When Survey Respondents Cheat: 
Internet Exposure and Ideological Consistency in the United States

BETHANY P. BRYSON¹
James Madison University, USA

Increasingly biased Internet news and information is frequently cited as a cause of opinion polarization in the United States. But is it that easy for media messages to influence political opinion? Matched samples of face-to-face versus online respondents in the 2012 and 2016 American National Election Studies reveal that about 23% of online respondents likely "cheated" by referencing the Internet to inform their answers. Doing so allowed those participants to provide more ideologically consistent responses to 41 survey questions, creating a strikingly bimodal distribution of reported opinion by pulling moderate answers to the political right. Quantile regression confirms these results. Probable cheating also increased the effect of Internet news source bias. These findings suggest that in-the-moment Internet messages can influence reported opinions, not because Internet media consumers are duped, but because online information empowers them to give answers consistent with dominant political schemata in survey options and online information.

Keywords: media influence, polarization, political knowledge, media bias, online surveys

If you spent two minutes researching the U.S. budget for foreign aid, could that experience influence your opinion on U.S. foreign policy? Social theorists have long worried about whether cultural products can control us by shaping our beliefs and worldviews, whether culture can obscure an oppressive economic system (Marx & Engels, 1972), or whether radio and television could conduct centrally controlled mass deception (Horkheimer & Adorno, 2002).

Recent concerns about political beliefs, however, have focused less on single messages than on polarized extremism in the U.S. (Baldassarri & Goldberg, 2014; Morris & Morris, 2017; Mutz & Young, 2011; Pacewicz, 2015) and globally (Duyvendak, Geschiere, & Tonkens, 2016; Lamont et al., 2016; Mij, Bakhtiari, & Lamont, 2016), especially relative to Brexit (Bastos & Mercea, 2019) and political instability in the Middle East (Eltantawy & Wiest, 2011). Media effects are also changing as rapidly as technologies of mass communication (DiMaggio, Hargittai, Neuman, & Robinson, 2001; Hampton, 2017), creating gaps in our knowledge about the effects of political mass messaging that urgently need to be filled. For example,

Bethany P. Bryson: brysonbp@jmu.edu
Date submitted: 2019-09-17

¹ I owe thanks to Gabriel Giacalone, Shaneda Destine, Benjamin Brewer, David Truille, the American National Election Studies, and the 2012–2016 respondents who made the effort to research those survey questions.

Copyright © 2020 (Bethany P. Bryson). Licensed under the Creative Commons Attribution Non-commercial No Derivatives (by-nc-nd). Available at http://ijoc.org.
evidence that Russia used social media to produce and disseminate political polarization among American voters during the 2016 U.S. presidential election changed the calculus of mass deception from that of single-message propaganda to a set of conflicting ideas designed to undermine social cohesion. Moreover, that strategy quickly and easily reached more than 126 million users of Facebook alone (Weise, 2017)—a mass-media sort of accomplishment. Disinformation strategies drawing on social media have also been used by Iran and Venezuela to influence politics across the globe (Frenkel, Conger, & Roose, 2019), and at least 51 nations have proposed or implemented action against Internet misinformation (Funke & Flamini, 2019).

But could it be that simple? Could exposure to the Internet generate political polarization over the course of a few months—or even instantaneously? Would it be possible to broadcast conflicting messages and expect both to spread within the same social system? In this analysis, I take advantage of a rare opportunity to test hypotheses about the effect of Internet exposure on political polarization in a large national survey. The 2012 and 2016 waves of the American National Election Studies (ANES) Time-Series Surveys each added an online sample to their traditional face-to-face methods, and I use propensity score matching to achieve a quasi-experimental design in which taking the survey online is the “treatment.” In each year, questionnaires for the two modes were identical, and ANES provided connected laptops to online respondents who lacked Internet access. An unspoken side effect of taking the survey online, however, was that, unlike those in the face-to-face sample, online respondents could consult the Internet while answering the questions. Using evidence from an included vocabulary quiz and eight questions about political knowledge, I show that hundreds of them did so, making basic political knowledge accessible to a wider range of respondents.

Knowledge, per se, was not the only phenomenon influenced, however. Using a measure of liberal-versus-conservative ideological consistency across 41 opinion questions in nine topic areas, I found that respondents who likely consulted the Internet for political knowledge provided more ideologically consistent answers, creating polarization. They also favored the conservative end of the spectrum (like mass deception). I employ quantile regression techniques to analyze both polarization and left-versus-right directional shifts, allowing me to reconcile these two explanations. In so doing, I contribute to the nascent literature on survey cheating by demonstrating that artificially inflated knowledge scores also influenced reported opinion. Finally, I confirm this result by demonstrating that the effect of left-versus-right-biased Internet news consumption is significantly stronger among respondents who likely accessed the Internet during their survey. Although the statistical tests reported here only make inference to a larger theoretical population of survey respondents, results of this quasi-experiment support further research into the speed with which Internet exposure might influence opinion formation and expression in other contexts.

**Background**

*Survey Cheating*

I use the term *cheating* in reference to respondents who likely accessed the Internet during their survey—not because consulting the Internet is immoral, but because it is the accepted term in this literature and a good metaphor for getting assistance from outside sources during a quiz. Moreover, seeking Internet information when asked to engage in any kind of decision-making process is fast becoming a normal behavior among a privileged class of people who have ample Internet access and good skills for using it (Hargittai,
Just as in exam cheating, accessing the Internet during an opinion survey has the potential to convey advantage—in this case, political advantage. As with academic exams, the answers provided by cheaters are incompatible with those provided by noncheaters. This study leverages that difference—perhaps making the most valuable contribution one can glean from these ultimately incompatible modes of data collection.

We already know that, given the opportunity, some online respondents will cheat on knowledge surveys of all sorts (Burnett, 2016), and even more will do so with a little motivation. Jensen and Thomsen (2014), for example, found that more than 22% of respondents to an online political knowledge survey self-reported using the Internet to research answers, and Motta, Callaghan, and Smith (2016) found comparable results across six different surveys. Prior and Lupia (2008) found that extra time and small financial incentives increased political knowledge scores by 24%. However, it is not clear that cheating takes more time than deciding on an answer; Jensen and Thomsen (2014) found that cheating reduced response times.

We know even less about the focus of the current study: how probable cheating might influence reported opinions. In particular, our current understanding of the way political knowledge increases opinion polarization (e.g., Gamson & Modigliani, 1966; Mutz & Young, 2011) has been grounded in methods that use face-to-face surveys to assess the political information stored in respondents’ memories.

Online surveys, however, have become more accepted among researchers and now provide important public opinion data even though a sizeable portion of those respondents cheat (Motta et al., 2016). Moreover, when opinion polls are considered a form of political voice or an identity statement, answers matter to respondents (Baldassarri & Goldberg, 2014; Mason, 2015). Thus, when opinion polls invoke a person’s identity, values, or self-interest, respondents are motivated to seek information that is congruent with those concerns (Lavine, Borgida, & Sullivan, 2000). This finding suggests both motivations for survey cheating and a reason to hypothesize that cheaters might give more ideologically consistent responses. Those whose concerns have been triggered by a survey question may consult the Internet, not so that they can give biased answers, but so they can give answers that better reflect the views they want to convey. What we don’t know is whether seeking political information online can influence responses in the-moment. That is the overarching question this article addresses. I will begin, however, by establishing that survey cheating does indeed occur in the ANES 2012 and 2016 online surveys:

**H1:** Survey Cheating. Respondents who took the ANES survey online will have higher political knowledge scores than respondents who took the face-to-face survey.

**A Hegemonic Internet**

What happens when people seek political information online? Classical theories about the influence of major cultural producers posit conservative effects that imbue viewers with hegemonic ideologies sympathetic to the interests of the ruling class (e.g., Gramsci, 1992), and the Internet is no exception. The Internet was once touted for its democratic, even anarchic, potential and nonhierarchical structure;

---

2 Although having consistent views could create a motivation to cheat rather than the other way around, this study controls that effect by comparing online respondents with those in the face-to-face sample who did not have the opportunity to access the Internet during their surveys, regardless of ideology.
however, critical media scholars have demonstrated that today’s Internet works, at least partly, in service to the capitalist interests that pay for its operation (Mager, 2014). Computing arrangements are often conservative forces deployed in the service of powerful people and institutions while reproducing the social arrangements that brought them into being (Cammaerts & Mansell, 2020; Pedro, 2011). Moreover, nations and political interests have regulated Internet information to further their own interests, be that repressing opposition in Syria (Matar, 2019) and China (Lu & Zhao, 2018), reducing unrest in Greece and the UK (Pedro, 2011), or endorsing political candidates in the U.S. (Serazio, 2015). These conditions might lead us to expect that exposure to political information on the Internet would result in conservative pressure on reported opinion.

H2: **A Conservative Internet. Those who likely accessed the Internet during their survey will report more consistently conservative opinions than other respondents.**

**Polarized Survey Responses**

In contrast to theories about dominant ideologies, political polarization has been a long-standing concern across the globe: in Africa (e.g., Manning, 2005), the Middle East (e.g., Hilal, 2010), Asia (e.g., Yee, 2017), South America (e.g., Bornschier, 2019), and Europe (e.g., Knutsen, 1988). In recent years, however, opinion researchers have also demonstrated polarization in the U.S. on matters involving climate change (McCright & Dunlap, 2011), abortion (Davis & Robinson, 1996; DiMaggio, Evans, & Bryson, 1996; Evans, Bryson, & DiMaggio, 2001), other topics framed as moral (Baldassarri & Gelman, 2008; Davis & Robinson, 1996; Evans, 2003), and take-off issues such as health care (Baldassarri & Bearman, 2007).

In addition, many of those polarizing forces have involved shifts to the right end of the political spectrum. Conservative Protestants, for example, have become more extreme on average, but other religious groups have not made similar moves in the opposite direction (Brooks & Manza, 2004; Evans, 2003). Thus, even when polarization has been demonstrated, that trend may be caused by a pull in a single direction (Bryson, 2020).

In this article, I focus on ideological consistency across a set of opinion questions—a type of polarization also called constraint (Converse, 1964; DiMaggio et al., 1996). Baldassarri and Gelman (2008), for example, found this form of polarization among privileged respondents and on moral issues, while the Pew Research Center (2014) and Layman, Carsey, and Horowitz (2006) have found ideological consistency in the U.S. to be on the rise overall.

Although political elites and media outlets in the U.S. have become more partisan over time (Mansbridge & Martin, 2013), and although those shifts largely preceded similar shifts toward polarized popular opinion, causal links between the media and opinion polarization have been elusive (see Prior, 2013). Nevertheless, given the increasingly polarized landscape of political ideas promulgated on the Internet, this study asks whether likely Internet exposure generates in-the-moment polarization in the answers provided to a range of opinion questions.
H3: A Polarizing Internet. Those who likely accessed the Internet during their survey will evidence more ideological consistency than other respondents, on both ends of the political spectrum.

**Internet News Source Bias**

The online sources of extreme and counterhegemonic ideas seem limited only by the imagination of content creators, with examples from the left such as #MeToo, #BlackLivesMatter, and the Arab Spring (Clark, 2016; Eltantawy & Wiest, 2011; Pollack, Allern, Kantola, & Ørsten, 2018) countered by others on the right, such as hate speech (Ben-David & Matamoros Fernández, 2016), right-wing conspiracy theories (Tripodi, 2019), and a posttruth era (Benkler, Faris, Roberts, & Zuckerman, 2017; Yee, 2017). Thus, the forces of online polarization appear to be gaining both strength and popularity. It is not surprising, then, that politically motivated information seeking not only increased during the 2004 U.S. presidential campaign (Stroud, 2008), but also led to opinion polarization (Stroud, 2010). Internet tracking data, moreover, show that one in four Americans visited a fake news site at least once during the 2016 presidential campaign (Guess, Nyhan, & Reifler, 2018).

In short, finding accurate political information is not easy, and that makes understanding news source bias especially important for explaining political polarization. Seeking political information was the most difficult of five Internet search tasks tested by Hargittai (2002), and users vary widely in their ability to critically evaluate the material they encounter in the process (Hargittai, Fullerton, Menchen-Trevino, & Thomas, 2010). In addition, media consumers do not appear to distinguish the credibility of campaign messages delivered through newspapers from those delivered via Twitter (Morris, 2018)—a platform that is especially susceptible to polarized patterns of content sharing (Stewart, Arif, & Starbird, 2018). The Internet now plays a significant role in disseminating campaign information and opening new pathways of participation (Hargittai & Shaw, 2013), but despite apparently open access, disparities in access and skill continue to make the Internet a place where advantage reproduces itself (Zillien & Hargittai, 2009).

By 2010, of course, the information economy made unintentional exposure to advertising and political messages the business model for social media platforms, which developed algorithms designed to hold the attention of users beyond their intended time and purpose (Charlesworth, 2014). Thus, ANES respondents seeking the answers to factual questions (such as the length of one Senate term) would almost certainly encounter other political ideas (such as term limits) that could influence their answers to related opinion questions. Moreover, online search results are likely channeled through algorithms that draw on one’s previous search history and browsing practices (Noble, 2018). That effect is expected to be strongest among likely cheaters, regardless of whether their biases lean to the left or to the right:

H4: Internet News Source Bias. Internet news bias, whether leaning toward the left or the right, will have a stronger influence on reported opinions among likely cheaters versus other respondents.
Methods

Data

In total, 5,470 respondents completed hundreds of political opinion questions in the preelection and postelection waves of the 2012 or 2016 ANES and answered all the questions used in this study, including all 41 variables in the ideological-consistency scale and eight questions used as controls. In each year, the online and face-to-face surveys were identical to each other, and, if necessary, participants in the Web samples were provided laptops and Internet access for the duration of the multisurvey process. “Don’t know” options were not provided to either sample; however, my analysis takes advantage of the fact that not knowing an answer affected the two samples differently. Face-to-face respondents were prompted by trained interviewers to answer anyway (except for knowledge questions described next). If that person still failed to provide an answer, the missing data would be coded as “don’t know” or “refused.” Despite greater pressure to choose an answer, people taking the face-to-face survey were more likely to skip questions than those in the online sample, who could skip or research their answers online.

Although the ANES uses probability sampling, this is not a formal experiment in that cases were not randomly assigned to test and control groups (face-to-face versus online delivery). In 2012, online respondents were selected via random-digit dialing and address-based random sampling, while the 2016 online mode exclusively used a random sample of U.S. residential addresses. In both years, face-to-face respondents were drawn from a stratified, multistage cluster sample of addresses. Response rates in 2016 were 50% in the face-to-face sample and 44% online. Probability-based response rates are not available for 2012, but the ANES promises that a future release of the codebook will report those calculations (American National Election Studies, 2012, 2016).

To reduce the influence of potential differences between online and face-to-face samples, the R package MatchIt (Ho, Imai, King, & Stuart, 2011) was used to match cases in the face-to-face sample with those in the “treated” online sample. Matching was conducted with replacement using logit-based nearest-neighbor propensity scores—the probability of taking the survey online, given age, income, average Internet news source bias, vocabulary quiz score and dummy variables for college degree, home Internet connection, political activity, and being White, Latinx, or male. The nearest-neighbor method, however, maximizes overall balance across the full set of specified variables. Therefore, I required exact matches on three variables of special importance: college education, political activity, and vocabulary score;3 cases were otherwise matched using the nearest-neighbor method. The resulting matched data served as the basis for all figures and analyses, reducing model assumptions, and approximating an experimental design.

As recommended by Lenis, Nguyen, Dong, and Stuart (2019), the analyses also employ survey weights provided by the ANES, especially to account for nonresponse bias. The resulting data’s t tests also show no significant differences (p > 0.05) between the two groups on three variables of interest that are

---

3 Vocabulary scores were included in the matching procedure to approximate biases related to political knowledge without holding political knowledge constant between the samples, which would make comparisons impossible.
Independent Variables

Political Knowledge. To identify likely cheaters, I used a scale of eight political knowledge questions that asked respondents to identify Joe Biden, Paul Ryan, and John Roberts, the length of a Senate term, House and Senate majorities, whether and which party is more conservative, and which category of federal spending is the smallest from among the following choices: Medicare, Social Security, national defense, and foreign aid. Four questions provided the options of Democrats or Republicans, while the Senate term and the three political-figure questions were open-ended to eliminate guessing. Face-to-face interviewers prompted respondents who said they did not recognize a political figure to reply anyway. Separate variables flag those events, and I used them to code resulting answers as missing/incorrect, given that there was no prompt for online respondents. There were no prompts for the questions about Senate terms, congressional majorities or federal spending. Comparable proportions of each sample had missing answers for the knowledge questions (4.98% of the face-to-face sample and 4.59% of the online sample). The resulting scale is the sum of correct answers, with “don’t know” responses coded as 0. It ranges from 0 to 8, but most analyses employ a flag for scoring more than one standard deviation above the total sample mean—at least seven of eight questions correct, which is above the 96th percentile in the face-to-face sample.⁴

In 2016, the answers to these questions were: (1) U.S. Vice President, (2) Speaker of the House, (3) Chief Justice of the U.S. Supreme Court, (4) six years, (5) Republicans, (6) Republicans, (7) Republican, (8) foreign aid.

⁴ Motta and colleagues (2016) also briefly examined the 2012 ANES knowledge questions and found correct scores to be significantly higher among online respondents.
Although this is an imperfect measure of an elusive phenomenon, the 23% of online respondents coded as likely cheaters closely matches the rate of survey cheating found in other studies (Jensen & Thomsen, 2014; Motta et al., 2016; Prior & Lupia, 2008). Other measures used in survey cheating research, such as self-reports, browser histories, and impossibly difficult questions, are not available in the ANES. There are other considerations, which vary according to question. These include the difficulty of finding information on each topic and how likely that information is to be biased by search algorithms. Because the goal is to measure the general propensity to consult the Internet for political information, a scale-based measure is best suited to the task.

Likely cheaters are online respondents with improbably high political knowledge—a dummy that flags scores of 7 or 8 on the political knowledge quiz within the Web sample. This is an interaction term that multiplies survey mode by the quiz score dummy. It is used with controls for the main effects of taking the survey online and for earning a high knowledge score.

Vocabulary Quiz. The vocabulary quiz is a proprietary product employed by several national surveys. The purpose is to establish basic English comprehension and verbal sophistication. The specific words used are unpublished, but they were chosen to produce a normal distribution of scores in the target population. I use it to demonstrate the extent of probable cheating on knowledge questions among online respondents. In the right panel of Figure 1, the normal distribution of vocabulary scores is altered by matching on the other variables. In all other analyses, vocabulary scores are matched exactly between survey modes—that is, there are no differences in vocabulary scores by survey mode.

Internet News Bias. The ANES asks respondents about their media consumption habits, including whether they visited about 20 specific Internet news sites. Because of the dynamic nature of the media landscape, these lists change in each survey year, so the overlapping options from 2012 and 2016 were matched with available news bias ratings. Although two of these were based on scholarly analysis, none was comprehensive or recent enough to make good use of the ANES options. Therefore, this measure employs a crowd-sourced bias score available from allsides.com. Those rankings range from 1 (strong liberal bias) to 5 (strong conservative bias) and appear to have a somewhat conservative slant. That is, right-leaning news sites had more centrist ratings than expected. Nevertheless, these scores were consistent with other sources, in relative terms. The measure used here is an average of bias scores for sites visited: ABC (2), CNN (3), Fox (4), Huffington Post (1), NBC (2), The New York Times (2), USA Today (3), The Washington Post (2), and Yahoo (2).

Analytic Strategy

After demonstrating the existence of cheaters, I present density distributions on the ideological consistency measure by survey mode and political knowledge score. In these analyses, I use survey mode as a control variable to determine whether political knowledge gained online has a different influence on polarization than retained knowledge that is available for recall in a face-to-face survey.

I support that graphic analysis with summary statistics for each group’s ideological consistency distribution and the industry standard for polarization measures—kurtosis. Sociologists have used variance
and kurtosis to quantify polarization in similar distributions at the aggregate level (Alwin & Tufiş, 2016; Baldassarri & Bearman, 2007; Hoffmann & Miller, 1998). Kurtosis, in particular, has helped to assess bimodality in contexts in which a large number of group-level observations are analyzed in relation to each other (see especially DiMaggio et al., 1996). Although kurtosis is sensitive to the tails of a distribution, it is not a direct measure of bimodality, and it can indicate different shapes, such as flat or U-shaped distributions, when very low (Westfall, 2014). In particular, bimodal distributions in which the two modes are not equal to the minimum and maximum values on a given scale are not easily distinguished by this measure. In short, I report kurtosis because it has some value of its own and for comparability with previous work, but the plots presented provide better information about the specific shape of each distribution and the bimodal character of the curve for likely cheaters.

Next, I use R’s Quantile Regression package (Koenker, 2015) on the matched data to model the shape and spread of these distributions while formally controlling demographic variables, political interest, and having Internet access at home. Instead of predicting the effect of each x on the mean of y, quantile regression can test whether the independent variables move respondents toward the tails of a distribution. I run models that predict decile points (i.e., the 10th, 20th, 30th, etc., percentiles). In these analyses, I analyze only respondents with political knowledge scores of 7 or 8, thus turning the survey mode dummy (online vs. face-to-face) into a flag for potential cheaters. H2 and H3 are evaluated together because they both draw on the density diagrams and quantile regression to distinguish between polarizing and rightward-pulling effects.

Finally, to assess whether likely cheaters were more influenced by potential bias in their own Internet news consumption habits (H4), I plot and report ordinary least squares (OLS) regression lines for likely cheaters, compared with other respondents, on a dot plot of the relationship between Internet news source bias and ideological consistency.

Analysis

Online Respondents Cheat

The center panel of Figure 1 shows that, despite being matched on the ten variables described above, online respondents performed much better on the eight political knowledge questions than the face-to-face respondents, with 23% answering seven or eight questions correctly. The clearest case of online political-knowledge seeking is evidenced in a question that asks respondents to identify John Roberts in an open-ended format. Note, in the left panel of Figure 1, that nearly equal proportions of the online and face-to-face samples gave the general answer “Supreme Court Justice,” but the specific answer “Chief Justice” was far more available to online respondents. More than 30% gave that answer, compared with only 9% in the face-to-face sample. This is an example of the sort of information that would appear in an Internet search for “John Roberts.”
Figure 1. Evidence that online respondents sought answers on the Internet.

This effect is even more evident in the third panel, which compares scores between the two samples on the 10-question vocabulary quiz. Data for that panel uses the same matching procedure as the other two except that online and face-to-face respondents are not matched on vocabulary scores. Whereas the face-to-face sample begins to taper off after seven correct answers, the distribution of scores among online respondents is linear, with a perfect score of 10 being the modal category for online respondents. In short, this evidence supports H1 by demonstrating that online respondents had ready access to factual answers and that many took advantage of that resource. Now the task is to determine how that may have influenced their opinions.

**Likely Cheaters Are More Conservative and More Polarized**

Figure 2 shows the density of ideological consistency scores for participants in the face-to-face versus Web samples by political knowledge scores. Decile divisions are also marked on the curve for likely cheaters (online respondents with high knowledge), and significant results from the quantile regression below are indicated with arrows and coefficients. In the left panel of Figure 2, respondents with less knowledge have tall, narrow distributions and smaller variances that indicate a predominance of middle-range consistency and mixed opinions. Knowledgeable respondents in the right panel, on the other hand, evidenced more consistent left or right opinions, creating a wider spread.
Figure 2. Density plots of ideological consistency by survey mode and political knowledge with significant effects of probable cheating, based on quantile regression results in Tables 2a and 2b.

Probable cheaters were the most polarized. Their distribution is short, broad, and strikingly bimodal. Table 1 confirms that they scored the highest and lowest observed values on the ideological consistency scale, and they evidenced the lowest (most extreme) kurtosis, as well as the greatest variance. The clear dip in the middle indicates a desertion of mixed and moderate views in favor of a new right-wing peak, which occupies a spot on the chart that knowledgeable face-to-face respondents did not. These results support both H3 (polarization) and H2 (a rightward pull).

In addition, retained knowledge among face-to-face respondents was notably associated with more liberal responses. This suggests that the two types of political knowledge have different effects. A comparison of means between the two groups of knowledgeable respondents, however, is complicated by their wider distributions, which, by definition, increases variance and standard errors around the mean, making significance tests unlikely to appear significant. I address this methodological challenge with quantile regression later. Finally, there is an overall rightward-pulling effect of taking the survey online, which is evident in significantly greater means and variances in the Web samples. Thus, likely cheaters may be wedged between the liberalizing effect of high political knowledge and the conservatizing effect of Internet exposure.
### Table 1. Descriptive Statistics and Tests for Figure 2.

<table>
<thead>
<tr>
<th>Knowledge:</th>
<th>Low Political Knowledge</th>
<th>High Political Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey Mode:</td>
<td>Face-to-face</td>
<td>Online</td>
</tr>
<tr>
<td>Mean</td>
<td>33.470</td>
<td>34.718</td>
</tr>
<tr>
<td>Diff. in Mean</td>
<td>−1.248****</td>
<td>−1.120</td>
</tr>
<tr>
<td>Variance</td>
<td>40.651</td>
<td>42.617</td>
</tr>
<tr>
<td>Diff. in Variance</td>
<td>−1.967*</td>
<td>−3.884</td>
</tr>
<tr>
<td>Median</td>
<td>34.183</td>
<td>34.976</td>
</tr>
<tr>
<td>Range</td>
<td>15.3–50.3</td>
<td>15.4–51.2</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>−0.447</td>
<td>−0.583</td>
</tr>
<tr>
<td>Matched &amp; Weighted N</td>
<td>835</td>
<td>3,452</td>
</tr>
</tbody>
</table>

Note. Means, medians, variances, and t tests for differences in means are weighted and use design-based standard errors. Differences in variance are significant at equal levels using both Levene's test for (non)homogeneity of variance and the Fligner-Killeen nonparametric test, which is robust for nonnormality. All analyses use matched data.

* p < .05. ** p < .01. *** p < .001, one-tailed tests.

The main effect of taking the survey online (net of its influence on knowledge) might indicate that online respondents consulted the Internet to help ensure that their answers to the opinion questions appeared ideologically consistent, even if they did not bother to do so for the knowledge questions. Although that issue cannot be explored further with these data, it is an interesting possibility that some survey respondents concern themselves more with political voice than with social desirability pressures frequently observed among face-to-face respondents, such as appearing politically knowledgeable.

In Figure 3, I demonstrate the usefulness of quantile regression for answering questions about the spread of these distributions by plotting political knowledge against ideological consistency. Two quantile regression lines follow the second and eighth deciles, while an OLS regression line estimates the mean. Variance in the dependent variable clearly increases with political knowledge, in violation of OLS regression assumptions. Moreover, that increasing spread is the very phenomenon I want to study. Quantile regression, then, allows me to predict spread as the way independent variables affect various points in the distribution.
Quantile regression analysis of only people with high political knowledge scores allows me to flag online respondents as potential cheaters and ask whether, for example, their top decile marker is in a significantly different place along the ideological consistency measure than the top decile marker for knowledgeable face-to-face respondents. The results of that analysis, reported in Tables 2a and 2b, were not breathtaking, but the positive and significant coefficients for online survey mode at the sixth and seventh deciles do support the conclusion that likely cheating contributed to a shift toward the political right, in support of H2. In addition, the lack of significant negative movement in the lower deciles confirms the visual analysis of Figure 2, demonstrating that the bimodal distribution results from a pull to the right, while the left-wing mode maintains its position. This supports H3 while clarifying that the process of polarization works through a pull to the right among likely cheaters (H2).

Other interesting results include an overall conservative effect of age, being White, and conservative Internet news bias; an overall liberal effect of high vocabulary scores (probably responsible for
the lack of effect for having a college degree); a liberalizing effect of income in the right-wing mode; a moderating effect of Latinx identity in the left-wing mode; and no effect for gender, political activity, or home Internet access.

Table 2a. Quantile Regression Results Predicting Ideological Consistency Among Politically Knowledgeable Respondents: Quantiles 1–4.

<table>
<thead>
<tr>
<th></th>
<th>1st Decile</th>
<th>2nd Decile</th>
<th>3rd Decile</th>
<th>4th Decile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online Mode</td>
<td>0.146 (0.512)</td>
<td>−0.053 (0.927)</td>
<td>0.251 (0.84)</td>
<td>1.096 (0.65)</td>
</tr>
<tr>
<td>Internet News Bias</td>
<td>3.393*** (0.298)</td>
<td>4.908*** (0.522)</td>
<td>6.327*** (0.435)</td>
<td>7.227*** (0.491)</td>
</tr>
<tr>
<td>Vocabulary Score</td>
<td>−0.927*** (0.165)</td>
<td>−0.865*** (0.208)</td>
<td>−0.670** (0.217)</td>
<td>−0.538* (0.249)</td>
</tr>
<tr>
<td>College Degree</td>
<td>−0.332 (0.524)</td>
<td>−0.302 (0.752)</td>
<td>−0.872 (0.65)</td>
<td>−0.699 (0.581)</td>
</tr>
<tr>
<td>Logged Income</td>
<td>0.157 (0.244)</td>
<td>0.692 (0.404)</td>
<td>0.360 (0.483)</td>
<td>−0.421 (0.518)</td>
</tr>
<tr>
<td>Year=2016</td>
<td>−1.968*** (0.54)</td>
<td>−1.523* (0.635)</td>
<td>−1.313* (0.585)</td>
<td>−1.361* (0.675)</td>
</tr>
<tr>
<td>Age</td>
<td>0.057*** (0.015)</td>
<td>0.043* (0.021)</td>
<td>0.060** (0.020)</td>
<td>0.070** (0.021)</td>
</tr>
<tr>
<td>White</td>
<td>2.330* (1.074)</td>
<td>2.097* (0.878)</td>
<td>1.248 (0.985)</td>
<td>1.651 (0.930)</td>
</tr>
<tr>
<td>Latinx</td>
<td>3.311* (1.405)</td>
<td>3.729** (1.331)</td>
<td>1.771 (1.119)</td>
<td>0.417 (0.998)</td>
</tr>
<tr>
<td>Male</td>
<td>0.505 (0.505)</td>
<td>1.038 (0.59)</td>
<td>1.225 (0.643)</td>
<td>1.195 (0.722)</td>
</tr>
<tr>
<td>Politically Active</td>
<td>−0.556 (0.535)</td>
<td>−0.091 (0.596)</td>
<td>0.757 (0.563)</td>
<td>0.699 (0.708)</td>
</tr>
<tr>
<td>Home Internet</td>
<td>−1.459 (0.788)</td>
<td>−0.638 (0.918)</td>
<td>0.885 (0.806)</td>
<td>2.304 (1.921)</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>19.218*** (2.346)</td>
<td>14.58*** (2.748)</td>
<td>10.557*** (2.846)</td>
<td>9.248* (3.776)</td>
</tr>
</tbody>
</table>

Note. N = 1,134 respondents with very high political knowledge scores. * p < .05. ** p < .01. *** p < .001, two-tailed tests.

Table 2b. Quantile Regression Results Predicting Ideological Consistency Among Politically Knowledgeable Respondents: Quantiles 5–9.

<table>
<thead>
<tr>
<th></th>
<th>5th Decile</th>
<th>6th Decile</th>
<th>7th Decile</th>
<th>8th Decile</th>
<th>9th Decile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online Mode</td>
<td>1.360 (0.913)</td>
<td>1.888* (0.881)</td>
<td>2.705* (1.127)</td>
<td>2.009 (1.320)</td>
<td>1.549 (0.993)</td>
</tr>
<tr>
<td>Internet News Bias</td>
<td>7.235*** (0.509)</td>
<td>6.596*** (0.528)</td>
<td>6.185*** (0.636)</td>
<td>5.428*** (0.587)</td>
<td>4.629*** (0.618)</td>
</tr>
<tr>
<td>Vocabulary Score</td>
<td>−0.628* (0.268)</td>
<td>−0.611** (0.228)</td>
<td>−0.493* (0.205)</td>
<td>−0.424 (0.232)</td>
<td>−0.392 (0.239)</td>
</tr>
<tr>
<td>College Degree</td>
<td>0.039 (0.704)</td>
<td>−0.265 (0.718)</td>
<td>−0.440 (0.754)</td>
<td>−0.823 (0.821)</td>
<td>−0.135 (0.845)</td>
</tr>
<tr>
<td>Logged Income</td>
<td>−0.985* (0.469)</td>
<td>−1.089* (0.434)</td>
<td>−1.159** (0.376)</td>
<td>−0.685* (0.327)</td>
<td>−0.483 (0.529)</td>
</tr>
<tr>
<td>Year=2016</td>
<td>−2.020** (0.672)</td>
<td>−2.640*** (0.642)</td>
<td>−3.249*** (0.666)</td>
<td>−2.254** (0.837)</td>
<td>−1.606* (0.776)</td>
</tr>
<tr>
<td>Age</td>
<td>0.090*** (0.022)</td>
<td>0.095*** (0.022)</td>
<td>0.093*** (0.025)</td>
<td>0.036 (0.026)</td>
<td>0.026 (0.026)</td>
</tr>
<tr>
<td>White</td>
<td>1.855* (0.819)</td>
<td>3.184*** (0.817)</td>
<td>2.183 (1.710)</td>
<td>1.236 (1.590)</td>
<td>0.304 (2.109)</td>
</tr>
<tr>
<td>Latinx</td>
<td>0.452 (2.050)</td>
<td>2.746 (1.528)</td>
<td>0.657 (2.883)</td>
<td>1.177 (2.114)</td>
<td>−1.928 (3.552)</td>
</tr>
<tr>
<td>Male</td>
<td>1.016 (0.747)</td>
<td>0.505 (0.712)</td>
<td>0.914 (0.616)</td>
<td>1.241 (0.679)</td>
<td>1.28 (0.989)</td>
</tr>
<tr>
<td>Politically Active</td>
<td>0.612 (0.823)</td>
<td>−0.537 (0.827)</td>
<td>−0.214 (0.618)</td>
<td>0.858 (0.784)</td>
<td>1.426 (0.78)</td>
</tr>
<tr>
<td>Home Internet</td>
<td>0.695 (2.75)</td>
<td>0.414 (1.252)</td>
<td>1.659 (1.731)</td>
<td>0.691 (1.498)</td>
<td>−1.079 (4.391)</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>14.211*** (4.288)</td>
<td>17.985*** (3.247)</td>
<td>19.344*** (3.847)</td>
<td>25.979*** (3.829)</td>
<td>32.748*** (5.947)</td>
</tr>
</tbody>
</table>

Note. N = 1,134 respondents with very high political knowledge scores. * p < .05. ** p < .01. *** p < .001, two-tailed tests.
**Biased Internet News Sources Influence Likely Cheaters More**

Although we cannot know exactly where probable cheaters sought information during the survey, respondents were asked whether they get news from the Internet, and if so, what specific sites they visit. Figure 4 shows that left-versus-right Internet news source bias is an important factor in ideological consistency, overall—with liberal news consumption encouraging liberal ideological consistency, and vice versa. That effect is even stronger among likely cheaters, however, where the steeper dotted line reflects stronger effects on both the liberal and conservative ends. Both effects are strong and significant ($p < .001$, in a two-tailed test), as detailed in the regression results reported following Figure 4 (Table 3).

*Figure 4. The effect of Internet news source bias on ideological consistency among probable cheaters with OLS regression lines. Shading represents 95% confidence intervals.*
Table 3. Linear Regression Results Represented in Figure 4.

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>(SE)</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>24.553</td>
<td>(0.807)</td>
<td>30.44</td>
<td>0.0000****</td>
</tr>
<tr>
<td>Internet News</td>
<td>3.556</td>
<td>(0.278)</td>
<td>12.79</td>
<td>0.0000****</td>
</tr>
<tr>
<td>Source Bias</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likely Cheaters</td>
<td>−9.522</td>
<td>(1.331)</td>
<td>−7.15</td>
<td>0.0000****</td>
</tr>
<tr>
<td>Internet Bias:</td>
<td>3.216</td>
<td>(0.474)</td>
<td>6.79</td>
<td>0.0000****</td>
</tr>
<tr>
<td>Likely Cheaters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 5,477  Degrees of freedom = 5,469  Residual deviance: 236,000  AIC: 37,700

Note. Estimates and standard errors are weighted and adjusted for complex sampling design using svyglm in R’s survey package (Lumley 2004).
* p < .05. ** p < .01. *** p < .001, two-tailed tests.

The results for likely cheaters are important because they show that Internet news bias is a source of left-versus-right-leaning ideological consistency in the moment, when respondents are researching their answers to opinion questions. Although such respondents do not necessarily consult news sources for that information, the slant of their media habits could affect more than the interpretive lens they bring to political information. Their everyday media habits could, in fact, influence the search results they get when looking for something as simple as Senate terms or John Roberts’ position, via search algorithms and selective attention (Bakshy, Messing, & Adamic, 2015; Hargittai, Gallo, & Kane, 2007) and by way of personalized political ads served to the respondent who searches for related information (Granka, 2010). In short, H4 is supported.

Conclusion

Survey cheaters are different from other respondents in that they were likely exposed to the Internet while answering a variety of opinion questions, and exposure increased polarization in their replies, as measured by ideological consistency. Moreover, likely cheaters reported more conservative views and fewer mixed or moderate responses, resulting in a clear bimodal distribution of ideological consistency across 41 survey questions. This effect was evident using matched data and held in quantile regressions that controlled Internet news bias, political activity, home Internet access, and demographic variables. Exposure to biased Internet news sources also increased polarization, and that effect was stronger among people who likely accessed the Internet during their survey. In short, all four hypotheses were supported in favor of the overall proposition that exposure to the Internet can influence reported opinion in real time. The tension between H2 (polarization) and H3 (a conservative pull) was clarified in the finding that polarization occurs by pulling moderate responses to the right.

This study contributes to the literature on survey cheating by demonstrating that the implications of consulting the Internet extend beyond our current focus on the mismeasurement of knowledge (e.g., Burnett, 2016; Clifford & Jerit, 2016; Jensen & Thomsen, 2014; Motta et al., 2016). Likely cheating affected opinion responses, as well. It did so within the time frame of a single survey, and it did so in ways predicted by the literature on mass media and opinion polarization. As such, likely cheaters reported opinions that were more conservative, more consistent, and more extreme. Although these results control for the overall
effect of political knowledge and for taking the survey online, it should be noted that the method used for identifying probable cheaters was not exact. Respondents who sought help answering questions may have consulted a nearby person for that assistance. As with any single study, more research will be required to confirm these results.

In addition, classical concerns about mass-media messages have recently shifted focus from dominant ideology to the dangers of losing those singular unified messages and replacing them with polarized echo chambers (e.g., Baum & Groeling, 2008). I reconcile the mass-media hypothesis with polarization hypotheses by showing that polarization occurs by pulling responses from the center to the right. In addition, because political knowledge is related to a leftward pull, whereas likely cheating is related to a rightward pull, these findings suggest that Internet sources might lean to the right. In this sense, Internet knowledge seems to impose conservative mass-media effects. The consequences of missing this point could be devastating, both for political processes and for the critical evaluation of new media communication.

To the polarization literature, this study offers two contributions. First, contemporary polarization might occur primarily through a pull to the political right. Second, some of that might be ephemeral—generated in the moment via exposure to political information on the Internet, specifically by respondents who do not retain much information about politics. More research will be required to explore the possibility of longer term effects. This study, however, offers an important road sign for future research on the relationship between Internet information and political polarization.

In the sizeable literature on biased media messages, rising polarization and rightward pulls are no surprise, taken separately. Their (apparently swift) connection, however, is an important new consideration. Still, the ability of opinion surveys to tap “real” belief is controversial, both in whether such belief systems exist and in whether individual respondents care about or subscribe to the survey options provided. Internet exposure could simply help respondents align their answers within a dominant political schema evident in both cultural domains (the Internet and the opinion survey). Even if polarization in survey responses is epiphenomenal, this study reveals the influence of Internet access on observed polarization in survey responses—a key source of evidence for claims about “real” opinion polarization, which have political influence of their own.

To that concern, Prior (2007) has argued that most media consumers are not so much polarized as they are uninterested in politics. Having deserted their collective posts in front of the six o’clock news, he argues, the majority merely leave politics to those with greater interest and, thus, stronger ideological affiliations. This study, however, asks what happens when respondents who don’t know the answer to a survey question go seeking political information on the Internet. The finding that Internet exposure pulls responses out of the middle and toward the right is also in keeping with some polarization research (e.g., Brooks & Manza, 2004; Bryson 2020; Evans, 2003; Mansbridge & Martin, 2013). Bourdieu (1984) might even argue that survey cheating pulls those who have been disenfranchised by our elite political culture into a conversation they would otherwise have avoided. That the conversation in question is polarized is likely something that will ebb and flow over time.
Since the early days of U.S. polarization research, when scholars found little to no evidence of increasing polarization in the general public, exceptions to that rule included respondents who were politically active (DiMaggio et al., 1996; Evans, 2003), knowledgeable (Layman et al., 2006), and sophisticated (Baldassarri & Gelman, 2008). Thus, it may be through the mechanism of political “knowledge” (however balanced or biased) that Internet consumers arrive at their reported opinions. Again, fast and easy access to information might have the potential to transform the politically uninterested into committed ideologues.

References


