UvA-DARE (Digital Academic Repository)

Survey context effects and implications for validity: Measuring political discussion frequency in survey research

Boukes, M.; Morey, A.C.

SURVEY CONTEXT EFFECTS AND IMPLICATIONS FOR VALIDITY: MEASURING POLITICAL DISCUSSION FREQUENCY IN SURVEY RESEARCH

MARK BOUKES*
ALYSSA C. MOREY

Although political discussion behavior is an important area of political communication research, analysis of the reliability and validity of political discussion survey measures has only recently begun. This study uses panel survey data to examine the effects of survey context on self-reported political discussion frequency measures (e.g., general political discussion, discussion about the economy), the moderating influence of individual difference variables, and implications for measurement validity. A quasi-experiment demonstrates that including a large (versus small) quantity of preceding issue-relevant questions leads to higher reports of political and economic discussion frequency, and this effect is greater for individuals with higher levels of political and economic interest, respectively. Results of a survey experiment reveal that inclusion of a preceding thought-listing question about the economy produces higher reports of economic talk frequency, but only among those who possess relatively expansive conceptualizations of the economy. Such survey context effects suggest problems with the construct validity of self-reported political discussion frequency measurements. Potential consequences of survey context for concurrent and predictive validity are assessed by examining relationships between discussion frequency and known correlates (e.g., education, interest) and outcomes (e.g., current affairs knowledge) of political talk. Results provide tentative evidence that discussion measures placed after a large (versus

small) quantity of issue-relevant questions may exhibit better criterion validity (quasi-experiment), whereas the order of a thought-listing question does not appear to impact criterion validity (experiment). Results of this study clearly underscore the need for additional research on self-report measures of political behavior, including political discussion frequency.

With political communication research relying extensively on survey measurements of political behavior, scholars have recognized the importance of examining the quality of these measures. This study contributes to a small but growing area of research on self-report measures of political discussion (e.g., Eveland, Morey, and Hutchens 2011; Fitzgerald 2013; Klofstad, Sokhey, and McClurg 2013; Sokhey and Djupe 2014) by analyzing the impact of survey context on general and economic political discussion frequency.

Using panel survey data, this study examines how two variations in survey context—the inclusion of a large (versus small) battery of preceding issue-relevant questions and a preceding (versus following) open-ended thought-listing task about a political issue—influence self-reported discussion frequency about politics (generally) and economics as a political topic (more specifically). Individual difference variables that may moderate survey context effects are also examined. Robust findings of survey context effects suggest problems with construct validity of political talk survey measures (see Cronbach and Meehl 1955). Examination of predictive and concurrent validity provide preliminary evidence that talk frequency measurements may exhibit higher validity when following a larger quantity of issue-relevant items, whereas validity seems unaffected by the order of a thought-listing task.

1. SELF-REPORTED POLITICAL DISCUSSION FREQUENCY

Political discussion generates higher quality opinions, promotes a civic-minded orientation, and lends legitimacy to political processes, outcomes, and systems (Habermas 1991; Chambers 2003; Barabas 2004; Delli Carpini, Cook, and Jacobs 2004). One of the most commonly studied aspects of political discussion, political discussion frequency is positively related to a number of democratically desirable outcomes, including political knowledge (Eveland 2004; Eveland and Hively 2009), opinion quality (Price, Cappella, and Nir 2002; Druckman and Nelson 2003; Conover and Searing 2005), and political participation (Leighley 1990; McLeod, Scheufele, and Moy 1999; Schmitt-Beck and Lup 2013).

Although a central concept in the literature, a standard measure of political discussion frequency has not yet been developed, nor has a thorough analysis

of measurement validity been undertaken. Across the many local, national, and cross-national surveys that assess political talk, political discussion frequency questions exhibit significant variability, including the specific words used to identify the content area of interest. Survey items may refer directly and exclusively to the term "politics," or include additional concepts that might either expand or qualify the term politics, such as "politics or public affairs" (Lee 2005; Valenzuela, Kim, and Gil de Zuniga 2012) or "government, elections, and politics" (Huckfeldt 2001; Mutz 2002; Huckfeldt 2007). Survey measures may also focus on specific political issues, ranging from broader concepts (economy) to more narrow concepts (Social Security) (Conover, Searing, and Crewe 2002; Barabas 2004; De Vreese and Boomgaarden 2006; Lee 2012).

With substantial variability in question wording evident across political discussion frequency measures, this study focuses on two separates measures of political talk frequency: (1) general political discussion and (2) discussion about the specific issue of the economy. The economy is not only one of the most prominent political issues (Fitzgerald 2013), but it has been used in a number of studies on political talk frequency (e.g., Barabas 2004; Lee 2012). The economy was also chosen for the practical reason that the panel survey data used in this study had a strong focus on political-economic issues.

2. SURVEY CONTEXT EFFECTS AND PRIMING

This study investigates the effects of survey context on self-reported political discussion frequency measures. Survey respondents are susceptible to contextual influences (Schwarz 1999). When ambiguity about a core concept exists, context effects are particularly likely to occur, as respondents are prompted to use the immediate context (including earlier encountered survey questions) to make assumptions about the pragmatic meaning of key concepts (Bishop, Oldendick, and Tuchfarber 1982; Schwarz 1999). Survey context effects on political discussion frequency measurements are thus likely to be observed, as politics is a highly flexible and porous construct that means different things to different people (Fitzgerald 2013; Morey and Eveland 2016), which would indicate problems with construct validity (Cronbach and Meehl 1955). When self-reported discussion frequency depends on survey context, the precision and validity of these measurements is called into question, as is the interpretation of these measurement instruments.

This study examines how two types of survey context effects influence self-reported political discussion frequency. First, this study examines the relationship between including a large, versus small, battery of preceding issue-relevant questions and self-reported political talk frequency. Second, this study examines the effects of including an issue-relevant (i.e., the economy) thought-listing task immediately before, versus after, general political and specific economic talk frequency items. Exposure to an external stimulus (e.g., a word or phrase in a survey question) activates the cognitive representation of

that concept in an individual's mind. This process not only increases the likelihood that this concept is used in subsequent judgments, a phenomenon known as *priming* (Aronson, Wilson, and Akert 2007), but cognitively associated concepts are also made more accessible through spreading activation. Research demonstrates that survey questions themselves serve as primes (Hyman and Sheatsley 1950; Tourangeau, Rips, and Rasinski 2000). Although thought-listing tasks are commonly used to measure the *extent* of priming, research demonstrates that the task of thought-listing itself also elicits priming of relevant concepts (Macrae, Stangor, and Milne 1994).

A large battery of issue-relevant questions or a preceding issue-relevant thought-listing task (i.e., about the economy) should, therefore, make a wide array of politically relevant concepts accessible in semantic memory. Increased accessibility of issue-relevant concepts may lead to higher levels of self-reported political discussion frequency. The boundaries of working memory conceptualizations of political discussion should temporarily expand, such that individuals "count" more discussions as having been political. Accessibility of politically relevant concepts might also prompt individuals to remember a larger number of political discussions in which they participated because the question encourages them to actively think about these conversations (Popping 2015).

> Hypothesis 1.A: A survey context with a larger (versus smaller) battery of preceding issue-relevant questions will be associated with higher levels of self-reported political and economic discussion frequency.

> Hypothesis 1.B: A survey context with a preceding (versus following) issue-relevant thought-listing question will lead to higher levels of self-reported political and economic discussion frequency.

3. INDIVIDUAL DIFFERENCES IN THE CONCEPTUALIZATION OF POLITICS

When survey context effects on self-report measurements vary across subgroups of the population, problems with construct validity become especially severe, suggesting that the exact same question may be interpreted differently across respondents (Cronbach and Meehl 1955). Individual differences that influence political priming may also moderate the impact of survey context effects. Studies examining potential moderators of political priming effects tend to focus on concepts related to political engagement (interest, sophistication; see McLeod, Kosicki, and McLeod 2009). Results have been mixed, with some researchers suggesting that the least politically engaged should be most vulnerable to priming effects (Iyengar, Kinder, Peters, and Krosnick 1984), and others arguing that priming effects should be strongest among the politically aware (Krosnick and Brannon 1993).

In this study, a politically oriented survey may increase the likelihood that politically relevant considerations are primed—relative to if no such survey were taken—more so among the less politically interested (who infrequently think about politics) than among the more politically interested (who regularly think about politics). Yet, answering issue-relevant questions will undoubtedly prime political concepts among all survey respondents. On the other hand, the extent to which politically relevant concepts are primed (the number of constructs primed and the strength of activation) will depend on the quantity and strength of associations between and among politically relevant considerations in semantic memory. After all, a large number of politically relevant constructs cannot be activated from memory by preceding survey questions if relevant constructs simply are not present (or are not strongly associated with other related concepts) in memory. Individuals interested in politics and the economy are particularly likely to have elaborate and extensive semantic models of political and economic concepts (Luskin 1990), respectively. Thus, preceding politically relevant survey questions should be most apt to influence the politically interested, while having little to no effect among those with low levels of political and economic interest (or who hold narrow conceptual definitions of these concepts).

> Hypothesis 2.A: The positive association between a survey context with a larger (versus smaller) battery of preceding issue-relevant questions and higher levels of self-reported political and economic discussion frequency will be strongest among those interested in the issue.

> Hypothesis 2.B: Positive priming effects of a survey context with a preceding (versus following) issue-relevant thought-listing question on levels of self-reported political and economic discussion frequency will be strongest among those with more expansive definitions of said issue.

4. SURVEY CONTEXT EFFECTS AND VALIDITY

Context effects are indicative of potential problems with construct validity. This study examines criterion-oriented validity to examine which measures of political discussion frequency—or more precisely, which survey contexts—produce more valid measurements (Cronbach and Meehl 1955). Criterion validity assesses a measure against some external criterion (Babbie 2005, p. 148): stronger relationships between a construct (i.e., political discussion) and variables to which it should be theoretically related (i.e., criteria) increase the confidence in the validity of that construct's measurement (Treiman 2009). Focusing specifically on concurrent and predictive validity (Cronbach and Meehl 1955), this study examines statistical relationships to assess the validity of construct x's measurement. That is, respondents who score high on the measurement of construct x (in this case, political talk frequency) should also score high on criterion y.

Concurrent validity assesses the strength of the relationship between two related variables, regardless of causal order. For example, media preferences are related to political orientation, although the direction of the causal relationship has not been definitively determined in the selective exposure literature. To examine concurrent validity, analyses in this study focus on the relationship between political and economic discussion frequency (constructs of interest) and education as well as political or economic interest (criterion variables). Previous research has demonstrated that education level is positively related to political talk frequency (see Straits 1991; McLeod, Scheufele, Moy, Horowitz, Holbert, et al. 1999; Bennett, Flickinger, and Rhine 2000; Scheufele 2000) and that higher levels of political interest correlate positively with the frequency of talking about politics (Straits 1991; McLeod, Scheufele, and Moy 1999; Gil de Zuniga, Valenzuela, and Weeks 2016).

Predictive validity focuses more specifically on an exogenous construct's (x's) ability to predict an endogenous criterion (y). For example, in studies documenting a relationship between age cohort and political attitude stability (Sears 1981; Krosnick and Alwin 1989), the former must necessarily cause the latter, whereas the reverse causal direction is not possible. To analyze predictive validity, this study examines the relationship between discussion frequency and current affairs knowledge, as research has consistently shown that political discussion leads to increased knowledge about political affairs (Scheufele 2000; Eveland 2004; Eveland and Hively 2009; Eveland and Schmitt 2015).

Extant research has not examined the validity of political discussion frequency measures, and it is unclear which of the survey context manipulations will produce more valid political talk frequency measurements *a priori*. Thus, the following exploratory research questions ask:

Research Question 1: Does (a) a larger versus smaller battery of issue-relevant questions or (b) a preceding versus following thought-listing question about a political issue (i.e., economy) influence the concurrent validity of political and economic discussion frequency measures?

Research Question 2: Does (a) a larger versus smaller battery of issue-relevant questions or (b) a preceding versus following thought-listing question about a political issue (i.e., economy) influence the predictive validity of political and economic discussion frequency measures?

5. METHOD

5.1 Sample

Data for this study include wave one and wave three of a three-wave panel survey among Dutch adults; questions about discussion frequency were not asked in wave two. Data were collected by I&O Research; respondents in this

company's database were randomly sampled from population registers and were primarily recruited between 2010 and 2013. The time lag between each of the online survey waves was eight weeks, with field dates including February 23, 2015 (wave one), April 20 (wave two), and June 15, 2015 (wave three). For each wave, respondents had twenty-four days to complete the survey.

In wave one, 22,879 people were invited to participate, of which 9,112 started the questionnaire (response rate, RR2 = 39.8 percent) and 6,386 completed the survey (RR1 = 27.9 percent; cooperation rate, COOP1 = 70.1 percent). Only respondents who completed the survey were invited to participate in subsequent waves. In wave two, 4,301 respondents completed the questionnaire (RR1 = 69.0 percent), and 3,270 individuals completed wave three of data collection (RR1 = 77.0 percent). The final sample deviated slightly from the general Dutch population, including an overrepresentation of male (66 percent), older ($M = 61.45$, $SD = 11.08$), and highly educated respondents (50.9 percent obtained a university degree).

5.2 Survey Context Operationalizations (Independent Variables)

5.2.1 Quasi-experiment.   In wave one, ten questions about various politically relevant considerations (e.g., interest, most important societal problem, ideological position, government performance, issue ownership, voting intentions) and twenty-two questions about the economy (e.g., interest, consumer confidence, government performance on the economy, feelings about the economic situation, responsibility for the financial crisis) preceded the discussion frequency questions (which were asked at the end of the survey). In wave three, by contrast, discussion frequency questions were asked in the beginning of the questionnaire, preceded by only a small battery of issue-relevant questions, including one pair of questions about the economy and two political questions.

Comparing differences in self-reported discussion frequency *within individuals* between the two survey waves effectively provides a quasi-experimental "manipulation" of survey contexts with a large, versus small, quantity of preceding issue-relevant questions. Of course, the current quasi-experiment can neither rule out external factors nor offer the same level of confidence about causal factors as true experiments, such that this first study provides preliminary findings and sets the stage for a more methodologically rigorous experiment.

5.2.2 Experiment.   The experiment was embedded in wave three to provide a more internally valid test of survey context priming effects (see Gaines, Kuklinski, and Quirk 2007). Holding all other factors constant (i.e., the questions asked across the two survey forms were identical), a randomized experiment allows the assessment of question order effects (a specific type of survey context effect).

In wave three, an open-ended question asked respondents, "Which thoughts come to your mind when you think of the economy?" This question was placed immediately before ($n = 1,612$) or immediately after ($n = 1,658$) the discussion frequency questions, randomized across participants.[1] Randomization checks confirmed that the two conditions differed neither in the number of words provided in response to the open-ended question ($p = 0.176$) nor on a range of variables assessed in wave one,[2] demonstrating that the experimental groups are not only balanced across this variable but also on a range of demographic and political variables assessed before the thought-listing manipulation. Analyses do not include control variables because of the successful randomization of participants across the two conditions (note: the direction and significance of findings were unchanged when control variables were included in the analyses).

5.3 Measurements

5.3.1 Discussion frequency (dependent variables).   Wave one and wave three included identical questions assessing how frequently respondents discussed each of four topics: "How frequently do you talk about the following topics with others? (a) politics, (b) economy, (c) show business and entertainment, and (d) sports." These four items were presented on a single page, with the order of item presentation randomized across participants to prevent structural assimilation or contrast effect biases for the sample as a whole (see Tourangeau et al. 2000). Response options included eight-point scales (1 = "never," 2 = "rarely," 3 = "once a month," 4 = "once a week," 5 = "twice a week," 6 = "three to six times a week," 7 = "daily," 8 = "several times a day"). Although more stable than findings reported elsewhere (e.g., Morey and Eveland 2016 report r_{xx} over-time reliability independent of stability between 0.43 and 0.50), over-time measures of political and economic discussion frequency show moderate correlations between wave one and wave three ($r = 0.64$ and $r = 0.60$, respectively).

5.3.2 Moderators.   Political and economic interest (also concurrent validity criteria) were assessed (only in wave one) separately using an eleven-point scale ranging from zero (totally not interested) to ten (very much interested) (political: $M = 6.79$, $SD = 1.99$; economic: $M = 7.03$, $SD = 1.69$). These interest variables are particularly well-suited for the study of criterion-related validity because they exhibit remarkable short- and long-term stability. Prior (2010) claims they are "exceptionally stable in the short run and over long periods of time" (p. 747).

1. The survey did not include a comparable open-ended question eliciting participants' thoughts about the concept of "politics" more generally.
2. Randomization checks were performed for age ($p = 0.421$), gender ($p = 0.347$), education ($p = 0.878$), interest ($p = 0.552$), knowledge ($p = 0.709$), news use ($p = 0.463$), and economic ($p = 0.297$) and political discussion frequency ($p = 0.846$).

A conceptual breadth variable (see Fitzgerald 2013) was created by summing the number of words provided in response to the open-ended thought-listing question ($M = 8.33$, $SD = 9.13$), an approach that provides a more objective measurement of this concept than manual coding options.[3] Stopwords—such as "the," "a," "is," "at," or "which," (on average 48 percent of the 16.11 average words written by respondents)—were automatically deleted to avoid measuring verbose language (see Boukes and Trilling 2017). The more substantive words an individual uses to describe what thoughts come to mind when contemplating the economy, the wider that individual's conceptual breadth of this political issue.

To verify the validity of the conceptual breadth operationalization, manual coding was performed on a subset of randomly selected thoughts ($n = 327$; ten percent of sample). Manual coding of the number of individual, economic thoughts each respondent expressed in the thought-listing task demonstrated satisfactory intercoder reliability (Krippendorff's $\alpha = 0.94$, on seventy-five thought-lists). This technique further validates our primary operationalization, with a strong correlation ($r = 0.83$) between the manually coded measure and the automated count measure of conceptual breadth.

5.4 Validity Criterion Measures

5.4.1 Concurrent validity. Measures of political and economic interest are described in the "Moderators" section above. Additionally, respondents were asked to report their highest level of education completed. The scale, adapted to the Dutch educational system, ranged from zero (no education; only primary education) to six (university degree) ($M = 4.02$, $SD = 1.56$).

5.4.2 Predictive validity. Current affairs knowledge was assessed by summing correct responses to eleven multiple-choice questions about political-economic issues ($M = 7.97$, $SD = 2.23$; exact items can be found in the appendix). Focusing on recent events that drew significant media attention, the topics of these questions were likely to be the subject of political discussion among citizens. Knowledge was measured dynamically, with the eleven knowledge items spread across the three survey waves, and new items asked in each wave. The knowledge measure is therefore indicative of respondents' general tendencies to acquire current affairs knowledge, and its correlation with talk frequency should be independent of time. Because our assessment of predictive

3. Manual coding for conceptual breadth suffers from some important limitations. Coding for opinion quality is complex, contentious and difficult to operationalize (including many dimensions, such as extensiveness, level of deliberation, and independence; see Price and Neijens 1997). Coding the number of thoughts expressed is more easily achieved through manual coding, yet identifying the degree of elaborated conceptual breadth is rather subjective and open to diverse interpretation of manual coders (Baek, Cappella, and Bindman 2011).

validity requires one stable criterion (all statistical tests include exactly the same variable), the summed-across-waves knowledge score is more appropriate than knowledge scores within any particular survey wave.

6. RESULTS

6.1 Context Effects: Quasi-Experiment

Results of the quasi-experiment are shown in table 1. Only respondents ($n = 1,612$) who were randomly assigned to receive the preceding (versus following) open-ended thought-listing task in wave three were included in the analyses (note: results did not change when all respondents were included). This approach provides the most conservative test of the hypotheses because all respondents had answered the open-ended question before the talk frequency questions in wave one. Analyses included paired-samples *t* tests to compare within-subject differences on self-reported political discussion frequency. Results (table 1) show that respondents reported higher levels of both political and economic discussion frequency in a survey context that includes a large (wave one), versus small (wave three), battery of preceding issue-relevant questions. Thus, hypothesis 1.A is supported, although it should be noted that effect sizes are rather small (see Cohen's *d*).[4]

   To ensure that results were specific to the political talk frequency items, differences in self-reported entertainment and sports talk frequency were also compared between wave one and wave three (table 1). Results indicate that respondents' self-reported sports talk did not differ significantly across the survey waves ($p = 0.305$). Responses to entertainment discussion frequency did differ across the two waves ($p < 0.001$). However, the association between survey wave and levels of discussion is in the opposite direction relative to the political and economic discussion frequency measurements: a large battery of preceding politically and economically relevant questions was associated with *lower* levels of entertainment discussion frequency (perhaps because the political and economic survey questions produced a more serious state of mind). Overall, these findings indicate that a survey context including a larger number of preceding, politically oriented questions was associated with higher self-reported frequency of discussion, but only about issue-relevant topics (politics and the economy).

   Moderating effects of political and economic interest were tested using regression models estimating discussion frequency (dependent variable) from the survey wave (i.e., survey context), individual-level interest variables, and interaction terms between survey context and interest (controlling for gender, age, education level, and news use). Robust clustered standard errors were

---

4. Ceiling effects might be an explanation for small effects; yet this does not appear to be the case here, as dependent variables were normally distributed, and less than one percent of respondents selected the highest option ("several times a day").

Table 1. Descriptive Statistics and *t* Test Analysis of Talk Variables in Wave One and Wave Three

Talking about:	Wave 1: Preceded by more issue-relevant items	Wave 3: Preceded by less issue-relevant items	Correlation	Difference	*t*	*df*	*p*	Cohen's *d*
Politics	*M* = 3.34 (*SD* = 1.60)	*M* = 3.24 (*SD* = 1.58)	*r* = 0.65	*M* = 0.10 (*SD* = 1.33)	3.03	1,611	0.003	0.08
Economy	*M* = 3.24 (*SD* = 1.53)	*M* = 3.11 (*SD* = 1.52)	*r* = 0.60	*M* = 0.12 (*SD* = 1.37)	3.63	1,611	< 0.001	0.09
Entertainment	*M* = 1.66 (*SD* = 1.43)	*M* = 1.88 (*SD* = 1.38)	*r* = 0.59	*M* = −0.13 (*SD* = 1.27)	−7.13	1,611	< 0.001	0.17
Sports	*M* = 2.70 (*SD* = 1.90)	*M* = 2.73 (*SD* = 1.80)	*r* = 0.78	*M* = −0.03 (*SD* = 1.24)	−1.03	1,611	0.305	0.02

used to correct for repeated measurements of individual respondents (i.e., within-person variation). Results indicate a significant interaction effect between survey wave and individual-level interest. In the model predicting political discussion frequency ($R^2 = 0.28$), the interaction between wave and political interest is significant, $b = -0.06$, $SE = 0.02$, $\beta = -0.04$, $p < 0.001$. Similarly, in the model predicting economic discussion frequency ($R^2 = 0.23$), the interaction between wave and economic interest is significant, $b = -0.10$, $SE = 0.02$, $\beta = -0.05$, $p < 0.001$. Thus, results demonstrate that the relationship between survey context (large versus small battery of preceding issue-relevant questions) and self-reported discussion frequency varied across levels of political and economic interest. This supports Hypothesis 2.A.

More concretely, the positive relationship between a large (versus small) battery of preceding (i.e., assessed before discussion frequency measures) issue-relevant questions and discussion frequency is more pronounced and statistically significant among individuals high in political interest, as shown in figure 1 (top). Confidence intervals show that politically interested respondents reported more frequent political discussion when discussion frequency questions were asked after a large (wave one) versus small (wave three) battery of issue-relevant questions. For the less politically interested, no difference in self-reported general political discussion frequency between survey waves emerged.

A similar pattern is revealed for economic interest and economic discussion frequency, as shown in figure 1 (bottom). Among the less economically interested, there was no relationship between survey context and self-reported economic discussion frequency, but among the more economically interested, a survey context that included a large (wave one) versus small (wave three) battery of preceding issue-relevant questions was associated with higher levels of self-reported economic discussion frequency.

Alternative analysis techniques confirm the robustness of these findings, including one-way repeated measures analysis of variance with covariates (ANCOVA) using a median split of political/economic interest (political discussion frequency: $F\,[1, 1,594] = 4.09$, $p = 0.043$; economic discussion frequency: $F\,[1, 1,594] = 3.87$, $p = 0.049$) and OLS regression examining the effect of political and economic interest on the change in political or economic discussion frequency (respectively) between waves (political discussion: $b = 0.04$, $SE = 0.02$, $\beta = 0.07$, $p < 0.001$; economic discussion: $b = 0.07$, $SE = 0.02$, $\beta = 0.09$, $p < 0.001$).

6.2 Context Effects: Experiment

Embedded in wave three of the panel survey, the experiment examined whether question order (preceding versus following issue-relevant thought-listing question) influenced levels of self-reported political and economic discussion frequency. The main effect of this question order manipulation was not significant (Hypothesis 1.B not supported). Independent sample *t* tests showed no differences in levels of

Frequency of Political Talk
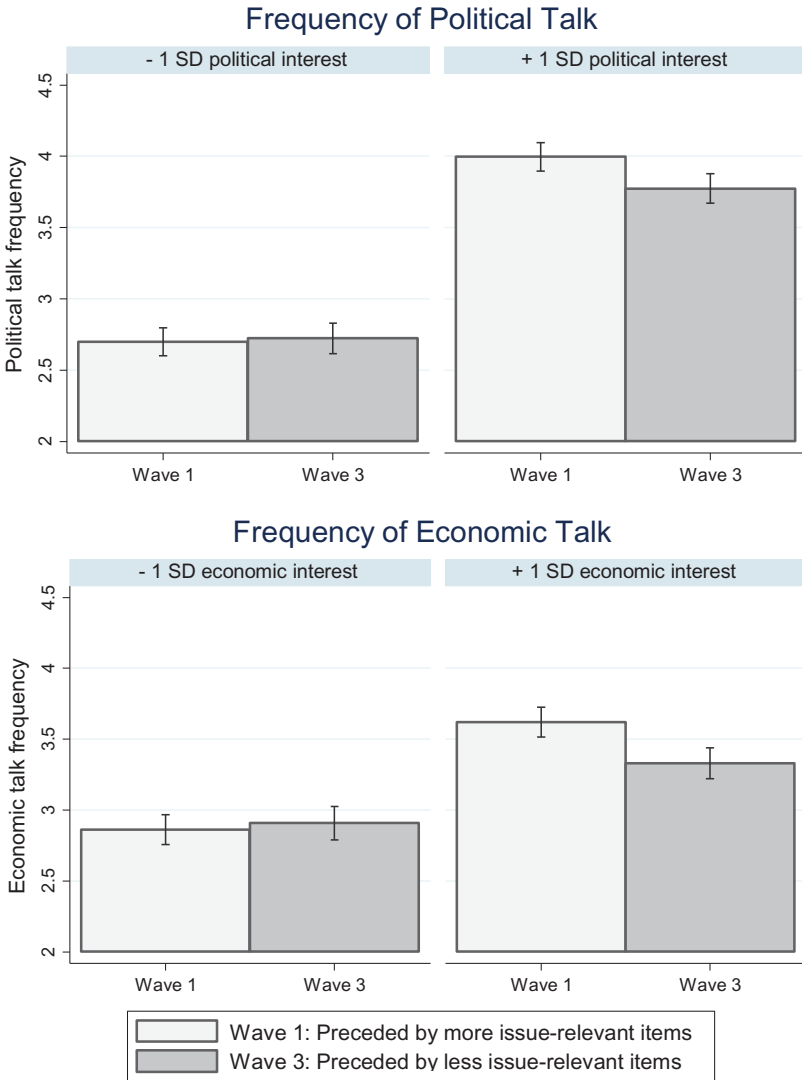

Frequency of Economic Talk

Figure 1. Regression Estimates with 95 Percent Confidence Intervals of Wave One and Wave Three Self-reported Political (Upper Graph) and Economic (Lower Graph) Discussion Frequency for Those More Interested (One Standard Deviation Above the Mean; Graphs on the Right) and Less Interested (One Standard Deviation Below the Mean; Left) in these Topics.

political or economic discussion frequency when the thought-listing question was asked before (politics: $M = 3.24$, $SD = 1.57$; economy: $M = 3.12$, $SD = 1.52$) or after (politics: $M = 3.24$, $SD = 1.57$; economy: $M = 3.11$, $SD = 1.54$) the discussion frequency items (politics: $p = 0.998$; economy: $p = 0.854$).

The moderating effect of conceptual breadth was examined using OLS regression. Conceptual breadth exhibited a significant main effect on both political ($\beta = 0.13$, $p < 0.001$) and economic discussion frequency ($\beta = 0.14$, $p < 0.001$). More pertinent to the hypotheses, the interaction effect between question order and conceptual breadth was *not* statistically significant for political discussion frequency, $b = 0.01$, $SE = 0.01$, $\beta = 0.04$, $p = 0.190$, but results did reveal a statistically significant, positive interaction effect between the experimental manipulation and economic conceptual breadth for economic discussion frequency, $b = 0.02$, $SE = 0.01$, $\beta = 0.10$, $p < 0.001$. This latter interaction effect proved robust, manifesting when control variables and even the lagged wave one dependent variable were included. Findings thus reveal that individual-level economic conceptual breadth moderated the order effect of an economically oriented thought-listing question on economic discussion frequency, but this interaction effect did not appear for general political talk frequency (Hypothesis 2.B partially supported).

The Johnson-Neyman technique was used to compute the values on the moderator (conceptual breadth) where the effect of the independent variable (question order) transitioned from insignificance to significance, and vice versa (see Hayes 2013). Findings showed that the positive effect of a preceding thought-listing question on self-reported economic discussion frequency occurred only among respondents with relatively more expansive conceptualizations of the economy. The survey context effect was significant with conceptual breadth scores of thirteen words or more and increased in strength with more expansive conceptualizations. Figure 2 graphically depicts the results of the regression model, illustrating the effect of placing the open-ended question before or after the discussion frequency questions (represented by the continuous line) for people who hold narrower (*x*-axis, left side) or more expansive (*x*-axis, right side) definitions of the economy. These results demonstrate that the survey context effect is more positive and increases in strength among those with more expansive conceptualizations of the economy.

For comparison purposes, the same analysis was run using self-reported entertainment and sports talk frequency (dependent variables). These topics are not directly related to the thought-listing task and should therefore not produce the significant interaction effect revealed for economic talk frequency. Results demonstrating insignificant interaction effects (both $p > 0.05$) support this conclusion.

6.3 Validity Assessments

Effects of survey context on self-reported political discussion frequency may indicate problems with construct validity. By comparing the strength of relationships
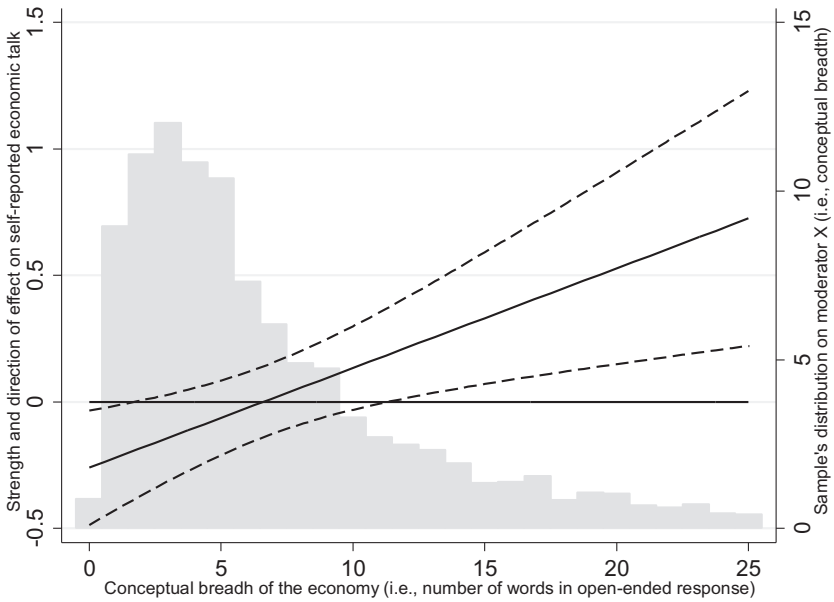
Figure 2. Marginal Effect of an Open-Ended Economic Thought-Listing Question Preceding (Versus Following) an Economic Discussion Frequency Measure, Conditional on an Individual's Conceptual Breadth of the Economy (*x*-axis). Grey background bars show the sample's distribution on the moderator (i.e., conceptual breadth). A significant effect of the open-ended question manipulation on self-reported economic discussion frequency at a particular value *x* (conceptual breadth of the economy) is evidenced when the upper and the lower boundary of the confidence interval (CI) are either both below or above zero on the *y*-axis (the effect on economic discussion frequency) at that particular point on the *x*-axis.

between political talk frequency and known criteria, this study examines the effects of survey context on concurrent validity (education, political or economic interest) and predictive validity (current affairs knowledge) of political discussion frequency items. Demonstrating that the selected variables are suitable for testing criterion validity, correlations between self-reported discussion frequency and education, interest (political or economic), and current affairs knowledge are consistently positive and statistically significant (tables 2a and 2b).

6.3.1 Concurrent validity. Comparing relationships between discussion frequency and education and between discussion frequency and interest, results (table 2a) show significantly stronger correlations in wave one than in wave three (research question 1.A). The difference between correlations for both education and interest are significant (marginally significant for economic discussion frequency and education), as demonstrated by Steiger's application of Fisher's Z-transformation, which allows for the formal comparison of

correlated correlations (Meng, Rosenthal, and Rubin 1992). A larger battery of preceding issue-relevant questions seems to improve concurrent validity of political discussion frequency measures.

In contrast, comparing correlations (discussion frequency and education, discussion frequency and interest) between the two conditions in the wave three survey experiment (table 2b) reveals no differences in the strength of relationships (research question 1.B). Using Fisher's *r*-to-Z-transformation (Kenny 1987) for testing the statistical difference between two independent correlations showed that the relationships with neither education nor interest were stronger in one experimental condition than in the other condition. Placement of an issue-relevant thought-listing question does not seem to affect concurrent validity.

6.3.2 Predictive validity. Similar to tests of concurrent validity, correlations between discussion frequency and current affairs knowledge show a similar pattern. Comparing correlations between wave one and wave three (table 2a), the relationship between economic discussion frequency and the current affairs knowledge scale (0–11) is stronger in the wave with a larger battery of preceding questions (wave one) than in wave three ($p < 0.001$). These results are robust, also manifesting in multivariate models including control variables. The comparison of correlations for general political discussion frequency between waves pointed in the same direction (research question 2.A), but was not statically significant ($p = 0.139$). Correlations between discussion frequency and knowledge were not significantly different in the two conditions of the wave three survey experiment (table 2b; $p = 0.697$ and $p = 0.416$, respectively).

6.3.3 Overall criterion validity. With stronger correlations in wave one than wave three, results suggest that political and economic talk frequency items demonstrate better concurrent and predictive validity in a survey context, including a large number of preceding issue-relevant questions compared to a context that includes only a handful of such questions (although predictive validity results for general political talk frequency are not definitive). In contrast, neither concurrent nor predictive validity seemed to vary according to the placement of a preceding, versus following, thought-listing task.

7. DISCUSSION

7.1 Survey Context Effects: Priming and Conceptualizations of Politics

In contrast to other difficult-to-measure concepts that are critical to the study of political communication—such as turnout rates or media exposure (e.g., Prior 2009)—extensive exploration of political talk survey measures has yet to be undertaken. This study examined survey context effects on political

Table 2a. Tests of Concurrent and Predictive Validity Using Correlation Analysis for Quasi-Experimental Results (Wave One and Wave Three)

Correlation with:	Political discussion frequency			Economic discussion frequency		
	Wave 1: more preceding issue-relevant items	**Wave 3:** less preceding issue-relevant items	**Difference:** $r_{w1} - r_{w3}$	**Wave 1:** more preceding issue-relevant items	**Wave 3:** less preceding issue-relevant items	**Difference:** $r_{w1} - r_{w3}$
	r	r	Steiger's z	r	r	Steiger's z
Concurrent validity						
Education	0.17 ($p < 0.001$)	0.13 ($p < 0.001$)	2.52 ($p = 0.012$)	0.13 ($p < 0.001$)	0.10 ($p < 0.001$)	3.22 ($p = 0.058$)
Interest	0.50 ($p < 0.001$)	0.43 ($p < 0.001$)	5.50 ($p < 0.001$)	0.42 ($p < 0.001$)	0.32 ($p < 0.001$)	6.93 ($p < 0.001$)
Predictive validity						
Knowledge	0.28 ($p < 0.001$)	0.26 ($p < 0.001$)	1.48 ($p = 0.139$)	0.28 ($p < 0.001$)	0.21 ($p < 0.001$)	4.25 ($p < 0.001$)

Table 2b. Tests of Concurrent and Predictive Validity Using Correlation Analysis for Experimental Results (Wave Three Only)

Correlation with:	Political discussion frequency			Economic discussion frequency		
	Condition 1: Open-ended thought-listing question first	**Condition 2:** Open-ended thought-listing question later	**Difference:** $r_1 - r_2$	**Condition 1:** Open-ended thought-listing question first	**Condition 2:** Open-ended thought-listing question later	**Difference:** $r_1 - r_2$
	r	r	Fisher's z	r	r	Fisher's z
Concurrent validity						
Education	0.15 ($p < 0.001$)	0.10 ($p < 0.001$)	1.44 ($p = 0.149$)	0.12 ($p < 0.001$)	0.09 ($p < 0.001$)	0.87 ($p = 0.383$)
Interest	0.44 ($p < 0.001$)	0.42 ($p < 0.001$)	0.59 ($p = 0.556$)	0.31 ($p < 0.001$)	0.34 ($p < 0.001$)	0.93 ($p = 0.350$)
Predictive validity						
Knowledge	0.27 ($p < 0.001$)	0.25 ($p < 0.001$)	0.39 ($p = 0.697$)	0.20 ($p < 0.001$)	0.23 ($p < 0.001$)	0.81 ($p = 0.416$)

discussion frequency using a quasi-experimental manipulation of a large (versus small) battery of preceding issue-relevant questions and an experimental manipulation of a preceding or following issue-relevant thought-listing task.

Quasi-experiments are, of course, limited in their ability to determine the causal direction of effects. In this study, answering questions in a previous survey wave may have affected responses or even actual behavior reported in subsequent waves (panel conditioning), and real-world events occurring between survey waves could have influenced actual levels of political discussion frequency. This study included an experiment to rule out such potential spurious or external factors. Although implementing an experiment within a later survey wave may raise concerns about validity (perhaps the less politically interested dropped out of subsequent waves at higher rates than the more politically engaged), key political characteristics were similar across waves. Moreover, all analyses included only those participants who completed all the waves. In other words, panel attrition cannot influence cross-survey wave comparisons.

Results indicate that a large (versus small) battery of preceding issue-relevant questions is associated with higher levels of self-reported political and economic discussion behavior. This relationship was driven primarily by respondents who were relatively more interested in politics or the economy (respectively), whereas the relationship weakened to statistical insignificance among the less interested. A preceding (versus following) issue-relevant thought-listing question generated higher reports of economic talk frequency, but this effect was constrained to those who held more expansive definitions of the economy.

Overall, these results lead to the robust conclusion that exposure to seemingly minor changes in survey context can influence levels of self-reported political and economic discussion frequency among the politically inclined. An optimistic finding for political discussion research is that the size of survey context effects found in this study are fairly small. However, findings that survey context effects vary across groups of individuals suggest that survey design variations may essentially represent different measurements for different groups (i.e., different subgroups of the population may interpret the exact same question differently depending on context), posing a problem for construct validity (Cronbach and Meehl 1955).

This study will hopefully encourage researchers to pursue a range of questions related to the effects of survey context on self-reported political discussion behavior, including which survey contexts elicit priming (and for whom or how strongly priming occurs) and the mechanisms through which priming impacts self-reported political discussion frequency. The effects of both context manipulations examined in this study were contingent on individual difference variables (interest and conceptual breadth). Yet, whereas the number of preceding issue-relevant questions in the quasi-experiment was directly associated with self-reported discussion frequency, the placement of an issue-relevant thought-listing task about the economy in the experiment did *not* elicit

a main effect. Although previous research shows that both close-ended survey questions and open-ended thought-listing tasks elicit priming effects, answering a long battery of questions and completing a specific thought-listing task are categorically distinct tasks. Perhaps a long battery of questions explicitly mentioning politically relevant concepts simply elicits stronger priming effects than an open-ended thought-listing task (Krosnick 1999); after all, thought-listing induced priming effects are arguably more dependent on a respondent's semantic memory structure and, particulary, on the strength of associations between politically relevant concepts. Thought-listing tasks may also impose a higher cognitive burden on respondents than answering a battery of close-ended questions. Thought-listing tasks require three cognitive steps (Popping 2015): (1) interpreting the question, (2) retrieving relevant information, and (3) translating this into a self-written answer. The third step differs from closed-ended questions and is particularly demanding and difficult (Blackwell, Galassi, Galassi, and Watson 1985; Popping 2015).

Thus, it is also possible that the effect of a preceding (versus following) thought-listing manipulation did not occur among those with less expansive conceptualizations because unengaged participants were more likely to satisfice on this question. Importantly, however, Geer (1988) showed that a lack of input in thought-listing tasks is predominantly driven by an actual absence of concrete thoughts on the topic (similar to narrow conceptual breadth) due to disinterest, rather than a lack of ability more generally (e.g., inability to articulate a response). Future research might use construct accessibility measures to directly assess the strength of priming elicited by various survey contexts (see Bargh and Pratto 1986; Shrum and O'Guin 1993) and continue to explore individual difference variables (likely related to the availability of political concepts in semantic memory) that constrain and enhance priming effects.

Survey context priming effects may be driven predominantly by temporarily expanded working memory definitions of politics (such that quick assessments of political talk frequency lead to higher reports via automatic processes) or by more careful memory searches (such that more effortful assessments lead to higher reports via controlled processes). Future research should explore how, exactly, priming influences self-reported political discussion frequency, including which of these mechanisms underlie the survey context effects observed here and whether the effect of priming on self-reported discussion frequency differs across particular types of survey context manipulations.

The results of this study further highlight the extremely limited extant body of knowledge on the structure of politically relevant concepts in the minds of citizens (Fitzgerald 2013; Morey and Eveland 2016). The thought-listing task used in this study influenced self-reported economic talk frequency (among those with expansive definitions of the economy), but not political talk frequency more generally. Future research should explore questions related to individuals' politically relevant semantic networks, such as whether political issues (e.g., the economy) represent subordinate categories within the

superordinate category of politics, whether political issues reside at the same "level" and simply share some conceptual overlap with the concept of politics, or whether the hierarchal arrangement of such concepts is subject to variability within individuals (depending, for instance, on survey context).

Developing more sophisticated measures of conceptual depth or breadth would be indisputably advantageous to such lines of inquiry. Although the fairly simple operationalization of conceptual depth employed in this study (e.g., number of non-stop words) was validated by manually coding a subset of open-ended responses, more sophisticated approaches to conceptual depth should not only be employed in studies replicating the findings reported here but also may ultimately illuminate the structure of political concepts in citizens' minds. Additionally, future research examining political conceptualizations as potential moderators should ideally measure this variable *before* (rather than *within)* any experimental manipulations. Importantly, in this study, conceptual breadth was independent of experimental condition, t (3268) $=$ 1.35, $p = 0.176$, minimizing concerns about potentially confounding effects of this study design limitation.

7.2 Implications for Validity

Criterion-oriented validation procedures in this study compared the strength of relationships between discussion frequency and two concurrent validity criteria (education and interest) and one predictive validity criterion (current affairs knowledge). Although previous work has established a causal relationship between knowledge and political talk (Scheufele 2000; Eveland 2004; Eveland and Hively 2009; Eveland and Schmitt 2015), reverse causality cannot be ruled out here. Future research should expand predictive validity assessments of political talk items to additional variables, including definitively endogenous criteria, such as political participation (McLeod et al. 1999). Talking about politics has been shown to stimulate election turnout likelihood (Knoke 1990; Zuckerman, Dasovic, and Fitzgerald 2007; Rolfe 2012).

Results of criterion-related validity tests from the quasi-experiment provide tentative evidence that survey contexts including a larger battery of preceding issue-relevant questions may elicit more valid political discussion frequency measures than contexts including a shorter battery of such questions. In contrast, validity tests of discussion frequency measures across the two experimental conditions suggest that placement of an open-ended, politically relevant thought-listing task immediately before or after talk frequency items may have no effect on validity.

Although provoking, these results cannot, of course, be taken as definitive evidence about which survey contexts produce the most valid measures of political discussion frequency. As stated previously, future research should explore the mechanisms underlying the priming effects observed here, including

whether survey context priming leads to more accurate and valid recollections of political discussions or simply expands working memory definitions of politics temporarily, such that more conversations are counted as political regardless of the number of political discussions in which a respondent has actually participated. For purposes of ecological validity (i.e., there are few, if any, surveys that open with political talk frequency questions), this study examined the presence of a large, versus small, battery of issue-relevant questions. Future research might test the strongest manipulation (i.e., maximal differences between conditions) by asking discussion frequency questions as the very first—versus very last—questions in the survey, effects of several different placements (e.g., in each quarter of a survey), or with surveys of varying length.

7.3 Lessons for Future Research

Surveys administered and analyzed by political communication scholars frequently include a large battery of politically oriented questions. Although findings from this study suggest that placing discussion frequency questions *after* rather than *before* a long battery of politically oriented survey items may elicit more valid measures of political talk, the most valid method of measuring political discussion frequency will depend on the focus of a particular research project. For instance, researchers interested in narrowly defined definitions of political discussion may be best served by survey formats that begin with questions about discussion frequency, followed by a range of other politically relevant survey items. This method should provide "uncontaminated" reports of discussion behavior by leaving little opportunity for preceding questions to prime subconcepts.

Demonstrating direct and moderated survey contexts effects, results of this study contribute to extant research demonstrating the complex and porous nature of the concept of politics. Researchers might consider including explicit (and perhaps narrowly constrained) definitions of key concepts (e.g., politics or politically relevant issues, such as the economy, immigration, or the environment) or decomposing the behavior into components that are easier to estimate (Schwarz 1999). By limiting concept ambiguity, explicit definitions might reduce or eliminate question-order priming effects, which should enhance not only the internal validity of political talk measures but also a range of politically oriented variables potentially vulnerable to survey context effects (e.g., political interest, see Robison 2015). However, providing detailed definitions is not an uncontested issue, and various concerns (e.g., space, time, and price) might discourage researchers from implementing this strategy.

This study reveals that effects of minor survey context changes representing just the sorts of variations that exist across political science and political communication surveys (e.g., specific placement of political talk items in public opinion surveys) and focuses on individual-level differences representing just

the sorts of characteristics often critical to political communication effects (e.g., political interest, the expansiveness of politically relevant conceptualizations). With main and moderated survey context effects suggesting problems with construct validity, results indicate that survey contexts do indeed influence criterion-related validity of political talk frequency measures. Overall, the impact of different survey context decisions on the measurement validity of self-reported political behaviors suggests that this area of research is deserving of much further attention by political communication scholars.

Appendix A:

Table A1. *Questions used to measure current affairs knowledge.*

Wave	Question	Answer
1	Looking at the current interest rates for saving accounts and mortgages, is the rate higher or lower than normal?	Lower
1	Who is the Dutch Minister of Finance	Jeroen Dijsselbloem
1	Who is the currently the Managing Director of the International Monetary Fund	Christine Lagarde
1	Which of the following five countries does not belong to Netherlands' most important trade partners?	Spain
1	What is the credit rating of the Netherlands according to Fitch and Moody's?	AAA
2	Which government-owned bank came into disrepute due to the bonuses of their directors?	ABN Amro
2	Which law was approved by Parliament that directly influences Dutch employees?	Allowing flexible working times
3	Which semi-public corporation did Timo Huges work for before he resigned after problems with public procurements?	NS Dutch Railways
3	What is the percentage of economic growth predicted by the Dutch National Bank?	2 percent

REFERENCES

Aronson, E., T. D. Wilson, and R. M. Akert (2007), *Social Psychology* (6th ed.), Upper Saddle River, NJ: Prentice Hall.

Babbie, E. (2005), *The Basics of Social Research*, Belmont, CA: Wadsworth.

Baek, Y. M., J. N. Cappella, and A. Bindman (2011), "Automating Content Analysis of Open-Ended Responses: Wordscores and Affective Intonation," *Communication Methods and Measures*, 5, 275–296.

Barabas, J. (2004), "How Deliberation Affects Policy Opinions," *American Political Science Review*, 98, 687–701.

Bargh, J. A., and F. Pratto (1986), "Individual Construct Accessibility and Perceptual Selection," *Journal of Experimental and Social Psychology*, 22, 293–311.

Bennett, S. E., R. S. Flickinger, and S. L. Rhine (2000), "Political Talk over Here, over There, over Time," *British Journal of Political Science*, 30, 99–119.

Bishop, G. F., R. W. Oldendick, and A. J. Tuchfarber (1982), "Political Information Processing: Question Order and Context Effects," *Political Behavior*, 4, 177–200.

Blackwell, R. T., J. P. Galassi, M. D. Galassi, and T. E. Watson (1985), "Are Cognitive Assessment Methods Equal? A Comparison of Think Aloud and Thought Listing," *Cognitive Therapy and Research*, 9, 399–413.

Boukes, M., and D. Trilling (2017), "Political Relevance in the Eye of the Beholder: Determining the Substantiveness of TV Shows and Political Debates with Twitter Data," *First Monday*, 22, available at http://firstmonday.org/ojs/index.php/fm/article/view/7031/6093.

Chambers, S. (2003), "Deliberative Democratic Theory," *Annual Review of Political Science*, 6, 307–326.

Conover, P. J., and D. D. Searing (2005), "Studying "Everyday Political Talk" in the Deliberative System," *Acta Politica*, 40, 269–283.

Conover, P. J., D. D. Searing, and I. M. Crewe (2002), "The Deliberative Potential of Political Discussion," *British Journal of Political Science*, 32, 21–62.

Cronbach, L. J., and P. E. Meehl (1955), "Construct Validity in Psychological Tests," *Psychological Bulletin*, 52, 281–302.

Delli Carpini, M. X., F. L. Cook, and L. R. Jacobs (2004), "Public Deliberation, Discursive Participation, and Civic Engagement: A Review of the Empirical Literature," *Annual Review of Political Science*, 7, 315–344.

De Vreese, C. H., and H. G. Boomgaarden (2006), "Media Message Flows and Interpersonal Communication: The Conditional Nature of Effects on Public Opinion," *Communication Research*, 33, 19–37.

Druckman, J. N., and K. R. Nelson (2003), "Framing and Deliberation: How Citizens' Conversations Limit Elite Influence," *American Journal of Political Science*, 47, 729–745.

Eveland, W. P. Jr. (2004), "The Effect of Political Discussion in Producing Informed Citizens: The Roles of Information, Motivation, and Elaboration," *Political Communication*, 21, 177–193.

Eveland, W. P. Jr., and M. H. Hively (2009), "Political Discussion Frequency, Network Size, and "Heterogeneity" of Discussion as Predictors of Political Knowledge and Participation," *Journal of Communication*, 59, 205–224.

Eveland, W. P. Jr., A. C. Morey, and M. J. Hutchens (2011), "Beyond Deliberation: New Directions for the Study of Informal Political Conversation from a Communication Perspective," *Journal of Communication*, 61, 1082–1103.

Eveland, W. P. Jr., and J. B. Schmitt (2015), "Communication Content and Knowledge Content Matters: Integrating Manipulation and Observation in Studying News and Discussion Learning Effects," *Journal of Communication*, 65, 170–191.

Fitzgerald, J. (2013), "What Does "Political" Mean to You?" *Political Behavior*, 35, 453–479.

Gaines, B. J., J. H. Kuklinski, and P. J. Quirk (2007), "The Logic of the Survey Experiment Reexamined," *Political Analysis*, 15(1), 1–20.

Geer, J. G. (1988), "What Do Open-Ended Questions Measure?" *Public Opinion Quarterly*, 52, 365–367.

Gil de Zuniga, H., S. Valenzuela, and B. E. Weeks (2016), "Motivations for Political Discussion: Antecedents and Consequences on Civic Engagement," *Human Communication Research*, 42, 533–552.

Habermas, J. (1991). *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*, (T. Burger Trans.), Cambridge, MA: MIT press.

Hayes, A. F. (2013), *An Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, New York: Guilford Press.

Huckfeldt, R. (2001), "The Social Communication of Political Expertise," *American Journal of Political Science*, 45, 425–438.

———. (2007), "Unanimity, Discord, and the Communication of Public Opinion," *American Journal of Political Science*, 51, 978–995.

Hyman, H. H., and P. B. Sheatsley (1950), "The Current Status of American Public Opinion," in *The Teaching of Contemporary Affairs*, ed. J. C. Payne, pp. 11–34, Washington: National Council of Social Studies.

Iyengar, S., D. R. Kinder, M. D. Peters, and J. A. Krosnick (1984), "The Evening News and Presidential Evaluations," *Journal of Personality and Social Psychology*, 46, 778–787.

Kenny, D. A. (1987), "Chapter 16: Testing Measures of Association", in *Statistics for the Social and Behavioral Sciences*, ed. D.A. Kenny, pp. 270–291, Boston, MA, Canada: Little, Brown and Company.

Klofstad, C. A., A. E. Sokhey, and S. D. McClurg (2013), "Disagreeing about Disagreement: How Conflict in Social Networks Affects Political Behavior," *American Journal of Political Science*, 57, 120–134.

Knoke, D. (1990), "Networks of Political Action: Toward Theory Construction," *Social Forces*, 68, 1041–1063.

Krosnick, J. A. (1999), "Survey Research," *Annual Review of Psychology*, 50, 537–567.

Krosnick, J. A., and D. F. Alwin (1989), "Aging and Susceptibility to Attitude Change," *Journal of Personality and Social Psychology*, 57, 416.

Krosnick, J. A., and L. A. Brannon (1993), "The Impact of the Gulf War on the Ingredients of Presidential Evaluations: Multidimensional Effects of Political Involvement," *American Political Science Review*, 87, 963–975.

Lee, F. L. F. (2005), "The Impact of Ordinary Political Conversation on Public Opinion Expression: Is Existence of Discord Necessary?" *Journalism and Mass Communication Quarterly*, 82, 891–909.

Lee, F. L. (2012), "Does Discussion with Disagreement Discourage All Types of Political Participation? Survey Evidence from Hong Kong," *Communication Research*, 39, 543–562.

Leighley, J. E. (1990), "Social Interaction and Contextual Influences on Political Participation," *American Politics Quarterly*, 18, 459–475.

Luskin, R. C. (1990), "Explaining Political Sophistication," *Political Behavior*, 12, 331–361.

Macrae, C. N., C. Stangor, and A. B. Milne (1994), "Activation Social Stereotypes: A Functional Analysis," *Journal of Experimental Social Psychology*, 30, 370–389.

McLeod, D. M., G. M. Kosicki, and J. M. McLeod (2009), "Political Communication Effects," in *Media Effects: Advances in Theory and Research* (3rd ed.), eds. J. Bryant and M. B. Oliver, pp. 228–251, New York: Routledge.

McLeod, J. M., D. A. Scheufele, and P. Moy (1999). "Community, Communication, and Participation: The Role of Mass Media and Interpersonal Discussion in Local Political Participation," *Political Communication*, 16, 315–336.

McLeod, J. M., D. A. Scheufele, P. Moy, E. M. Horowitz, R. L. Holbert, W. Zhang, S. Zubric, and J. Zubric (1999), "Understanding Deliberation: The Effects of Discussion Networks on Participation in a Public Forum," *Communication Research*, 26, 743–774.

Meng, X. L., R. Rosenthal, and D. B. Rubin (1992), "Comparing Correlated Correlation Coefficients," *Psychological Bulletin*, 111, 172–175.

Morey, A. C., and W. P. Eveland (2016), "Measures of Political Talk Frequency: Assessing Reliability and Meaning," *Communication Methods and Measures*. DOI: 0.1080/ 19312458.2015.1118448.

Mutz, D. C. (2002), "The Consequences of Cross-Cutting Networks for Political Participation," *American Journal of Political Science*, 46, 838–855.

Popping, R. (2015), "Analyzing Open-Ended Questions By Means of Text Analysis Procedures," *Bulletin of Sociological Methodology*, 128, 23–39.

Price, V., J. N. Cappella, and L. Nir (2002), "Does Disagreement Contribute to More Deliberative Opinion?" *Political Communication*, 19, 95–112.

Price, V., and P. Neijens (1997), "Opinion quality in public opinion research," *International Journal of Public Opinion Research*, 9(4), 336–336.

Prior, M. (2009), "The Immensely Inflated News Audience: Assessing Bias in Self-Reported News Exposure," *Public Opinion Quarterly*, 73, 130–143.

———. (2010), "You've either got it or you don't? The stability of political interest over the life cycle," *The Journal of Politics*, 72(3), 747–766.

Robison, J. (2015), "Gaps in Political Interest: Following Public Affairs in Surveys from Gallup, Pew, and the ANES," *International Journal of Public Opinion Research*, 27, 406–416.

Rolfe, M. (2012), *Voter Turnout: A Social Theory of Political Participation*, Cambridge, UK: Cambridge University Press.

Scheufele, D. A. (2000), "Talk or Conversation? Dimensions of Interpersonal Discussion and Their Implications for Participatory Democracy," *Journalism and Mass Communication Quarterly*, 77, 727–743.

Schmitt-Beck, R., and O. Lup (2013), "Seeking the Soul of Democracy: A Review of Recent Research into Citizens' Political Talk Culture," *Swiss Political Science Review*, 19, 513–538.

Schwarz, N. (1999), "Self-Reports: How the Questions Shape the Answers," *American Psychologist*, 54, 93–105.

Sears, D. O. (1981), "Life-Stage Effects on Attitude Change, Especially among the Elderly," in *Aging: Social Change*, eds. S. B. Kiesler, J. N. Morgan, and V. K. Oppenheimer, pp. 183–204, New York: Academic Press.

Shrum, L. J., and T. C. O'Guin (1993), "Processes and Effects in the Construction of Social Reality: Construct Accessibility as an Explanatory Variable," *Communication Research*, 20, 436–471.

Sokhey, A. E., and P. A. Djupe (2014), "Name Generation in Interpersonal Political Network Data: Results from a Series of Experiments," *Social Networks*, 36, 147–161.

Straits, B. C. (1991), "Bringing Strong Ties Back in: Interpersonal Gateways to Political Information and Influence," *Public Opinion Quarterly*, 55, 432–448.

Tourangeau, R., L. J. Rips, and K. Rasinski (2000), *The Psychology of Survey Response*, Cambridge, UK: Cambridge University Press.

Treiman, D. J. (2009). *Quantitative Data Analysis: Doing Social Research to Test Ideas*, San Francisco: John Wiley & Sons.

Valenzuela, S., Y. Kim, and H. Gil de Zuniga (2012), "Social Networks That Matter: Exploring the Role of Political Discussion for Online Political Participation," *International Journal of Public Opinion Research*, 24, 163–184.

Zuckerman, A., J. Dasovic, and J. Fitzgerald (2007). *Partisan Families: The Social Logic of Bounded Partisanship in Germany and Britain*, Cambridge, UK: Cambridge University Press.