level of esteem having been exposed to the town hall and having her familiarity set to the counterfactual under no exposure to the town hall. Thus, the ACME isolates the average change in esteem due specifically to changes in familiarity that were induced by exogenous exposure to the member, among those who were exposed, and is identified as a causal effect under assumptions that we state below.[6]

To estimate the ACME, we construct all its components (Imai et al., 2010). First, we simulate the distribution of familiarity conditional on being exposed to treatment ($\eta_4(1)$), and conditional on not being so exposed ($\eta_4(0)$). Second, we simulate the distribution of esteem (η_5) based on the estimated model parameters, and under two sets of assumptions: assuming treatment exposure $D = 1$ and using the distribution of familiarity ($\eta_4(1)$) assuming $D = 1$, which yields a posterior distribution for $\eta_5(1,\eta_4(1))$, and assuming treatment exposure $D = 1$ but now using the distribution of familiarity from the unexposed state ($\eta_4(0)$), which yields a posterior distribution for $\eta_5(1,\eta_4(0))$.[7] The ACME estimate is the conditional expected difference in the two η_5 posterior distributions, averaged over respondents. We retrieve point estimates of the ACME as well as Bayesian credible intervals based on the posterior distributions.

We estimate this model with the full sample of all 760 respondents who completed the background materials survey (including a fixed effect whether the respondent and the member's political party differ). We also estimate the model by subsetting these according to whether the member and constituent were same-partisan ($N = 439$) or opposite-partisan ($N = 303$). These subsetted models fully interact all elements of the model with the same- or opposite-partisan indicator. Each model also includes a

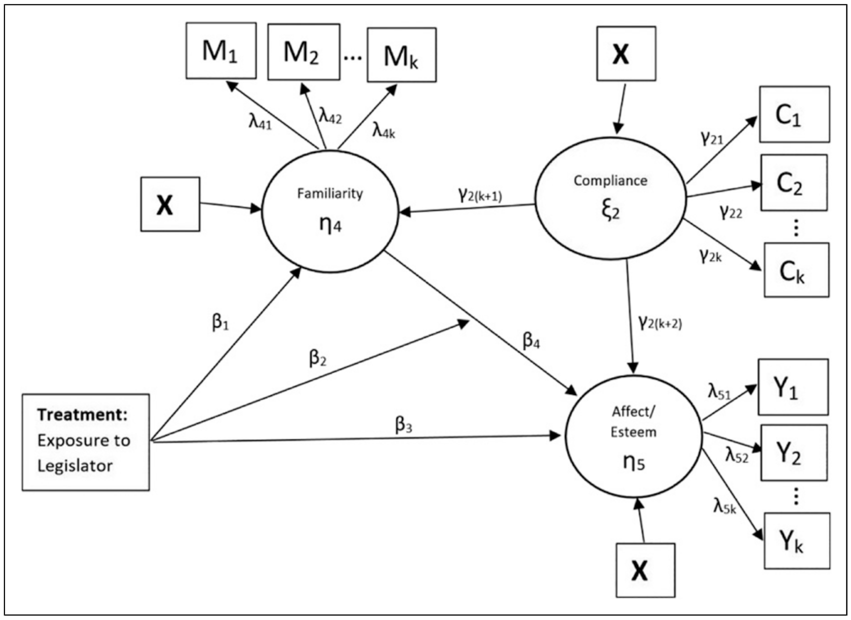

Figure 1. *Causal Mediation Model.* The figure presents the causal mediation model we deploy below. Notationally, **C** denotes indicators of compliance with experimental activities; **M**, of whether a participant responded to attitude items about the member as our measure of familiarity; **Y**, of esteem for the member; and **X** refers to covariates. In a slight abuse of graph representation, the arrow labeled β_2 indicates the multiplicative interaction between **Treatment** and **Familiarity**. See Supplemental Appendix A.5 for details and code. We estimated this model for the 760 respondents who completed the background materials survey, and for the same-partisan and opposite-partisan subgroups.

battery of covariates in the outcome, familiarity, and compliance equations to adjust for demographic and personality factors that are potential confounders; including these covariates also improves the precision of estimates.[8] These covariates were measured on the pretest survey or provided by the survey vendor. When a respondent is not eligible for a given task, the item that indicates complying with the task is set to missing.[9]

As with any field experiment that uses an encouragement design, we encountered noncompliance with treatment assignment and nonresponse on the post-test survey. That is, some participants either failed to attend a session when invited, failed to respond to the post-test survey, or both.[10] Rates of assignment, compliance and response are detailed in the appendix. To identify causal effects in the presence of noncompliance and nonresponse, we use the GET model for causal inference (Esterling et al., 2011a). The GET model serves two methodological purposes. First, GET is a generalization of the methods of instrumental variables and principal stratification (Angrist et al., 1996; Frangakis and Rubin, 1999, 2002) which are widely used methods for field experimental data. In field experiments it is common for some participants to fail to comply with their randomly assigned treatment, and for some participants to fail to respond to the outcome survey. Generalized endogenous treatment generalizes instrumental variables and the method of principal stratification by explicitly measuring and conditioning on participants' compliance type, and hence can recover the counterfactual comparison both among compliers as well as among non-compliers within the randomized design.

Second, GET provides a modeling solution to justify the assumption of sequential ignorability to identify mediation causal effects described in Imai et al. (2011), which are

$$\{\eta_5(d',\eta_4),\eta_4(\mathrm{d})\} \perp D \mid \xi_2, X \tag{2a}$$

$$\{\eta_5(d',\eta_4) \perp \eta_4(\mathrm{d})\} \mid D, \xi_2, X \tag{2b}$$

for $d,d' = 0,1$ and for all realizations of η_4. We justify assumption 2a using the design of the GET model by conditioning on ξ_2, which, as we explain above, holds constant the "compliance type" principal stratification of the respondent (Frangakis and Rubin 2002). We justify assumption 2b since ξ_2 creates the conditional independence between the mediator and the outcome required for sequential ignorability. As Esterling et al. (2011b) demonstrate, the GET model parametrically models all stochastic dependence between the mediator and the outcomes due to omitted variables. In this parametric approach, we do not need to condition on covariates in order to justify an assumption of sequential ignorability; instead, the model captures any endogeneity between our measure of familiarity and our measure of esteem by including the random effect ξ_2 in both equations.

Esterling et al. (2011b) is an example of an application of the GET model to mediation.

Latent Variables

The three elements (η_4,η_5,ξ_2) indicated as ovals in Figure 1, are latent variables, and are estimated in a measurement model based on sets of indicators.[11] The model is a structural equation model (Bollen, 1989) in that all latent variables, missing data, and structural parameters are estimated simultaneously.

We measure esteem (η_5) based on a set of indicators **Y**. The first item is the member feeling thermometer score, which serves as a summary of the constituent's warmth toward the member (Wilcox et al., 1989; Winter and Berinsky, 1999). The respondent was presented with a feeling thermometer slider, with scores scaled to range from 0 to 100, and the text "[MOC], your Member of Congress," where the member's name appeared in place of [MOC]. Higher scores indicate "warmer" attitudes toward the member.[12] The remaining items measure trust in the member ("How much of the time do you think you can trust [MOC], your Member of Congress, to do what is right?" Always, Most of the time, Some of the time, Not at all); approval of the member ("Do you approve of the way that [MOC] is handling [MOCPRONOUN2] job as Congressperson?" five-point response scale); approval of the member on immigration policy ("Do you approve or disapprove of the way [MOC], your Member of Congress, is handling the issue of immigration?" five-point response scale); and vote intent ("If the vote for the House of Representatives were held today, who would you vote for?" Definitely [MOC] [MOCPARTY], Probably [MOC] [MOCPARTY], Probably [CHALLENGER] [CHALPARTY], Definitely [CHALLENGER] [CHALPARTY], with Undecided, Other candidate, and Would not vote set to missing). These last four items included DK options. Therefore, we only observed a response if the respondent did not select DK; DK responses and skips are set to missing in Y.

As justified above, we measure familiarity with the member (η_4) as the latent propensity to respond *DK* to post-test items about the member herself or her actions in D.C., collected in the vector **M**, with each having value 0 if the respondent responded *DK*, and 1 if the respondent reported an attitude *DK*. Nonresponses due to skips are set to missing. Specifically, we construct a variable indicating *DK* for items measuring trust of the member, approval of the member on immigration policy,[13] and a knowledge question regarding how the member voted on important immigration legislation ("How about [MOC], your Member of Congress? Do you think [MOCPRONOUN] voted for or against making it a felony to assist illegal immigrants in entering or remaining in the U.S.?").

Finally, to implement the GET causal model (Esterling et al. 2011a), we must condition on the propensity to comply with the experimental protocol (ξ_2). We constructed this scale from **C**, a vector of three indicator variables that track whether the respondent complied with an assigned experimental activity. That is, we had an indicator for whether she: participated in the town hall (if assigned to the DG condition); responded to the post-test survey; and responded to a supplemental survey administered in November of 2006. In instances where a respondent was ineligible for an experimental activity (e.g., participating in the town hall, if not assigned to the DG condition), the relevant indicator is set to missing and the missing value is imputed dynamically as a probability distribution conditioned on the ξ_2 latent variable (Tanner and Wong, 1987).

Results

With our measures for familiarity, esteem and compliance in hand, we now turn to our statistical findings. Consistent with the expectations based on overgeneralization in judgment, familiarity has very large and positive effects on esteem. These findings hold for both the latent esteem variable as well as for each item we use to measure esteem. The estimates are particularly large for same-partisans, but also positive (though insignificant) for opposite partisans.

The estimated ACMEs appear in Figure 2. The ACME is the effect of exposure to the member through a town hall on esteem that is due to changes in the familiarity mediator. The first row of Figure 2 shows the posterior distributions of the ACME for the full sample, and the first column presents the ACME for latent esteem, η_5. Bars indicate 95%, 90%, and 50% distributions, in the direction increasing thickness. In the full sample, the posterior distribution of the ACME has a mean of about 1, which is an increase of roughly a standard deviation. Filling out the other rows, we see a somewhat larger effect for same-partisans and an attenuated effect for opposite-partisans. While the ACME estimate is not statistically significant for opposite-partisans, the posterior distribution mostly sits on the positive side, indicating that we have no evidence that familiarity breeds contempt among this group.

The remaining columns show the ACME results for the individual items. Technically, we retrieve the ACME results for the k^{th} item by scaling the ACME for η_5 by the corresponding factor coefficient for the item, λ_{5k}. The factor coefficients are reported in Supplemental Appendix A 5.1.

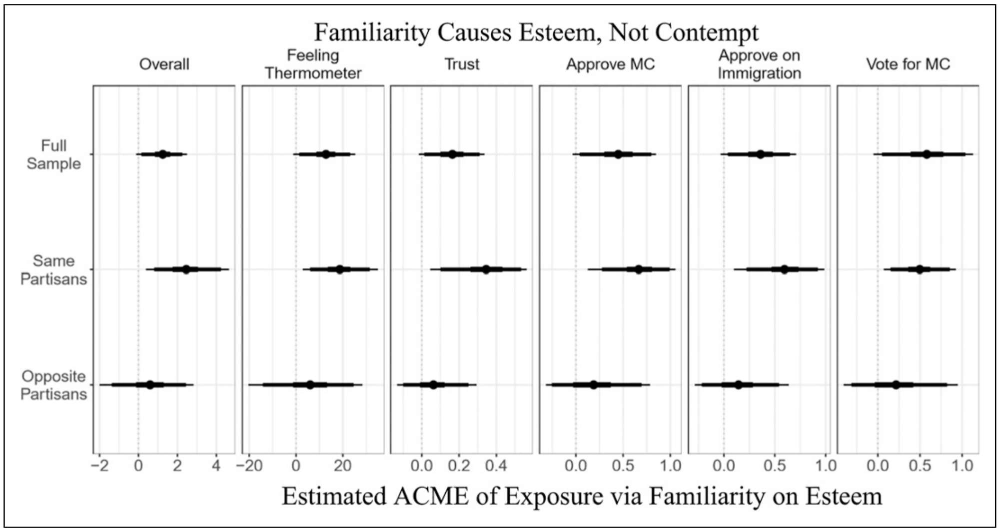

Figure 2. ACME estimates Each cell presents a summary of the posterior distribution for the ACME for a sample and outcome. These estimates condition on exposure to treatment (D = 1), varying only in the (mediating) value of familiarity (η_4). Bars indicate 95%, 90%, and 50% distributions, in the direction increasing thickness. Overall, the results indicate evidence consistent with overgeneralization in judgment—an increasing mediated effect of exposure, via familiarity, on esteem and its constituent outcomes. Full sample N = 760; same-partisan sample N = 439; opposite-partisan sample N = 303.

Feeling thermometer scores were scaled to have a mean of 62 and a standard deviation of 21 (minimum 0, maximum 100), with higher scores indicating a "warmer" attitude toward the member. Since the latent variable has a unit scale, the parameter estimate suggests that a one standard deviation increase in the latent propensity to offer a response on attitude items is associated with increased feeling thermometer ratings by about 12 points for the full sample, or about a half of a standard deviation, and nearly 20 points among same-partisans—nearly a standard deviation.

Across all items, the ACME is substantially different from zero among the same-partisans. For this group, the ACME for trust is just over a standard deviation, for approval is about one standard deviation, for approval on immigration policy nearly a standard deviation and for planning to vote for the member is just over a standard deviation. While the ACME estimated posterior distributions from the opposite-party sample overlap zero, the point estimates and the bulk of the distribution is positive. By that same reasoning, however, they overlap substantially with estimates for the same-party sample. At the very least, we certainly have no evidence to suggest that familiarity diminishes esteem among opposite-partisans.

Overall, the figure shows that among both same- and opposite-partisans, increased familiarity toward a member is correlated with more positive or "warmer" attitudes particularly among same-partisans. Familiarity does *not* breed contempt, even among constituents who oppose their member's party (a finding similar to Broockman and Butler, 2017). This suggests that on balance, enhanced familiarity with a member enhances constituents' feelings toward the member, a finding consistent with Alvarez (1997), Fenno (1978) and Lammers et al. (2021), and inconsistent with Hibbing (2002).

As we mention in the introduction, there is something of a paradox that the model must resolve to identify mediation effects in this context: respondents with lower familiarity tend to select *DK*, and hence, their attitudes toward the member are often missing. The model resolves this paradox by estimating each respondent's *propensity* to respond *DK* or *DK*, rather than treating familiarity as either-or. In so doing, we identify the structural parameters by leveraging the *DK* responses among those who have low familiarity but responded to the member esteem items anyway, and the *DK* responses among those who have high familiarity and did not respond on the esteem items, along with the linearity assumptions of the model. This approach relying on estimated propensities and functional form assumptions is common to any parametric model that relies on propensity scores.

This strategy for identification is aided by the fact that the feeling thermometer scale did not contain a *DK* filter and so nearly all respondents provided an answer, including 74% of those at the lowest level of familiarity who also completed the post-test survey. Because of this design, we have at least some esteem-related attitude data on nearly all respondents including those with low familiarity. The remaining imputations are done via the covariates and the linear assumptions of the mediation model, and again these are normal assumptions in any model that uses a propensity score—that is, using low propensity responders who do respond, and high propensity responders who do not respond, to identify the counterfactual.

Finally, we note that the total treatment effect for each item (complier average treatment effects as reported in Minozzi et al. (2015)) is relatively small compared to the ACME—on the order of 0.2–0.5 on Cohen's *D* scale. The average direct effect (ADE) is the remaining set of mediating effects other than familiarity. Given that the total effect is positive but small while the effect of familiarity is large and positive, the effect of the remaining mediators captured by the ADE is both negative and smaller than the effect of familiarity in absolute value. Since we do not know the content of these remaining mediators we do not dwell on the ADE estimates.

Discussion and Conclusion

Heretofore we have not had a good way of settling an important, long-standing theoretical and empirical disagreement in political science regarding the relationship between familiarity and esteem in democratic representation. We resolve this impasse by offering a definition of familiarity as the condition of having engaged sufficiently to realize some latent attitudes, and positing that, in the absence of familiarity, constituents overgeneralize. Thus, even when they are unfamiliar with their representative, we should expect constituents to maintain some latent affect. We draw on uncertainty aversion and the social psychological framework of mere exposure to explain how this overgeneralized, negative latent affect will brighten upon exposure to the member's outreach efforts. Finally, we develop a novel, robust, and compelling method to measure the relationship between familiarity and latent attitudes as substantive opinions.

Our findings demonstrate a basic consistency between the normative theory of democratic representation and the practice of representation in mass democracy (Madison, 1961; Mansbridge, 2003; Pitkin, 1967). That members enhance esteem among their constituents by increasing constituents' familiarity with them creates incentives for members to connect with their constituents. In this experiment, a sample of constituents were invited to a direct interaction with their MC, and by drawing contrasts with a control group who were not invited, we identify the effect of increased familiarity on constituents' attitudes toward their member, as well as their propensity to offer an attitude response at all. This enhanced familiarity dramatically increased constituents' willingness and ability to report an attitude, and at the same time typically led them to hold more positive attitudes toward their MC—familiarity emphatically did *not* breed contempt. As expected, the effect of familiarity was strongest among same-partisans.

We acknowledge, of course, the scope conditions of our findings: the familiarity that we induce in the experiment is the kind which elected officials themselves cultivate in what Fenno (1978) called "presentation of self." Since such self-presentation is envisioned within normative theories of democratic representation (Pitkin, 1967), we consider it an important cause of familiarity. Our results do not speak, however, to effects of exposure to information through the mass media, such as campaign messages, negative advertising, or newspaper reports (Lau et al., 2007). Our results instead demonstrate the effect of familiarity on attitudes that arises directly within the practice of democratic representation—that is, from member outreach efforts—which is the contrast to the normal exposure to mediated information about the member available to those in the control condition.

We offer a further caveat that the study occurred in summer of 2006, and so we must be cautious in extrapolating the findings to today. There are two possible reasons why the findings may differ. First, while partisan polarization is not new (Hetherington, 2001; Poole and Rosenthal, 1984), and its current scope is possibly overstated (Levendusky and Malhotra, 2016), it is possible that the nature of partisanship has changed since the study was completed (Darr and Dunaway, 2018), and that may matter in particular for the response of opposite-partisans to a Deliberative Town Hall. However, there is no evidence to suggest that the magnitude of treatment effects for deliberation interventions, such as the one we used in this study, have changed over time. For example, Abernathy (2018) reports results with a telephone town hall experiment with members of Congress in 2016 that had similar responses from constituents. Second, the nature and extent of technology-enabled communication has changed over time (Settle, 2018). However, the platform used in this study does not have an embedded social algorithm nor did it give users the opportunity to distribute false news stories (Lazer, 2015), and so a similar platform used today would likely have similar effects.

We assert that these results demonstrate the role of familiarity as fundamental to democratic representation. Given the increasing importance of digitally mediated communication, the present results demonstrate that familiarity within democratic representation can be generated via online interactions, which are likely to become more pervasive in democratic practice going forward.

Declaration of Conflicting Interests

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

Funding

This project was funded by a grant from the Digital Government Program of the National Science Foundation (IIS-0429452). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The replication package for this paper can be found on the Open Science Framework at https://osf.io/3y48x/.

ORCID iD

Kevin M. Esterling https://orcid.org/0000-0002-5529-6422

Supplemental Material

Supplemental material for this article is available online.

Notes

1. https://www.pewresearch.org/politics/2014/06/12/political-polarization-in-theamerican-public/, accessed April 7, 2021.
2. Knowledge Networks (KN, now GfK) maintains a panel of respondents that is representative of the U.S. population. To meet sample size requirements in each congressional district, KN subcontracted with two other vendors, Survey Sampling International (SSI, now Dynata) and Global Mapping International (GMI). In the models below we include fixed effects to account for any differences between these panels. Including SSI and GMI respondents enables us to generalize only to Internet-connected constituents in the study's congressional districts. See the appendix for more details.
3. Because of this conditioning, the statistical results we report only generalize to the subpopulation of respondents who are engaged enough to read the reading materials. This subpopulation is of direct interest in that it likely represents constituents that have at least a minimal level of political engagement. In addition, in the appendix we show that 67% of respondents choose to do the reading when given the opportunity, and in Supplemental Appendix Table 3 we show the covariate distribution of this subsample is not particularly different from those who chose not to read the background material.
4. We explain the measure of familiarity as a latent variable above, which is measured by DK responses to affect items. Our theory posits affect as a latent construct, and so the outcomes also must be measured as a latent variable in the structural equation model. The results replicate, however, when we model the feeling thermometer (the only affect item that lacks systematic missingness) as a single outcome.
5. Our notation for latent variables follows that of Bollen (1989), where ξ indicates an exogenous latent variable and η indicates an endogenous latent variable.
6. The case where $d = 0$ is the corresponding ACME estimand for those who were not exposed to the town hall. If the estimates under the two cases differ, that would indicate an interaction between the mediator and the treatment exposure.
7. We also estimated the corresponding ACME conditional on maintaining $D = 0$ in the second step, and the results are identical to those we report. This indicates an absence of a treatment-mediator interaction.
8. Very few of the covariate parameters are significant in any of the equations (see Supplemental Appendix Table 8). None of the demographic covariates reach statistical significance to explain compliance with attending the online Deliberative Town Hall (see Neblo et al. 2010), which indicates that the town hall participants are broadly representative of their congressional districts. Notably however, consistent with Mondak and Anderson (2004), men score higher on the familiarity measure.
9. All missing data is imputed as missing at random, and the parameter estimates are marginal to the missing data distributions (Tanner and Wong 1987).
10. Here we refer to unit nonresponse, that is, nonresponse on the survey itself, as opposed to item nonresponse, skipping or responding DK on a specific attitude item.
11. We provide the results of the measurement models in Supplemental Appendix A.2.
12. This feeling thermometer was presented among a battery of thermometer measures, and the order of the thermometers rotated randomly. There was no DK option given for this item; nonresponses (skips) on the outcome variable are set to missing.
13. As we note above, the data have two indicators of approval of the member: the general question about job approval, and the specific question about approval of the member's handling of immigration policy. Unfortunately, because trust and the general approval measure are so highly correlated, the mediation model fails to converge for empirical reasons in the opposite-partisan subsample when we include both to measure the familiarity latent. To maintain consistency across all models, and to ensure that we only report results that are robust and stable, we report the results from the models that do not include the general approval indicator. This restriction has no substantive impact. The mediation models that include the general approval item that do converge (the full sample and the same-partisan sample) yield results that are identical to the restricted model to two significant figures.

References

Abernathy, Claire. 2018. "Understanding and Responding to Constituent Opinion on Capitol Hill." *Congress and the Presidency* 45 (3): 254–278.

Alvarez, R. Michael. 1997. *Information and Elections*. Ann Arbor, MI: University of Michigan Press.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects using Instrumental

Variables." *Journal of the American Statistical Association* 91 (June): 444–455.

Arnold, R. Douglas. 1990. *The Logic of Congressional Action*. New Haven, CT: Yale University Press.

Atir, Stav, Emily Rosenzweig, and David. Dunning. 2015. "When Knowledge Knows No Bounds: Self-Perceived Expertise Predicts Claims of Impossible Knowledge." *Psychological Science*.

Berinsky, Adam J 1999. "The Two Faces of Public Opinion." *American Journal of Political Science* 43 (Oct): 1209–1230.

Berinsky, Adam J., and Joshua A. Tucker. 2006. "'Don't Knows' and Public Opinion Towards Economic Reform: Evidence from Russia." *Communist and Post-Communist Studies* 39: 73–99.

Bollen, Kenneth A 1989. *Structural Equations with Latent Variables*. New York, NY: John Wiley and Sons, Ltd.

Box-Steffensmeier, Janet M., David C. Kimball, Scott R. Meinke, and Katherine. Tate. 2003. "The Effects of Political Representation on the Electoral Advantages of House Incumbents." *Political Research Quarterly* 56 (3): 259–270.

Brady, David W., and Sean M. Theriault. 2001. "A Reassessment of Who's to Blame: A Positive Case for the Public Evaluation of Congress." In *What is it About Government that Americans Dislike?*, edited by John R. Hibbing and Elizabeth Theiss-Morse, 175–192. New York, NY: Cambridge University Press.

Broockman, David E., and Daniel M. Butler. 2017. "The Causal Effects of Elite Position Taking on Voter Attitudes: Field Experiments with Elite Communication." *American Journal of Political Science* 61 (1): 208–221.

Cover, Albert D., and Bruce S. Brumberg. 1982. "Baby Books and Ballots: The Impact of Congressional Mail on Constituent Opinion." *American Political Science Review* 76 (2): 347–359.

Darr, Joshua P., and Johanna L. Dunaway. 2018. "Resurgent Mass Partisanship Revisited: The Role of Media Choice in Clarifying Elite Ideology." *American Politics Research* 46 (6): 943–970.

Delli Carpini, Michael X, and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.

Esterling, Kevin M., Michael A. Neblo, and David M.J. Lazer. 2011a. "Estimating Treatment Effects in the Presence of Noncompliance and Nonresponse: The Generalized Endogenous Treatment Model." *Political Analysis* 19 (Spring): 205–226.

Esterling, Kevin M., Michael A. Neblo, and David M.J. Lazer. 2011b. "Means, Motive, and Opportunity in Becoming Informed about Politics: A Deliberative Field Experiment Involving Members of Congress and their Constituents." *Public Opinion Quarterly* 75 (Fall): 483–503.

Fenno, Richard F. 1978. Homestyle: House Members in Their Districts. Boston, Mass. Little, Brown and Co.

Fiorina, Morris, Bruce Cain, and John Ferejohn. *The Personal Vote: Constituency Service and Electoral Independence*. Cambridge, MA: Harvard University Press, 1987.

Firestone, Chaz, and Brian J. Scholl. 2015. "When do Ratings Implicate Perception Versus Judgment? The "Overgeneralization Test" for Top-Down Effects." *Visual Cognition* 23 (910): 1217–1226.

Frangakis, Constantine E., and Donald B. Rubin. 1999. "Addressing Complications of Intention-to-Treat Analysis in the Combined Presence of All-or-None Treatment Noncompliance and Subsequent Missing Outcomes." *Biometrika* 86 (2): 365–379.

Frangakis, Constantine E., and Donald B. Rubin. 2002. "Principal Stratification in Causal Inference." *Biometrics* 58 (March): 21–29.

Grant, J. Tobin, and Thomas J. Rudolph. 2004. "The Job of Representation in Congress: Public Expectations and Representative Approval." *Legislative Studies Quarterly* 29 (Aug): 431–445.

Grimmer, Justin, Solomon Messing, and Sean J Westwood. 2012. "How Words and Money Cultivate a Personal Vote: The Effect of Legislator Credit Claiming on Constituent Credit Allocation." *American Political Science Review* 106 (4): 703–719.

Harmon-Jones, Eddie, and John J. B. Allen. 2001. "The Role of Affect in the Mere Exposure Effect: Evidence from Psychophysiological and Individual Differences Approaches." *Personality and Social Psychology Bulletin* 27 (7): 889–898.

Hetherington, Marc J. 2001. "Resurgent Mass Partisanship: The Role of Elite Polarization." *American Political Science Review* 95 (3): 619–631.

Hibbing, John R. 2002. "How to Make Congress Popular." *Legislative Studies Quarterly* 27 (2): 219.

Hibbing, John R., and Elizabeth Theiss-Morse. 1995. *Congress as Public Enemy*. New York, NY: Cambridge University Press.

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis." *Psychological Methods* 15 (4): 309–334.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105 (Nov): 765–789.

Jessee, Stephen A 2017. "'Don't Know' Responses, Personality, and the Measurement of Political Knowledge." *Political Science Research and Methods* 5 (4): 711–731.

King, Gary. 1991. "Constituency Service and Incumbency Advantage." *British Journal of Political Science* 21 (1): 119–128.

Krosnick, Jon A 2002. "Survey Nonresponse." In *Technometrics*, edited by R M Groves, D A Dliiman, J L Eltinge, and R J A Little, Vol. 44. New York: Wiley.

Krosnick, Jon A., and Michael A. Milburn. 1990. “Psychological Determinants of Political Opinionation.” *Social Cognition* 8 (1): 49–72.

Lammers, Joris, Eileen Pauels, Alexandra Fleischmann, and Adam D. Galinsky. 2021. “Why People Hate Congress but Love Their Own Congressperson: An Information Processing Explanation.” *Personality and Social Psychology Bulletin* forthcoming, 1–14.

Lau, Richard R., Lee Sigelman, and Ivy Brown. Rovner. 2007. “The Effects of Negative Political Campaigns: A Meta-Analytic Reassessment.” *The Journal of Politics* 69 (4): 1176–1209.

Lazer, David M.J 2015. “The Rise of the Social Algorithm.” *Science* 348 (6239): 1090–1091.

Lee, Angela Y 2001. “The Mere Exposure Effect: An Uncertainty Reduction Explanation Revisited.” *Personality and Social Psychology Bulletin* 27 (10): 1255–1266.

Levendusky, Matthew, and Neil Malhotra. 2016. “(Mis)perceptions of Partisan Polarization in the American Public.” *Public Opinion Quarterly* 80: 378–391.

Luskin, Robert C., and John G. Bullock. 2011. ““Don’t Know” Means “Don’t Know”: DK Responses and the Public’s Level of Political Knowledge.” *Journal of Politics* 73 (April): 547–557.

Madison, James. 1961. “Federalist 57: The Alleged Tendency of the New Plan to Elevate the Few at the Expense of the Many Considered in Connection with Representation.” In *The Federalist Papers*, edited by Clinton Rossiter, New York, NY: New American Library.

Mansbridge, Jane. 2003. “Rethinking Representation.” *American Political Science Review* 97 (4): 515–528.

Minozzi, William, Michael A. Neblo, Kevin M. Esterling, and David M.J. Lazer. 2015. “Field Experiment Evidence of Substantive, Attributional, and Behavioral Persuasion by Members of Congress in Online Town Halls.” *Proceedings of the National Academy of Sciences* 112 (13): 3937–3942.

Mondak, Jeffery A 1999. “Reconsidering the Measurement of Political Knowledge.” *Political Analysis* 8 (1): 57–82.

Mondak, Jeffery J., and Mary R. Anderson. 2004. “The Knowledge Gap: A Reexamination of Gender-Based Differences in Political Knowledge.” *Journal of Politics* 66 (2): 492–512.

Mondak, Jeffrey J., Edward G Carmines, Robert Huckfeldt, Mitchell Dona-Gene, and Scot Schraufnagel. 2007. “Does Familiarity Breed Contempt? The Impact of Information on Mass Attitudes Toward Congress.” *American Journal of Political Science* 51 (Jan): 34–48.

Neblo, Michael A., Kevin M Esterling, Ryan Kennedy, David Lazer, and Anand Sokhey. 2010. “Who Wants to Deliberate – and Why?” *American Political Science Review* 104 (3): 566–583.

Neblo, Michael A., Kevin M. Esterling, and David M.J. Lazer. 2018. *Politics with the People: Building a Directly Representative Democracy*. New York, NY: Cambridge University Press.

Norpoth, Helmut, and Milton Lodge. 1985. “The Difference between Attitudes and Nonattitudes in the Mass Public: Just Measurements.” *American Journal of Political Science* 29 (2): 291.

Pierce, John C., and Douglas D. Rose. 1974. “Nonattitudes and American Public Opinion: The Examination of a Thesis.” *American Political Science Review* 68 (2): 626–649.

Pitkin, Hanna Fenichel. 1967. *The Concept of Representation*. Berkeley, CA: University of California Press.

Poole, Keith T, and Howard Rosenthal. 1984. “The Polarization of American Politics.” *Journal of Politics* 46 (4): 1061–1079.

Settle, Jamie M 2018. *Frenemies: How Social Media Polarizes America*. New York, NY: Cambridge University Press.

Steenbergen, Marco R, Andre Bachtiger, Markus Sporndli, and Jurg Steiner. 2003. “Measuring Political Deliberation: A Discourse Quality Index.” *Comparative European Politics* 1 (1): 21–48.

Tanner, Martin A., and Wing Hung Wong. 1987. “The Calculation of Posterior Distributions by Data Augmentation.” *Journal of the American Statistical Association* 82 (398): 528–540.

Turgeon, Mathieu. 2009. “’Just Thinking’: Attitude Development, Public Opinion, and Political Representation.” *Political Behavior* 31: 353–378.

Visser, Penny S., Allyson L. Holbrook, and Jon A. Krosnick. 2007. “The Sage handbook of public opinion research.” In *The Sage Handbook of Public Opinion Research*. Thousand Oaks, CA: Sage Publications.

Wilcox, Clyde, Lee Sigelman, and Elizabeth Cook. 1989. “Some Like it Hot: Individual Differences in Responses to Group Feeling Thermometers.” *Public Opinion Quarterly* 53: 246–257.

Winkielman, Piotr, and John T Cacioppo. 2001. “Mind at Ease Puts a Smile on the Face: Psychophysiological Evidence that Processing Facilitation Elicits Positive Affect.” *Journal of Personality and Social Psychology* 81 (6): 989–1000.

Winter, Nicholas, and Adam Berinsky. 1999. *What’s Your Temperature? Thermometer Ratings and Political Analysis*. Tech. rept. Atlanta, GA: Presented at the Annual Meeting of the American Political Science Association.

Zajonc, R.B. 2001. “Mere Exposure: A Gateway to the Subliminal.” *Current Directions in Psychological Science* 10 (6): 224–228.

Supplemental Materials for: “Does Familiarity Breed Esteem? A Field Experiment on Emergent Attitudes toward Members of Congress”

Kevin M. Esterling
(Corresponding Author)
Professor, Public Policy and Political Science
UC – Riverside
kevin.esterling@ucr.edu

William Minozzi
Associate Professor, Political Science
Ohio State University
minozzi.1@osu.edu

Michael A. Neblo
Professor, Political Science
Ohio State University
neblo.1@osu.edu

December 17, 2021

A Appendix

Contents

This appendix describes the experimental design and statistical analysis for the paper "Does Familiarity Breed Contempt (or Esteem)? A Field Experiment on Emergent Constituent Attitudes toward Members of Congress."

A.1 Research Design

The research design and data collection for this field experiment are summarized in the paper. Here we describe the design and the data collection in more detail. In this section, we also discuss some deviations from the ideal experimental design – the kinds of complications that can often occur in a large field experiment – and how our methods address these issues.

A.1.1 Subject Recruitment and Selection

The data we report in this paper come from a larger study centered on a field experiment from the summer of 2006 in which constituents from a set of U.S. Congressional districts were randomly selected to participate in a deliberative online town hall with their current member of Congress. The study proceeded in four stages.

At the first stage we recruited a representative cross-section of each congressional district to participate in a pre-treatment (or baseline) survey; this survey provides the data for the construct validity test. At the conclusion of this pretest survey, we provided the dates and times of town hall sessions that were scheduled and invited respondents to RSVP for an online town hall. The respondents who RSVP'd "yes" were randomly assigned to one of three groups: 1) a "deliberative group" (DG) that is scheduled to receive background reading materials and an invitation to participate in the town hall; 2) an "information only group" (IO) this is scheduled to receive the same background reading materials but is not invited to the town hall; and 3) a "true control group" (TC) that does not receive the reading materials or an invite. The respondents who indicate an interest in participating in the study but RSVP "no" are randomly assigned to either IO or TC. The remaining respondents are excluded from the study.

At the second stage, after about 10 days, respondents assigned to either the IO or the DG group are provided the background reading materials and asked to complete a short survey. At the third stage, respondents in the DG group attend their scheduled town hall, about three days after receiving the reading materials. At the fourth stage, all respondents (TC, IO and DG) in the congressional district are administered a posttreatment (or post-test) survey about one week following the conclusion of the town hall. This post-test suvey has many of the same questions as appeared on the pretreatment survey. We use the data from the post-treatment survey to estimate the causal mediation model.[1] As we state in the paper, to ensure comparability we focus our

[1] We also administered a survey after the November 2006 midterm elections. We do not use the responses from the November survey in this study other than an indicator of their response to measure the compliance type latent variable.

analyses on the subsample of respondents who were assigned either to the IO or to the DG conditions and who completed the background readings survey. A comparison between the DG and the TC respondents would introduce an unwanted confound, in that the DG respondents also read policy background material that was unrelated to the member while the TC respondents did not. The counterfactual of interest here is only with respect to exposure to the member.

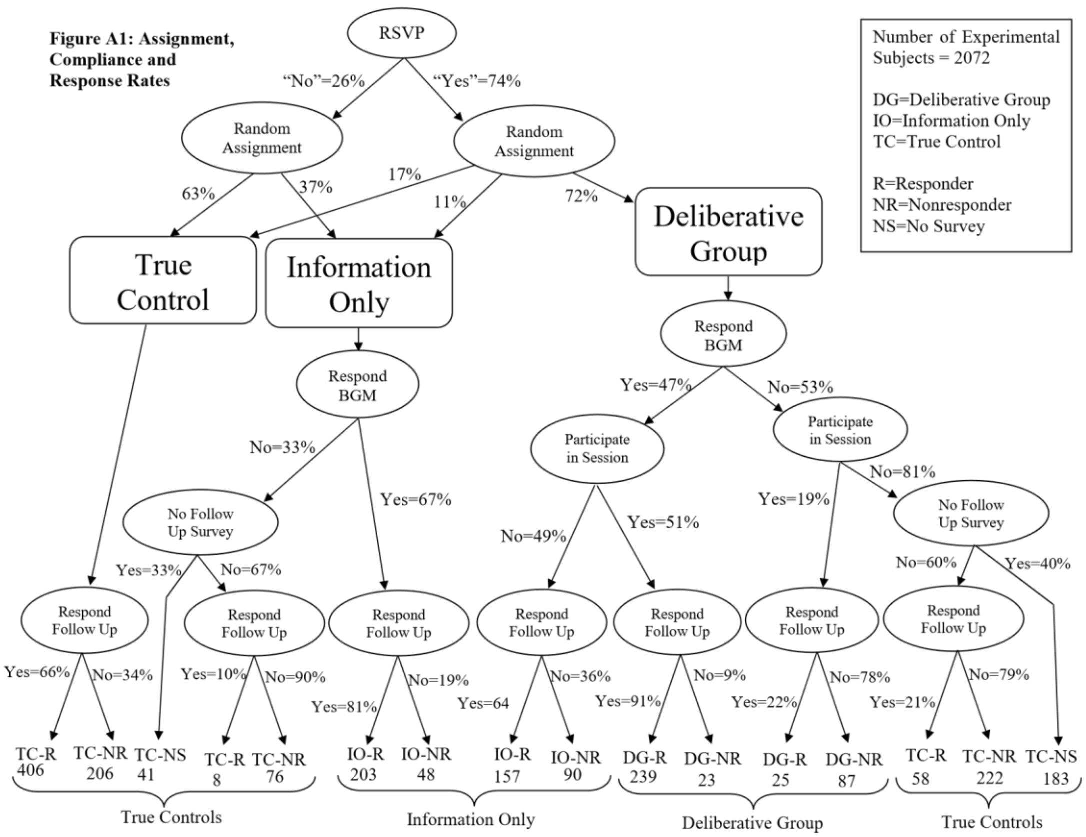

Figure 3: Assignment, Compliance and Response Rates

Figure 3 gives an overview of the study flow through these four stages, and the assignment, compliance and response rates among the 2072 subjects in the experiment. The flow chart has four stages: an RSVP and assignment, and then (if eligible) exposure to the background materials (BGM), exposure to a deliberative session, and a post-test survey. Knowledge Networks (KN, now GfK) administered all aspects of the study.

A.1.2 RSVP and Assignment

In the pretest survey we included an RSVP filter question, which indicated the time the session would take place for the subject's congressional district, and that the session would last approximately an hour. We then allowed subjects to indicate whether they 1) would be willing and able to attend the session; 2) would only complete surveys for the project; or 3) refused to participate in the project. Only 10.7 percent of respondents refused to participate in the study; we discard these observations and do not consider them further.[2] Among those who agreed to participate in some way, 73 percent of subjects indicated they would be willing and able to attend a session, and these subjects were randomized among the three treatment arms (72 percent to deliberative group (DG) condition, 11 percent to the information only (IO) condition, and 17 percent to the true control (TC) condition). The remaining 26 percent of subjects who indicated they wanted to participate in the study but would not attend the session were randomized among the information only condition (37 percent) and the true control condition (63 percent).

We chose these assignment rates in an attempt to create as-treated groups of sufficient size, for each treatment. For example, we assigned relatively few subjects to the IO condition, since we assumed many subjects assigned to the DG condition would read and complete the background materials survey but fail to attend their assigned session. These subjects would then receive the IO treatment. Of course, we did not know the compliance rates in advance, so the as-treated cell sizes are not identical to each other.

We included the RSVP filter question to improve the information we had available at the initial assignment stage. When we began the study, KN did not know the rate at which subjects would attend the session in practice. As a result, at the beginning of the study, we simply did not know what proportion of subjects to assign to the deliberative condition in order to ensure enough subjects in the cell of subjects who complete the deliberative session and who respond to the post-test survey. Asking the RSVP filter question gave us some of this information. And of course, once we asked the RSVP question, it would have been odd to invite subjects after they have indicated they would not attend. Hence, we randomize these participants among the other two arms.

We retain the participants who RSVP'd "No" (self-reported they wished to participate in the study but would not attend a session) for use in the statistical analysis. Asking the RSVP question simply gave us more information regarding the likely take up rate of the deliberative sessions, but otherwise does not affect our analyses, under one assumption: that the RSVP self-reports are true, that the respondent indeed would not have attended a session had she been given the opportunity.[3] To see this, imagine the case if we had not asked this filter question. Assuming those who RSVP'd "No" would have failed to attend a session if invited, these respondents would have selected themselves into one of the control groups through non-compliance. As we emphasize in the paper, noncompliance

[2] An additional 299 subjects did not respond to the pretest survey, for an AAPOR RR6 response rate of 0.76 (see Callegaro and Disogra, 2008).

[3] In our statistical model, the converse of this assumption, that those who say they will attend in fact attend, does not need to be true.

is inevitable in a field experiment and we use statistical methods to identify causal effects in the presence of noncompliance. Thus, under the assumption that the RSVP "No's" are accurate, asking the RSVP gave us information for basing assignment rates, but otherwise is irrelevant to the study.[4]

A.1.3 Background Reading Materials

The topic of the deliberative online town halls was U.S. immigration policy. We provided non-partisan, factual reading materials to the respondents to the IO and DG conditions on this topic that we distilled from reports written by the Congressional Research Service. About half of respondents in each treatment arm filled out the background survey. We restrict the analysis in the causal mediation model to those in the IO or DG conditions who complete the background survey because these self-selected respondents are most comparable to each other.[5]

A.1.4 Treatment Compliance and Response Rates

The nodes below the assignments in figure 3 indicate compliance with each task (excluding responses to an additional survey we administered the following November, which we use only as an indicator of compliance type). Cell sizes for the treatment actually received (the cell "Ns") are indicated in the terminal nodes (the bottom row) of the diagram. The cell labels have two components. For the first component, DG (deliberative group) indicates the subject participated in a deliberative group; IO (information only) indicates the subject completed the informational background materials survey but not a session and hence received the information only treatment; and TC (true control) indicates the subject attended neither a session nor read the background material and hence is a true control subject. For the second component, R (responder) indicates the subject responded on the post-test survey; NR (nonresponder) indicates the subject was offered the posttest survey but chose not to respond; and NS (no survey) indicates the subject was not administered the post-test survey.

A.1.5 Administration of the Post-test Survey

About half way through the study, in negotiations over unexpectedly high costs with our survey vendor, we agreed to discontinue sending post-test surveys to subjects with the strongest histories of nonresponse. These were the subjects who, *ex ante*, were least likely to respond to the post-test survey and so most likely to need to have their responses imputed anyway. As a result, a total of 224 subjects, or about 11 percent of the

[4] In Table 4 we show that the pretreatment covariate marginals between those who select out at the filter stage (and hence can only be in one of the control groups) and those who do not (but end up in one of the control groups) are nearly identical.

[5] This statement relies on an exclusion restriction assumption that assignment does not affect respondents' propensity to comply with the survey.

sample, did not receive a post-test survey. But note that even if we had sent these subjects the survey, most of them would not have filled them out. We know this because the revision to the survey procedures occurred after we had fielded the study to more than half of the sample. As a result, 364 of the subjects we later identified as "chronic nonresponders" were sent a post-test survey, and among these, only 17.6 percent responded. Thus, among the 224 who were not sent the survey (assuming that the order of the districts does not matter), we likely would have observed only an additional 39 surveys returned, which is only 2 percent of the total sample.

In principle, this revision to the survey procedures poses little problem for our statistical analysis. In the compliance model, response to the post-test survey is missing for these subjects; we do not treat these "nonresponses" as behavioral data, in that the statistical model does not assume this "failure" to respond to the post-test survey reveals any additional information at all about these subjects' compliance type. Instead, the model imputes their probability distribution of compliance based on their observed behavioral data and covariates. The model below accommodates these imputations by allowing our uncertainty about whether and how the respondent would have responded to propagate through all estimated parameters of the statistical model (Tanner and Wong, 1987).[6]

A.1.6 District Panel Sizes and Subcontracting for Subjects

As we mention in the text, Knowledge Networks (now GfK) maintains panels of potential survey respondents that are demographically representative. Because of the effort and resources required to maintain these panels, however, the panels themselves are relatively small. Since our study blocked on congressional districts (all treatment and control subject came from the 12 congressional districts in our study), KN's panels were not large enough in each congressional district to meet our size requirements. As a result, KN subcontracted with two other high quality online survey vendors, Survey Sampling International and GMI. While both of these subcontracting vendors maintain high quality panels, neither goes to the same lengths as KN to maintain representative samples. Since the panels that SSI and GMI maintain are very similar, we use panel fixed effects (KN versus SSI/GMI) to account for the differences between KN panels and the other two panels.[7]

[6] We only administered the November survey to participants who completed a post-test survey, and/or who participated in a deliberative session. We do not use responses to the November survey questions in any way in this analysis. Instead, we only use an indicator of whether or not participants returned this survey as an additional behavioral compliance indicator. For those who are not administered a November survey, we impute a probability distribution for their response based on pretreatment data and the latent compliance variable, just as we do for any other missing compliance indicators, such as for compliance with the treatment for those assigned to the control. The logic of this restriction, the imputation of their response if missing, and the consequences for estimation are identical to that of the post-test survey.

[7] Using separate fixed effects for SSI and GMI never yielded significant differences, so we collapse these to a single category.

That KN recruited some of our subjects from these other vendors means that the inferences we make in this study are limited to the population of subjects who join online survey panels, which themselves reflect the larger population that is well-connected to the Internet. We certainly cannot make inferences about how the average American would respond to our experiment if exposed to it, particularly those who stand on the other side of the digital divide. We do not see this as much of a limitation. If members of Congress were to one day adopt deliberative online town halls more broadly, it is the connected population who would likely be the ones to attend, and our sample is well-designed to assess the impacts on this population. And even if our study were limited to KN panelists, we likely would make this restriction in any case.

A.2 Variables

This section describes the esteem outcomes, the familiarity indicators, the compliance indicators and the covariates that are included in the statistical models. All variables are listed in appendix tables 2 and 3. The data describing congressional district indicators are listed in the text table 1.

A.2.1 Esteem Outcomes

We use the following outcome measures from the post-treatment (post-test) survey that, taken together, measure the constituent's esteem for the member, which is the main outcome of the causal mediation model. We represent the esteem scale with η_5.

The Feeling Thermometer (fol11c) The feeling thermometer score is a summary of the constituent's attitude (Wilcox et al., 1989; Winter and Berinsky, 1999). The respondent was presented with a feeling thermometer slider, with scores scaled to range from 0 to 100, and the text ''[MOC], your Member of Congress'' where the member's name appeared in place of [MOC]. Higher scores indicate "warmer" attitudes toward the member.[8] Nonresponses on the outcome variable are set to missing.

Trust in Member (fol17) "How much of the time do you think you can trust [MOC], your Member of Congress, to do what is right?" (1 if responded always or most of the time, 0 otherwise, and NA if missing. Dichotomized to maintain marginals in the same and opposite party subsamples).

Approval of Member (fol22) "Do you approve of the way that [MOC] is handling [MOCPRONOUN] job as Congressperson?" (3 if strongly/approve, 2 if neither, 1 if strongly/disapprove, and NA if missing).

[8] This feeling thermometer was presented among a battery of thermometer measures, and the order of the thermometers rotated randomly.

Approval of Member on Immigration Policy (fol29) “Do you approve or disapprove of the way [MOC], your Member of Congress, is handling the issue of immigration?” (3 if strongly/approve, 2 if neither, 1 if strongly/disapprove, and NA if missing).

Intend to Vote for Member “If the vote for the House of Representatives were held today, who would you vote for?” (4 if definitely member, 3 if probably member, 3 if probably opponent, 4 if definitely opponent, missing otherwise).

A.2.2 Compliance Indicators

In order to satisfy the sequential ignorability assumption in the causal mediation model, we condition both post-treatment familiarity and post-treatment esteem with a compliance type variable. This latent variable adjusts for the respondent’s propensity to comply with the experimental protocol and in addition it models any dependence between the familiarity latent and the esteem latent that remains after conditioning on covariates (Esterling et al., 2011). We estimate compliance type (represented by ξ_2) using the following indicators. In each case, we indicate compliance with a 1 and noncompliance with a 0.

Participated in session Attended the deliberative online town hall, if assigned (1 if yes, 0 if no, NA if assigned to the information only arm).

Completed post-test survey Responded to the post-test survey, if administered (1 if yes, 0 no, NA if not administered the survey).

Completed November survey Responded to the November post-election survey, if administered (1 if yes, 0 if no, NA is not administered the survey).

A.2.3 Familiarity Indicators

Familiarity is the mediating construct (η_4) in our causal mediation model and we estimate this using the post-treatment survey. We also explore the construct validity of our measure of familiarity in the construct validity model (via η_3) using the pretreatment survey. We use the following identical set of items from each of these two surveys to measure pretreatement and post-treatment familiarity. Note that all of these items included a Don’t Know option (partial filter) as a response option. In each case, a substantive response is coded 1, a *DK* is coded 0, and otherwise set to missing.

Trust (fol17) Did not respond “Don’t Know” to (fol17) “How much of the time do you think you can trust [MOC], your Member of Congress, to do what is right?” (1 if a response other than “Don’t Know,” 0 if responded “Don’t know,” and NA if missing).

Approval (fol29) Did not respond “Don’t Know” to (fol29) “Do you approve or disapprove of the way [MOC], your Member of Congress, is handling the issue of immigration?" (1 if a response other than "Don't Know," 0 if responded "Don't know," and NA if missing).

Knows How Member Voted Did not respond "Don't Know" to (fol33a) "How about your [MOC], your Member of Congress? Do you think [MOCPRONOUN] voted for or against making it a felony to assist illegal immigrants in entering or remaining in the US?" (1 if a response other than "Don't Know," 0 if responded "Don't know," and NA if missing).

As we note in the list of esteem outcome measures, the data have two indicators of approval of the member, the general question about job approval, and the specific question about approval of the member's handling of immigration policy. Unfortunately, the mediation model fails to converge for empirical reasons in the opposite-partisan subsample when we include both the trust and the general approval $\neg DK$ items to measure the familiarity latent. The tetrachoric correlation between the general and the specific approval $\neg DK$ measures are correlated at 0.84 in the full sample (0.85 for same partisans and 0.82 for opposite partisans), while the specific approval measures correlates with trust at 0.60 (0.64 among same partisans, 0.52 among opposite partisans), which indicates both approval $\neg DK$ measure the same latent construct. Further, the mediation models that include the general approval item that do converge (the full sample and the same-partisan sample) yield results that are identical to the restricted model to two significant figures.
To maintain consistency across all models, and to ensure that we only report results that are robust and stable, we report the results from the models that do not include the general approval indicator.

A.2.4 Policy Preference Familiarity and Policy Knowledge Familiarity

In the construct validity model we compare the properties of our familiarity latent (η_3) with latent measures of the respondent's policy preference familiarity (η_1) and her policy knowledge familiarity (η_2). The policy preferences are regarding approval of President Bush, which political party is preferable on immigration, a pathway to citizenship and making assisting illegal immigrants a felony. The knowledge questions are on a range of factual questions relevant to immigration debates in 2006. Each of these items had a partial *DK* filter option. In each case, a substantive response is coded 1, a *DK* is coded 0, and otherwise set to missing. We show the full wordings and responses for the policy preference and policy knowledge items in the last section of the appendix.

Table A1: Descriptives for Outcomes and Indicators

	Pretest	Post-test	Latent Scale, Parameter
	Mean (SD) N	Mean (SD) N	
Familiarity Indicators			η_1, λ_{11}
Approve Bush Immigration ¬*DK*	0.95 (0.21) 757		
Best Party on Immigration ¬*DK*	0.80 (0.40) 755		η_1, λ_{12}
Pathway to Citizenship ¬*DK*	0.96 (0.19) 760		η_1, λ_{13}
Felony to Assist ¬*DK*	0.94 (0.23) 760		η_1, λ_{14}
Number Illegal Immigrants Reside US ¬DK	0.77 (0.42) 759		η_2, λ_{21}
Number Illegal Immigrants Enter US ¬DK	0.66 (0.47) 756		η_2, λ_{22}
Fraction from Mexico ¬*DK*	0.81 (0.40) 754		η_2, λ_{23}

Felony to Reside ¬*DK*	0.74 (0.44) 757		η_2,λ_{24}	
Companies Employment Rules ¬DK	0.66 (0.48) 757		η_2,λ_{25}	
Eligible to Apply for Citizenship ¬DK	0.63 (0.48) 757		η_2,λ_{26}	
Trust ¬*DK*	0.79 (0.40) 752	0.85 (0.36) 580	η_3,λ_{31}	η_4,λ_{41}
Approve ¬*DK*	0.62 (0.49) 756	0.82 (0.44) 582	η_3,λ_{32}	η_4,λ_{42}
Knows How Member Voted ¬*DK*	0.35 (0.48) 753	0.52 (0.50) 581	η_3,λ_{33}	η_4,λ_{43}
Outcomes Member Feeling Thermometer		61.89 (21.26) 580	η_5,λ_{51}	
Trust Member		0.49 (0.50) 494	η_5,λ_{52}	
Approve Member		2.39 (0.75) 533	η_5,λ_{53}	
Approve Member on Immigration		2.22 (0.79) 475	η_5,λ_{54}	
Will Vote for Member		0.75 (0.43) 398	η_5,λ_{55}	
Compliance Indicators Participated in Discussion		0.51 (0.50) 509	ξ_2,γ_{21}	
Completed Post-test Survey		0.79 (0.41) 760	ξ_2,γ_{22}	

Completed November Survey	0.87 (0.34) 622	ξ_2, γ_{23}

Note: Compliance indicators set to missing for respondents not eligible for the corresponding study activity.

A.2.5 Covariates

We condition every element of the causal mediation model (the familiarity, compliance and esteem scales) using the following set of covariates. The covariates and scales are a relevant subset of those found in Neblo et al. (2010). In addition, we explore the covariance between each of these covariates and the three familiarity latents in the construct validity model (η_1, η_2, and η_3).

Treatment Exposure If randomized to receive an invitation, coded 1 if participated in the deliberative online town hall, 0 otherwise. If no invitation set to missing. This is the indicator for our causal intervention.

Same-partisan, Opposite-partisan correspondence between members party and constituents Party ID on a 3 point scale. We use this to condition in the full sample, and to split the sample to test for differences by partisanship.

Education College degree or more education; NA if missing

Income Reported income at or above $50k (the median category); NA if missing

Basic Demographics Indicator variables for gender and race; NA each if missing.

General Political Knowledge From Delli Carpini and Keeter (1996). Sum of Delli Carpini and Keeter 5 dichotomized to 0 is fewer than 3 correct answers, 1 is 3 or more correct.

KN panelist 1 if respondent is from the Knowledge Networks panel; 0 if from a GMI or SSI panel.

Self Efficacy First principal component of the items, Please tell us how much you agree or disagree with the following statements, (bas13) I don't think public officials care much what people like me think. (bas14) Sometimes politics and government seem so complicated that a person like me can't really understand what is going on. And (bas15) I have ideas about politics and policy that people in government should listen to. Scaled so high values indicate high efficacy. Dichotomized so that high values indicate above median value. Missing values set to median.

Need for Cognition From Cacioppo et al. (1984). Additive scale constructed from (bas45) Would you say you have opinions about 1 if responds almost everything or about many things, 0 otherwise, NA if missing, and from (bas46) Some people like to have responsibility for handling situations that require a lot of thinking, and other people don't like to have responsibility for situations like that. Do you 1 if responds like them a lot or like them somewhat 0 otherwise, NA if missing. Dichotomized so that 2 is recoded to 1, and 1 and 0 are zero. Missing responses set to median.

Need to Evaluate From Bizer et al. (2004). Additive scale constructed from (bas47) Please tell us how much the statement below describes you: It is very important to me to hold strong opinions. 1 if extremely or somewhat characteristic, 0 otherwise, NA if missing, and from Please tell us how much the statement below describes you: I often prefer to remain neutral about complex issues 1 if somewhat or extremely uncharacteristic, 0 otherwise, NA if missing. Dichotomized so that 2 is recoded to 1, and 1 and 0 are zero. Missing responses set to median.

Strong Partisan Self reports as either strong Democrat or Strong Republican on a 7 point scale; NA if missing

Interest in immigration news reporting 1 if responded followed closely or somewhat to (bas26) Recently there has been a lot of reporting about the issue of illegal immigration. Would you say you have 0 otherwise, missing set to median.

Congressional District Indicators for congressional district, see table 1

The descriptives for these covariates are in table 3. Table 3 gives a breakdown of subjects' pretreatment covariate averages by the treatment they actually received. We present this table as descriptive information for the sample only; it is not intended to indicate balance between the different groups across these covariates, nor does the statistical model require balance.[9] We do note, however, that when comparing the DG to the IO exposure groups, the raw data do not begin with any marked departures from balance across these groups. There is a slight difference on dimensions for which one would expect to observe imbalance between deliberators and controls, especially in college degree, political knowledge, and need for cognition and need to evaluate. In any case, the statistical model condition this full set of covariates, so the causal model is fully stratified.

[9] As we note above, the treatment is ignorable since the model controls for compliance type principal strata.

Table 3: Covariates by Treatment Group Exposure

	Combined DG+IO	DG	IO	TC
	Mean (SD) N	Mean (SD) N	Mean (SD) N	Mean (SD) N
Treatment (DG) Exposure	0.34 (0.48) 760	1 262	0 498	0 1200
Opposite Party	0.40 (0.49) 760	0.37 (0.48) 262	0.42 (0.49) 498	0.34 (0.47) 1200
College or More	0.45 (0.50) 760	0.52 (0.50) 262	0.41 (0.49) 498	0.38 (0.49) 1200
Income	0.62 (0.49) 760	0.61 (0.49) 262	0.62 (0.48) 498	0.58 (0.49) 1200
Male	0.32 (0.47) 760	0.36 (0.48) 262	0.31 (0.46) 498	0.27 (0.45) 1200
White	0.83 (0.38) 760	0.84 (0.37) 262	0.83 (0.38) 498	0.79 (0.41) 1200
Political Knowledge	0.70 (0.46) 760	0.78 (0.41) 262	0.65 (0.48) 498	0.60 (0.49) 1200

KN Panel	0.46	0.50	0.44	0.21
	(0.50)	(0.50)	(0.50)	(0.41)
	760	262	498	1200
Self-Efficacy	0.52	0.63	0.47	0.46
	(0.50)	(0.48)	(0.50)	(0.50)
	760	262	498	1200
Need for Cognition	0.55	0.66	0.50	0.48
	(0.50)	(0.48)	(0.50)	(0.50)
	760	262	498	1200
Need to Evaluate	0.56	0.65	0.52	0.48
	(0.50)	(0.48)	(0.50)	(0.50)
	760	262	498	1200
Strong Party ID	0.40	0.39	0.40	0.35
	(0.49)	(0.49)	(0.49)	(0.48)
	760	262	498	1200
Follows News	0.83	0.89	0.80	0.76
on Immigration	(0.38)	(0.32)	(0.40)	(0.43)
	760	262	498	1200

The DG+IO sample includes all respondents who completed the background materials survey.

The statistical results we report in the paper are local to the subsample of respondents that read the background materials, and hence the results only generalize to the population of respondents who are engaged enough to read the reading materials. As we show above, 67 percent of respondents who were given the opportunity to read the background materials did so, and so represent a majority of respondents. The final column in table 3 shows the covariate averages for respondents who did not read the background materials and were excluded from the analysis (that is, the respondents in the true control (TC) condition of the larger experiment.) Comparing the TC respondents to the IO respondents shows there are differences on the same dimensions where we observe imbalance between deliberators and controls, such as college degree, political knowledge, and need for cognition and need to evaluate. While we cannot make an inference about the causal effect of those in the TC condition, we do note that the covariate differences indicate that the TC subsample is not radically different from the IO subsample.

Table 4: Covariates by Design Components

	RSVP Yes	RSVP No	Assigned DG	Assigned IO	Assigned TC
	Mean (SD) N	Mean (SD) N	Mean (SD) N	Mean (SD) N	Mean (SD) N
Opposite Party	0.37 (0.48) 1524	0.35 (0.48) 548	0.35 (0.48) 1084	0.43 (0.50) 376	0.35 (0.48) 612
College or More	0.41 (0.49) 1524	0.41 (0.49) 548	0.43 (0.49) 1084	0.38 (0.49) 376	0.39 (0.49) 612
Income	0.59 (0.49) 1524	0.62 (0.49) 548	0.58 (0.49) 1084	0.61 (0.49) 376	0.60 (0.49) 612
Male	0.28 (0.45) 1524	0.31 (0.46) 548	0.29 (0.46) 1084	0.28 (0.45) 376	0.29 (0.46) 612
White	0.79 (0.41) 1524	0.84 (0.36) 548	0.78 (0.42) 1084	0.82 (0.39) 376	0.84 (0.37) 612
Political Knowledge	0.63 (0.48) 1524	0.68 (0.48) 548	0.65 (0.48) 1084	0.61 (0.49) 376	0.62 (0.49) 612
KN Panel	0.25 (0.43) 1524	0.42 (0.49) 548	0.25 (0.43) 1084	0.34 (0.48) 376	0.34 (0.48) 612

Self-Efficacy	0.49	0.47	0.50	0.49	0.44
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
	1524	548	1084	376	612
Need for Cognition	0.55	0.44	0.56	0.45	0.47
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
	1524	548	1084	376	612
Need to Evaluate	0.53	0.47	0.55	0.48	0.47
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
	1524	548	1084	376	612
Strong Party ID	0.31	0.34	0.37	0.38	0.34
	(0.48)	(0.48)	(0.48)	(0.49)	(0.48)
	1524	548	1084	376	612
Follows News	0.81	0.73	0.82	0.76	0.77
on Immigration	(0.39)	(0.44)	(0.39)	(0.43)	(0.42)
	1524	548	1084	376	612

A.3 Bayesian Estimation Procedures

We implement both the compliance and the causal mediation statistical model in the MCMC sampling software OpenBUGS (Spiegelhalter et al., 1996). This sampler sequentially uses Bayes' Rule to update the parameter values until the posterior distribution converges to a stationary distribution; the resulting marginal stationary posterior distributions serve as the parameter estimates (see Jackman, 2000). One can characterize the point estimates and standard errors of all structural parameters, as well as functions of these parameters such as those we report in the manuscript, using the resulting posterior distribution. We approximate maximum likelihood (ML) estimates by assigning flat priors for all parameters.

Below we reproduce the OpenBUGS code for the construct and mediation models, graphically shown in figures 4 and 1 of the paper. The code contains the "computer" variable names. To see how the data for the outcome (O), familiarity indicators (S), compliance indicators (C) and covariates (X.O, X.S, and X.C), and the initial values are created, and how they correspond to the notation in the paper, refer to the set up file included with the replication data.

For identification, we must scale each latent variable to one observed or measured variable; the choice here is arbitrary. We scale each of the scales to one of the indicators by assigning a value of one for the corresponding factor coefficient for the latent variable. Since the response functions have no natural scale, the variance of the latent variables are not identified; we follow standard practice in item response models and set these variances to one.

A.3.1 Imputing Missing Data for the Endogenous Variables

It is widely recognized that discarding observations that contain missing data may cause biased estimates in any statistical method unless the data happen to be missing completely at random (MCAR) (e.g., Barnard et al., 2003). Frangakis and Rubin (1999) proposed the method of principal stratification to address the problems of both noncompliance and missing outcome data. Under the latent ignorability assumption,[10] the missing esteem and familiarity outcome responses and the compliance indicators are conditionally independent within strata of the compliance type variable, and hence imputation through data augmentation enables unbiased estimates of treatment effects. This conditional independence assumption is standard in latent variable models, such as IRT models (see Treier and Jackman, 2013). With data augmentation, the uncertainty inherent in the imputation is propagated through the posterior model parameters (Tanner and Wong, 1987).

A.4 Construct Validity Model

While an epistemic lack of familiarity is a primary driver of the *DK* response, it need not be the only driver and indeed there may be confounding factors that reduce the construct validity of this measure (Atir et al., 2015; Mondak, 1999). We use the presurvey data to understand the construct validity of our measure of member familiarity by comparing this measure with similar measures of familiarity toward two additional objects. First, a respondent might choose *DK* due to general disinterest in politics and policy. In this case, an individual would be unwilling to respond to attitude questions on even the most important policy items. Second, a respondent might choose *DK* because of a selfperceived lack of knowledge. In this case, the individual would be unwilling to respond to knowledge-based questions on policy issues. We are able to identify the systematic component of our member familiarity measure by comparing its properties to those of the other two.

We isolated these constructs using the pretest data and a confirmatory factor model (Bollen, 1989). For survey items that contain *DK* in the response set, we label substantive responses $\neg DK$ (coded 1) to contrast with *DK* responses (coded 0).[11] In short, the model

[10] As we discuss above, latent ignorability assumes that compliance type is correlated both with the outcomes and with the missing data process.

[11] Each of these questions offered the respondent a partial *DK* filter, which is useful for these analysis in that including a filter signals the acceptability of satisficing (Krosnick, 2002); that is, the filter increases the rate of *DK*s and hence increases variability on our response indicator variables.

enables us to estimate respondents' propensity to respond DK or $\neg DK$ across a wide range of items, that is, across three different substantive scales, as well as a second-level generalized disposition to respond DK.

More specifically, the model estimates three latent variables[12] (η_1,η_2,η_3) that measure the propensity to respond $\neg DK$ rather than DK on three distinct sets of items, where ¬ is the symbol for "not" or "false." First, the items measuring η_1 are indicators of whether the respondent chose DK (= 0) or $\neg DK$ (= 1) on a set of items measuring *policy preferences* on immigration policy options. Second, the items measuring η_2 indicate whether the respondent chose DK or $\neg DK$ on a set of questions on immigration *policy knowledge*. And third, the items measuring η_3 indicate whether the respondent chose DK or $\neg DK$ on a set of questions regarding the *personal attitudes toward their own member of Congress*.[13] The three latent variables are themselves nested within a fourth scale (ξ_1) which captures the systematic component.

The construct validity model is diagrammed in figure 4. This model measures three latent variables η and regresses these on a common factor ξ_1. The code that we reproduce here shows the full model likelihood and priors.

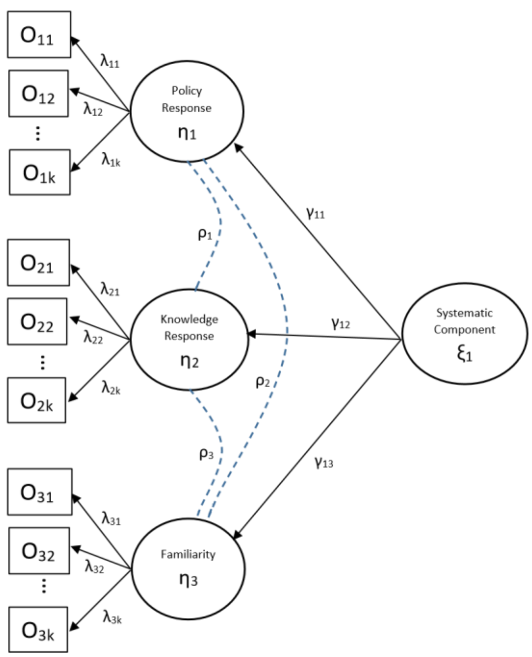

Figure 4: Construct Validity Model

[12] Our notation for latent variables follows that of Bollen (1989), where ξ indicates an exogenous latent variable and η indicates an endogenous latent variable.

[13] As we explain in detail in the appendix, this set of variables excludes one of the two measures of approval of the MC that we include in our model of esteem below. The reason is that if we include it, the model fails to converge in some instances. Since the two measures of approval have identical properties for measurement, excluding one has no impact on the construct validity of the measure.

In the confirmatory factor model, a typical indicator equation set is:

$$\begin{aligned} O^*_{mk} &= \alpha_{mk} + \lambda_{mk}\eta_m + \epsilon_{mk} \\ \eta_m &= \gamma_m \xi_1 + \mathbf{X}\boldsymbol{\beta} + \epsilon_{\eta_m} \end{aligned} \tag{3}$$

for m = 1 to 3, k = 1 to K_m, and K_m is the number of indicators on the m^{th} scale. $O^*_{\cdot}$ is either a continuous observed indicator variable, or it is the latent index for a discrete indicator variable. In this approach, $\gamma_m\xi_1$ scales the systematic component of η_m, and ϵ_{η_m} is the residual that is relevant to the items in set $\mathbf{O_m}$. The η_m equation optionally includes covariates $\mathbf{X}$ and a parameter vector β.

The first column of table 5 summarizes the content of each latent variable in the construct validity model. The latent variables in the second column are defined in the mediation model of the next section.

Table 5: Latent Variables

	Factor Model	Mediation Model
Systematic Component	ξ_1	
Knowledge Item Response	η_1	
Policy Preference Item Response	η_2	
Familiarity	η_3	η_4
Esteem		η_5
Compliance		ξ_2

A.4.1 Variables

Each of the η_m scales is measured by a set of items $\mathbf{O_m}$ indicating whether or not the respondent chose the ¬*DK* option. The disclosure appendix provides details regarding the coding and the exact wording of all questions and response sets.

The $\mathbf{O_1}$ items measuring η_1 are with respect to policy attitudes. These are, “Do you approve or disapprove of the way George W. Bush is handling the issue of immigration?” (five point approve-disapprove response scale); “Regardless of how you usually vote, do you think the Republican Party or the Democratic Party is more likely to make the right decisions when it comes to dealing with immigration issues?” (Democratic Party, Republican Party, or No Difference); “Now we”d like to ask about proposals to give some illegal immigrants who have resided in the U.S. for many years the opportunity to eventually become legal citizens. Some argue that providing opportunities for citizenship would reward illegal behav ...” (for or against); and “Now we”d like to ask about proposals to make it a felony to assist illegal immigrants to enter or remain in the US.

Some argue that we need tough penalties to stop people who smuggle illegal immigrants into the US, and help them remain here. Others ...” (for or against).

The $\mathbf{O_2}$ items measuring η_2 are with respect to policy knowledge specific to U.S. immigration and border control policy. These are, “About how many illegal immigrants do you think currently reside in the United States?” (five response options); “Do you know about how many new illegal immigrants come into the U.S. each year?” (five response options); “Do you know about what fraction of illegal immigrants in the U.S. are from Mexico?” (five response options); “Under current law, is it a felony to reside illegally in the United States?” (yes or no); “Under current law, do companies who want to employ non-citizen immigrants have to prove that doing so will not hurt the employment of U.S. citizens?” (yes or no); and “Under current law, are illegal immigrants who have lived in the U.S. for five years or more eligible to apply for citizenship?” (yes or no).

The $\mathbf{O_3}$ items measuring η_3 are with respect to attitudes regarding the member him or herself. These are, “How much of the time do you think you can trust [MOC], your Member of Congress, to do what is right?” (always, most of the time, some of the time, not at all); “Do you approve or disapprove of the way [MOC], your Member of Congress, is handling the issue of immigration?” (five point approve-disapprove response scale); and “How about [MOC], your Member of Congress? Do you think [MOCPRONOUN] voted for or against making it a felony to assist illegal immigrants in entering or remaining in the US?” (for or against). We use these items to measure our familiarity latent variable.

A.4.2 Model Code

The following code implements the construct validity structural model. To see how the computer notation corresponds to the notation in the paper, see the replication material.

```
### BEGIN OpenBUGS MODEL model{

for (i in 1:n){ ###

## DK scale equations:

        S1[i,1] ~ dbern(p.S1.b[i, 1]) # Policy Attitude Familiarity p.S1.b[i, 1] <- min(.999, max(p.S1[i, 1], .001)) logit(p.S1[i, 1]) <- (constant.S1[1] + 1*eta.S[i,1]) # coef set to 1 for identification for (j in 2:n.S1) {
        S1[i, j] ~ dbern(p.S1.b[i, j])
```

```
            p.S1.b[i, j] <- min(.999, max(p.S1[i, j], .001)) logit(p.S1[i, j]) <- (constant.S1[j] +
            lambda.eta.S1[j]*eta.S[i,1]) }

            S2[i,1] ~ dbern(p.S2.b[i, 1]) # Policy Knowledge Familiarity p.S2.b[i, 1] <- min(.999,
max(p.S2[i, 1], .001)) logit(p.S2[i, 1]) <- (constant.S2[1] + 1*eta.S[i,2]) # coef set to 1 for identification
for (j in 2:n.S2) {
            S2[i, j] ~ dbern(p.S2.b[i, j])
            p.S2.b[i, j] <- min(.999, max(p.S2[i, j], .001)) logit(p.S2[i, j]) <- (constant.S2[j] +
            lambda.eta.S2[j]*eta.S[i,2]) }

            S3[i,1] ~ dbern(p.S3.b[i, 1]) # Member Attribute Familiarity p.S3.b[i, 1] <- min(.999,
max(p.S3[i, 1], .001)) logit(p.S3[i, 1]) <- (constant.S3[1] + 1*eta.S[i,3]) # coef set to 1 for identification
for (j in 2:n.S3) {
            S3[i, j] ~ dbern(p.S3.b[i, j])
            p.S3.b[i, j] <- min(.999, max(p.S3[i, j], .001))
            logit(p.S3[i, j]) <- (constant.S3[j] + lambda.eta.S3[j]*eta.S[i,3]) }
      }


for (j in 1:n){ eta.S[j,1:3] ~ dmnorm(mu.S[j,1:3], Phi[,])
mu.S[j,1] <- (0 + 1*eta.F[j]) mu.S[j,2] <- (0 +
lambda.eta.F[2]*eta.F[j]) mu.S[j,3] <- (0 +
lambda.eta.F[3]*eta.F[j])
      eta.F[j] ~ dnorm(mu.F[j],1)I(-5,5) mu.F[j] <- 0 #
}


Phi[1:3, 1:3] ~ dwish(Phi0[,], nu0)

Phi0[1,1]<-1
Phi0[2,2]<-1
Phi0[3,3]<-1
Phi0[1,2] <- 0
Phi0[1,3] <- 0
Phi0[2,3] <- 0
Phi0[2,1]<-Phi0[1,2]
Phi0[3,1]<-Phi0[1,3] Phi0[3,2]<-
Phi0[2,3] nu0<-3


## Priors for parameters

for (m in 1:3) {
   lambda.eta.F[m] ~ dlnorm(0, 5) #
}
```

```
for (m in 1:n.S1) {
   lambda.eta.S1[m] ~ dlnorm(0, 5) #
}
for (m in 1:n.S2) {
   lambda.eta.S2[m] ~ dlnorm(0, 5) #
}
for (m in 1:n.S3) {
   lambda.eta.S3[m] ~ dlnorm(0, 5) #
}

for (m in 1:n.S1) {
    constant.S1[m] ~ dnorm(0, .0001)
}
for (m in 1:n.S2) { constant.S2[m] ~
    dnorm(0, .0001)
}
for (m in 1:n.S3) { constant.S3[m] ~
    dnorm(0, .0001)
}

### end }
```

A.4.3 Estimation

We estimate the structural parameters, the latent variables, and missing data parameters using Bayesian MCMC methods (Spiegelhalter et al., 1996). For each model we draw repeatedly from a candidate posterior distribution until the posterior distribution of all parameters is stationary by the Gelman and Rubin (1992) diagnostic, and then we sample 10,000 draws from the stationary distribution (thinning by 10) in order to simulate a posterior distribution of the model parameters. The standard errors we report are the Bayesian credible intervals of the marginal posterior distribution of each parameter or function of parameters.

A.4.4 Results from the Construct Validity Model

Table 6 presents the estimates for the structural parameters of the construct validity model. The first column shows the results for the complete sample of all 2,072 respondents who participated in the experiment. This includes those who completed the background materials survey that we focus our analysis on, as well as those assigned to the true control group and those assigned to the information-only and deliberative town hall conditions who did not complete the background materials survey. The second column shows the structural parameter results for the subsample that we focus on, those who completed the background materials survey. Note that in comparing these two columns the results are nearly identical in the subsample as they are in the full sample.

Table 6: Construct Validity Results

	Parameter	Full Sample	IO Only†
		Mean (SE)	Mean (SE)
ξ_1 Coefficients			
Policy Pref. ¬DK (η_1)	γ_{11}	1 (–)	1 (–)
Policy Knowledge ¬DK (η_2)	γ_{12}	1.87 (0.80)	1.43 (0.49)
Member Familiarity ¬DK (η_3)	γ_{13}	1.15 (0.31)	1.10 (0.33)
Factor Coefficients‡			
Approve Bush Immigration ¬DK	η_1,λ_{11}	1 (–)	1 (–)
Best Party on Immigration ¬DK	η_1,λ_{12}	0.80 (0.11)	0.80 (0.19)
Pathway to Citizenship ¬DK	η_1,λ_{13}	1.32 (0.23)	0.87 (0.23)
Felony to Assist ¬DK	η_1,λ_{14}	1.64 (0.30)	1.23 (0.33)
Number Illegal Immigrants Reside US ¬DK	η_2,λ_{21}	1 (–)	1 (–)
Number Illegal Immigrants Enter US ¬DK	η_2,λ_{22}	1.31 (0.16)	1.42 (0.26)
Fraction from Mexico ¬DK	η_2,λ_{23}	0.84 (0.09)	1.15 (0.19)
Felony to Reside ¬DK	η_2,λ_{24}	0.27 (0.03)	0.35 (0.06)
Companies Employment	η_2,λ_{25}	0.34	0.47

Rules ¬*DK*		(0.03)	(0.08)
Eligible to Apply for Citizenship ¬*DK*	η_2, λ_{26}	0.33 (0.03)	0.46 (0.08)
Trust Member ¬*DK*	η_3, λ_{31}	1 (–)	1 (–)
Approve Member on Immigration ¬*DK*	η_3, λ_{32}	1.09 (0.13)	1.08 (0.22)
Knows How Member Voted ¬*DK*	η_3, λ_{33}	1.13 (0.16)	1.00 (0.22)

†The IO-only sample includes only respondents who completed the background materials survey. ‡Constants for each indicator equation are not reported.

Our primary interest is in η_3, which we take as our main measure of member familiarity. We call attention to three sets of findings in the construct validity model. First, we use the model to decompose the propensity to respond *DK* on each of the three dimensions (η) into its systematic component (ξ_1) and a residual component. This decomposition reveals that the systematic component accounts for much of the variance in each η: 42% for policy preference response, 70% for policy knowledge response, and 96% for member familiarity. Thus, nearly all of the variance in member familiarity is systematic, rather than due to some omitted or confounding element.

Second, including ξ_1 in the three η equations allows estimation of the correlations among the three scales that is due to their common dependence on ξ_1. We note that the estimated factor coefficients γ, giving the relationship between η_m and ξ_1, exceed 1 $(\widehat{\gamma_{11}} \equiv 1,\ \widehat{\gamma_{12}} = 1.43,\ (SE = 0.49),\ \widehat{\gamma_{13}} = 1.10,\ (SE = 0.33))$. Under the identification constraint $cov(\xi_1, \epsilon_m) = 0$ for $m = 1$ to 3, the dependence between the η_k scale and the η_m scale can be retrieved with:

$$\rho_{\eta_k,\eta_m} = \frac{\gamma_k \gamma_m \xi_1^2}{\sqrt{(\mathrm{var}(\eta_k))\,(\mathrm{var}(\eta_m))}} \tag{4}$$

for $k = 1$ to 3 and $m = 1$ to 3. This correlation between the η_k that is driven by the systematic component indicates the latent concepts that are most closely related. Here we find $\rho_{\eta1,\eta2} = 0.54$, $\rho_{\eta1,\eta3} = 0.63$, and $\rho_{\eta2,\eta3} = 0.82$, indicating that the systematic components of familiarity and policy knowledge are most closely associated in their systematic component. That is, member familiarity seems primarily related to factual retrieval in its similarity to willingness to respond to policy knowledge questions, rather than the (perhaps) more expressive policy preference dimension.

Third, the model estimates covariances among the unique $_k$ components by specifying the joint distribution among the η scales as multivariate Normal, and the vector ρ captures the covariances among the residuals components.[14] Figure 5 shows the posterior distribution for the three estimates for ρ. Notice that the residual component of having policy preferences covaries with the residual component of having policy knowledge. This is sensible and indicates some residual dependence between knowledge and preferences. Also notice that the residual component of the member attribute familiarity scale η_3 is correlated with the policy preference familiarity scale but not the policy knowledge response scale. This shows that the residual component of the member attribute familiarity scale is more closely related to that of the respondent's policy preference than it is to her policy knowledge.

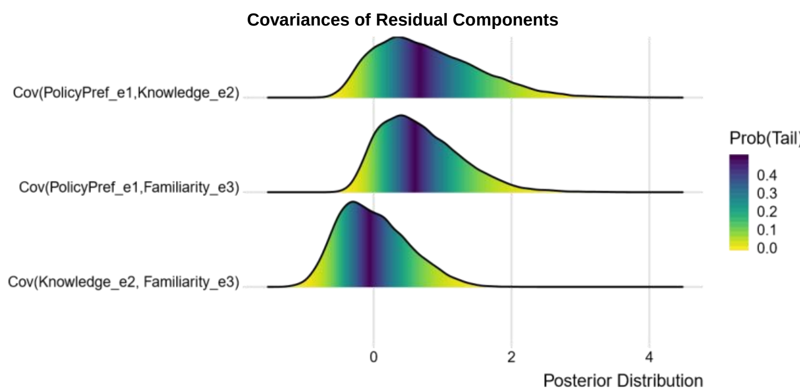

Figure 5: Covariances between the epistemic components

We expanded this model to include covariates in each equation for η_m, to determine the patterns of respondent attributes that might be correlated with each familiarity scale.[15]We include 13 covariates plus fixed effects for each congressional district. Overall, all three latent measures of familiarity are broadly unrelated to common demographic measures, political knowledge, party identification, level of partisanship and personality measures. Of the 39 estimated coefficients across the three equations, only eight were statistically significant (even without correcting for multiple comparisons), and only the covariate indicating interest in political news reporting appeared significant in all three

[14] Note that mathematically we must constrain the precision matrix in the joint distribution to be positive-definite and we do this by assigning it a diffuse Wishart prior.

[15] These covariates are Treatment Assignment, Opposite Party, College or More, Income, Male, White, Political Knowledge, KN Panelist, Self-Efficacy, Need for Cognition, Need to Evaluate, Strong Party ID, and Interest in Immigration News Reporting. See Appendix A.2 for details.

equations. Thus, familiarity is not confounded with commonly reported demographic and attitudinal measures (Jessee, 2017).

We re-estimated the model including the full sample of 2,072 participants (i.e., including all true controls as well as those who did not complete the background materials survey) and got substantively identical results. See table 6.

This confirmatory factor analysis warrants the measurement of familiarity using *DK* responses to items about attributes of *MC*'s.

A.5 Causal Mediation Model

The code we reproduce here shows the full model likelihood as well as all priors we use to estimate the causal mediation model. This model measures two latent variables, the familiarity mediator η_4, the esteem outcome η_5, and regresses these on a common factor ξ_2 that measures the respondent's propensity to comply with the experimental protocol. To see how the computer notation corresponds to the notation in the paper, see the replication materials.

```
### BEGIN WinBUGS MODEL model{

for (i in 1:n){ ###

## Compliance equations: C[i,1] ~
        dbern(p.C.b[i, 1])
        p.C.b[i, 1] <- min(.999, max(p.C[i, 1], .001))
        p. C[i, 1] <- phi(constant.C[1] + 1*eta.C[i]) # coef set to 1 for firstindicator, for identification
for (j in 2:n.C) {
        C[i, j] ~ dbern(p.C.b[i, j])
        p.C.b[i, j] <- min(.999, max(p.C[i, j], .001))
        p. C[i, j] <- phi(constant.C[j] + lambda.eta.C[j]*eta.C[i])}

## Mediation (familiarity) equations:
        S[i,1] ~ dbern(p.S.b[i, 1])
        p.S.b[i, 1] <- min(.999, max(p.S[i, 1], .001)) logit(p.S[i, 1]) <- (constant.S[1] + 1*eta.S[i]) #
        same as above; coef set to 1 for identification
for (j in 2:n.S) {
        S[i, j] ~ dbern(p.S.b[i, j])
        p.S.b[i, j] <- min(.999, max(p.S[i, j], .001)) logit(p.S[i, j]) <- (constant.S[j] +
        lambda.eta.S[j]*eta.S[i]) }


## Outcome equations:
        O[i,1] ~ dnorm(mu.O1[i], tau.O) # ft.membr.fol mu.O1[i] <- (lambda.eta.O[1]*eta.O[i] +
        k1) # be sure NOT to include a constant in X.O
```

```
        O[i,2] ~ dbern(p1.bound[i]) # trust.fol -- DKs set to missing p1.bound[i] <-
                        max(0,min(1,p1[i])) logit(p1[i]) <- (k2 + 1*eta.O[i])

        O[i,3] ~ dcat(p.O3[i,1:3]) # approve.fol -- DKs set to missing p.O3[i,1] <- 1-q2[i,1]
                p.O3[i,2] <- q2[i,1]-q2[i,2]
                p.O3[i,3] <- q2[i,2]
        logit(q2[i,1]) <- (-k3[1] + lambda.eta.O[3]*eta.O[i])
        logit(q2[i,2]) <- (-k3[2] + lambda.eta.O[3]*eta.O[i])

        O[i,4] ~ dcat(p.O4[i,1:3]) # appimmig.fol -- DKs set to missing p.O4[i,1] <- 1-q3[i,1]
                p.O4[i,2] <- q3[i,1]-q3[i,2]
                p.O4[i,3] <- q3[i,2]
        logit(q3[i,1]) <- (-k4[1] + lambda.eta.O[4]*eta.O[i]) logit(q3[i,2]) <- (-k4[2] +
        lambda.eta.O[4]*eta.O[i])

        O[i,5] ~ dcat(p.O5[i,1:4]) # whovoteplan -- DKs=3 set to missing, and 6 and 7
                p.O5[i,1] <- 1-q5[i,1]
                p.O5[i,2] <- q5[i,1]-q5[i,2]
                p.O5[i,3] <- q5[i,2]-q5[i,3]
                p.O5[i,4] <- q5[i,3]
        logit(q5[i,1]) <- (-k5[1] + lambda.eta.O[5]*eta.O[i]) logit(q5[i,2]) <- (-k5[2] +
        lambda.eta.O[5]*eta.O[i]) logit(q5[i,3]) <- (-k5[3] + lambda.eta.O[5]*eta.O[i])
    }

for (j in 1:n){ eta.O[j] ~ dnorm(mu.O[j],tau.eta.O)I(-5,5) mu.O[j] <- (inprod(g.X.O[1:n.X.O], X.O[j,]) +
    lambda.eta.C[n.C+1]*eta.C[ j] + lambda.eta.S[n.S+1]*eta.S[j] + g.X.O[n.X.O+1]*X.O[j,1]*eta.S[j
            ]) eta.C[j] ~ dnorm(mu.C[j], tau.eta.C)I(-5,5)
    mu.C[j] <- (inprod(g.X.C[], X.C[j,]))
    eta.S[j] ~ dnorm(mu.S[j], tau.eta.S)I(-5,5) mu.S[j] <- (inprod(g.X.S[], X.S[j,]) +
        lambda.eta.C[n.C+2]*eta.C[j])
}

## Priors for parameters

tau.O <- pow(sigma.O, -2) sigma.O ~ dunif(0, 100) tau.eta.C <- 1 # note this parameter is
not identified when the scaling indicator is dichotomous
tau.eta.S <- 1 # note this parameter is not identified when the scaling indicator is dichotomous
tau.eta.O <- 1 # note this parameter is not identified when the scaling indicator is dichotomous

for (m in 1:n.O) { lambda.eta.O[m] ~
dunif(0,100)
} for (m in 1:n.X.O+1) {

    g.X.O[m] ~ dnorm(0, .0001)
        }
```

```
# We need to set relatively tight priors for the compliance measurement model in order to endure
      convergence; this is simply to
# allow the scale to be estimated dynamically rather than setting weights as constants a priori
for (m in 1:n.C) {
   lambda.eta.C[m] ~ dlnorm(0, 5) # lambda.eta.C[1] # not updated in the model
} lambda.eta.C[n.C+1] ~ dnorm(0, .0001)
lambda.eta.C[n.C+2] ~ dnorm(0, .0001)


for (m in 1:n.C) { constant.C[m] ~ dnorm(0,
     .0001)
}

for (m in 1:n.X.C) {
       g.X.C[m] ~ dnorm(0, .0001)
              }

for (m in 1:n.S) {
   lambda.eta.S[m] ~ dlnorm(0, 5) #lambda.eta.S[1] not updated in the model
} lambda.eta.S[n.S+1] ~ dnorm(0, .0001)

for (m in 1:n.S) { constant.S[m] ~ dnorm(0,
     .0001)
}

for (m in 1:n.X.S) {
       g.X.S[m] ~ dnorm(.0, .0001)
              }

k1 ~ dnorm(.0, .0001) k2 ~ dnorm(.0, .0001)
k3[1]~dnorm(-1, 0.1)I(-10, k3[2])
k3[2]~dnorm(1, 0.1)I(k3[1], 10)
k4[1]~dnorm(-1, 0.1)I(-10, k4[2])
k4[2]~dnorm(1, 0.1)I(k4[1], 10)
k5[1]~dnorm(-1, 0.1)I(-10, k5[2])
k5[2]~dnorm(0, 0.1)I(k5[1], k5[3])
k5[3]~dnorm(1, 0.1)I(k5[2], 10)

### end
}
```

A.5.1 Results of the Causal Mediation Model

Appendix table 7 presents the structural parameters for the causal effect estimands and the factor coefficients, λ, for each measurement model. In appendix table 7, cells give the estimated parameters along with standard errors. All coefficients are statistically significant at $p < 0.05$ with the exception of the treatment interaction term. Statistically significant factor coefficients show the reliability of each indicator (see Bollen, 1989).

As a robustness check, we re-estimated the statistical model excluding the 137 respondents in our estimation sample who indicated they were unable or unwilling to attend a town hall for the study. As we describe above, our pretreatment survey began with a filter question asking respondents if they were able and willing to attend one of the upcoming town halls scheduled for their congressional district (i.e., RSVP'd yes), and if not (i.e., RSVP'd no), whether they still would be willing to take part in the surveys for the project. Among those who RSVP'd no but were willing to take the surveys, some were randomized to the information-only (IO) condition, and hence some ended up in the estimation sample for the present analysis.

Of course, in any encouragement design some of those assigned to treatment do not comply with the treatment and the method principle stratification in the GET model is designed to allow causal inference in the presence of noncompliance. Including this initial filter question did not change the design of our statistical test but instead only gave us some information at the start of the study about the likely compliance rate. Had we not asked the question, those who RSVP'd no would have been ordinary noncompliers, so including them in our model should not affect our causal inference.

Testing the model's robustness by excluding 137 IO respondents who RSVP'd no allows us to check the plausibility of the assumptions of our causal model. When we re-estimated the model excluding those in the IO condition who RSVP'd no, the results do not change in the slightest. The parameter estimate for T $\rightarrow$ Familiarity is 1.90 (0.30); that for Familiarity $\rightarrow$ Esteem is 0.97 (0.17), which combined yield the identical mediation effect. The estimate for the interaction T $\times$ Familiarity $\rightarrow$ Esteem is 0.11 (0.14) which also is identical.

Finally, Table 8 shows the regression parameter estimates for the covariates included in the model, for the familiarity, esteem and compliance equations. Overall, the covariates have very little explanatory power in the model. Opposite partisans have lower esteem for their members. Need for Cognition and Need to Evaluate are negatively related to esteem but positively related to familiarity. Consistent with Mondak and Anderson (2004), men score higher on the familiarity measure. Finally, none of the demographic covariates reach statistical significance to explain compliance with attending the online Deliberative Town Hall; this result is consistent with Neblo et al. (2010), which previously analyzed these data to find that participation at the online Deliberative Town Halls is broadly representative of constituents in the included congressional districts.

Table 7: Mediation Model Results (part 1 – main results)†

	Parameter	Overall	Same-Partisans	Opposite-Partisans
T → Familiarity	β_1	2.03 (0.25)	2.20 (0.32)	1.63 (0.46)
T → Esteem	β_3	-1.65 (0.41)	-1.89 (0.54)	-1.67 (0.95)
Familiarity → Esteem	β_4	0.91 (0.13)	1.23 (0.17)	0.77 (0.52)
T × Familiarity → Esteem (Interaction Term)	β_2	0.17 (0.14)	0.04 (0.18)	0.23 (0.27)
Factor Coefficients				
Trust Member $\neg DK$	λ_{41}	1 (–)	1 (–)	1 (–)
Approve Member on Immigration $\neg DK$	λ_{42}	1.01 (0.18)	0.96 (0.21)	0.95 (0.25)
Knows How Member Voted $\neg DK$	λ_{43}	0.83 (0.13)	0.84 (0.14)	0.89 (0.24)
Member Feeling Thermometer	λ_{51}	10.31 (0.64)	7.67 (0.59)	10.20 (0.87)
Trust Member	λ_{52}	1 (–)	1 (–)	1 (–)
Approve Member	λ_{53}	2.56 (0.36)	2.20 (0.46)	2.39 (0.75)
Approve Member on Immigration	λ_{54}	1.10 (0.11)	0.93 (0.13)	0.89 (0.15)
Will Vote for Member	λ_{55}	1.98 (0.23)	1.01 (0.19)	1.67 (0.30)

Compliance Indicators				
Participated in Discussion	γ_{21}	0.57	0.49	0.71
		(0.10)	(0.11)	(0.17)
Completed Post-test Survey	γ_{22}	1	1	1
		(–)	(–)	(–)
Completed November Survey	γ_{23}	0.55	0.78	0.57
		(0.13)	(0.27)	(0.15)
Familiarity	γ_{24}	-0.89	-0.71	-0.85
		(0.17)	(0.26)	(0.51)
Esteem	γ_{25}	1.39	0.75	2.06
		(0.24)	(0.35)	(0.39)

†Constants and covariate coefficients not reported.

Table 8: Mediation Model Results (part 2 – covariates)†

	Esteem Equation	Familiarity Equation	Compliance Equation
Opposite Party	-1.26	-0.19	-0.02
	(0.26)	(0.21)	(0.13)
College or More	0.71	-0.37	-0.09
	(0.28)	(0.22)	(0.15)
Income	0.35	-0.31	-0.19
	(0.26)	(0.21)	(0.14)
Male	-0.82	0.39	0.06
	(0.28)	(0.22)	(0.15)
White	0.17	-0.23	0.33
	(0.33)	(0.27)	(0.17)
Political Knowledge	-0.15	0.29	0.04
	(0.29)	(0.24)	(0.15)

KN Panel	-1.57	1.10	1.08
	(0.43)	(0.31)	(0.20)
Self-Efficacy	-0.55	0.36	0.25
	(0.27)	(0.22)	(0.14)
Need for Cognition	-0.66	0.45	0.25
	(0.27)	(0.22)	(0.14)
Need to Evaluate	-0.74	0.70	0.13
	(0.29)	(0.23)	(0.14)
Strong Party ID	0.23	-0.03	-0.07
	(0.24)	(0.20)	(0.13)
Follows News Immigration	0.01	0.39	-0.08
	(0.33)	(0.27)	(0.18)

References

Atir, S., E. Rosenzweig, and D. Dunning (2015). When Knowledge Knows No Bounds: Self-Perceived Expertise Predicts Claims of Impossible Knowledge. *Psychological Science*.

Barnard, J., C. E. Frangakis, J. L. Hill, and D. B. Rubin (2003). Principal Stratification Approach to Broken Randomized Experiments: A Case Study of School Choice Vouchers in New York City. *Journal of the American Statistical Association 98*(462), 299–323.

Bizer, G. Y., J. A. Krosnick, A. L. Holbrook, S. C. Wheeler, D. D. Rucker, and R. E. Petty (2004). The Impact of Personality on Cognitive, Behavioral, and Affective Political Processes: The Effects of Need to Evaluate. *Journal of Personality 72*(Oct.), 995– 1027.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York, N.Y.: John Wiley & Sons, Ltd.

Cacioppo, J. T., R. E. Petty, and C. Feng˜Kao (1984). The Efficient Assessment of Need for Cognition. *Journal of Personality Assessment 48*(May), 306–307.

Callegaro, M. and C. Disogra (2008). Computing Response Metrics for Online Panels. *Public Opinion Quarterly 72*(5), 1008–1032.

Delli Carpini, M. X. and S. Keeter (1996). *What Americans Know about Politics and Why It Matters*. New Haven, Conn.: Yale University Press.

Esterling, K. M., M. A. Neblo, and D. M. Lazer (2011). Estimating Treatment Effects in the Presence of Noncompliance and Nonresponse: The Generalized Endogenous Treatment Model. *Political Analysis 19*(Spring), 205–226.

Frangakis, C. E. and D. B. Rubin (1999). Addressing Complications of Intention-toTreat Analysis in the Combined Presence of All-or-None Treatment-Noncompliance and Subsequent Missing Outcomes. *Biometrika 86*(2), 365–379.

Gelman, A. and D. B. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7*(Nov.), 434–455.

Jackman, S. (2000). Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo. *American Journal of Political Science 44*(April), 369– 398.

Jessee, S. A. (2017). 'Don't Know' Responses, Personality, and the Measurement of Political Knowledge. *Political Science Research and Methods 5*(4), 711–731.

Krosnick, J. A. (2002). Survey Nonresponse. In R. M. Groves, D. A. Dliiman, J. L. Eltinge, and R. J. A. Little (Eds.), *Technometrics*, Volume 44, Chapter The Causes, pp. 301–301. New York: Wiley.

Mondak, J. A. (1999). Reconsidering the Measurement of Political Knowledge. *Political Analysis 8*(1), 57–82.

Mondak, J. J. and M. R. Anderson (2004, 5). The Knowledge Gap: A Reexamination of Gender-Based Differences in Political Knowledge. *Journal of Politics 66*(2), 492–512.

Neblo, M. A., K. M. Esterling, R. Kennedy, D. Lazer, and A. Sokhey (2010). Who Wants to Deliberate – and Why? *American Political Science Review 104*(3), 566–583.

Spiegelhalter, D., A. Thomas, N. Best, and W. Gilks (1996). BUGS 0.5: Bayesian Inference Using Gibbs Sampling Manual (version ii). Technical report, MRC Biostatistics Unit.

Tanner, M. A. and W. H. Wong (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association 82*(398), 528–540.

Treier, S. and S. Jackman (2013). Democracy as a Latent Variable Democracy as a Latent Variable. *American Journal of Political Science 52*(1), 201–217.

Wilcox, C., L. Sigelman, and E. Cook (1989). Some Like it Hot: Individual Differences in Responses to Group Feeling Thermometers. *Public Opinion Quarterly 53*, 246–257.

Winter, N. and A. Berinsky (1999). What's Your Temperature? Thermometer Ratings and Political Analysis. Technical report, Presented at the Annual Meeting of the American Political Science Association, Atlanta, GA.