

# Improving Media Effects Research through Better Measurement of News Exposure

**Markus Prior** Princeton University

*Survey research is necessary to understand media effects, but seriously impeded by considerable overreporting of news exposure, the extent of which differs across respondents. Consequently, apparent media effects may arise not because of differences in exposure, but because of differences in the accuracy of reporting exposure. Drawing on experiments embedded in two representative surveys, this study examines why many people overstate their exposure to television news. Analysis indicates that overreporting results from unrealistic demands on respondents' memory, not their motivation to misrepresent or provide superficial answers. Satisficing and social desirability bias do not explain overreporting. Instead, imperfect recall coupled with the use of flawed inference rules causes inflated self-reports. To lower reports of news exposure and improve the validity of conclusions about media effects, researchers should help respondents with the estimation by providing population frequencies and encouraging comparison with others.*

People who watch news or read newspapers cast more informed votes. Exposure to political advertising affects turnout and vote choice. Political comedy and other infotainment blends increase political awareness among otherwise hard to reach young people. Watching *Fox News* distorts perceptions of reality. Incessant coverage of polls and strategy makes people cynical. These are only a few of the many proposed effects of media exposure. If true, they testify to the central role of mass media in politics. Yet not too long ago, Bartels concluded that “[t]he state of research on media effects is one of the most notable embarrassments of modern social science” (1993, 267). This article aims to improve one of the tools social scientists use to study media effects.

Media effects research has made great progress in the last two decades. Experimental research, both in the lab (e.g., Iyengar 1991; Iyengar and Kinder 1987; Neuman, Just, and Crigler 1992) and in the field (e.g., Gerber and Green 2000; Gerber et al. 2007), convincingly established that media or campaign messages can cause change in political attitudes and behaviors. Aggregate studies relate systematic variation in media or campaign content to outcomes such as turnout or vote choice (e.g., Johnston, Hagen, and Jamieson 2004). Survey research, however, the methodology that

provided many of the early insights about the role of media in politics (e.g., Lazarsfeld and Berelson 1944; Patterson and McClure 1976), has not kept up with this progress. Its key problem is that many survey respondents are bad at telling us accurately about their media exposure (Ansolabehere, Iyengar, and Simon 1999; Price and Zaller 1993; Prior, 2009). This casts serious doubt on survey-based evidence about the relationship between media exposure and political outcomes. Apparent media effects may arise not because of differences in exposure, but because of differences in the accuracy of reporting exposure. Null findings may reflect not the absence of media effects, but flawed measures of exposure.

Yet surveys remain indispensable for a full understanding of media and campaign effects. Outside the experimental lab, we need surveys to pinpoint who was exposed to which and how much media content—and thus who should be affected by it. Even experimental research needs self-reported exposure to rule out that experimental effects occur only among subjects who never encounter the stimulus in the real world. Rather than dismiss self-reports altogether, we need to understand better why people do not provide more accurate responses. The purpose of this article is to focus attention on the measurement of media

exposure in surveys as a major obstacle for the study of media effects and to devise new ways to counteract people's tendency to overreport exposure.

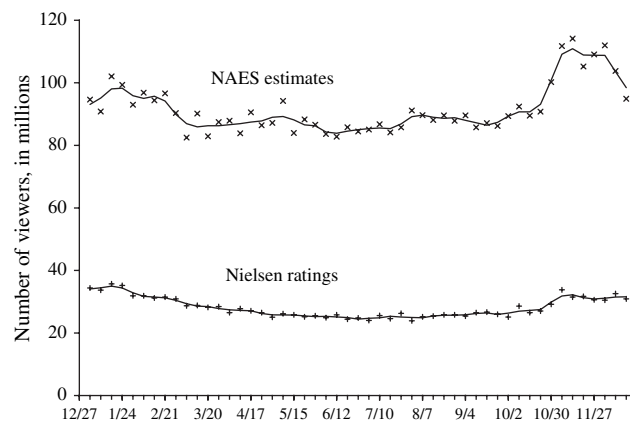
This is the first study to examine why many people dramatically overstate their news exposure and how this overstatement distorts our understanding of media effects. In this article, I draw on research on the cognitive mechanisms of survey responses to formulate hypotheses about the causes of overreporting. At the heart of this puzzle is the distinction between motivation and ability: Do respondents deliberately overstate their news exposure—to look good, or to finish the survey quickly? Or are they frustrated by the difficulty of the task despite their best intentions? Hypotheses are tested in a series of experiments embedded in two surveys, one conducted over the phone, the other online. As a practical benefit, this project develops new question formats that reduce overreporting. Following the presentation of the experimental results, I examine if items that reduce reports of news exposure change conclusions about the causes and consequences of news exposure.

Despite the known shortcomings of self-reported exposure, research into media effects on political behavior and public opinion continues to rely on these measures. In the last 10 years alone, the *Journal of Politics*, the *American Political Science Review*, and the *American Journal of Political Science* have published 42 studies that use self-reported media exposure in their empirical analysis.<sup>1</sup> Even though media use was not the theoretical focus in all of them, simply controlling for self-reported exposure easily biases the effects of other variables if the reporting errors are not random (see Brady 1986; Katz 2000). Vote validation studies have demonstrated how adjusting for self-report bias can lead to a different and more valid understanding of political behavior. For example, analyses of validated turnout show a much weaker, sometimes not even significant, effect of education on turnout because some of the more educated, but few of the less educated respondents falsely report that they voted (Belli, Traugott, and Beckmann 2001; Presser and Traugott 1992; Silver, Anderson, and

Abramson 1986). Whether or not overreporting of news exposure arises for the same reasons as turnout overreporting, measuring media exposure in surveys with less error promises more valid conclusions about media effects on political behavior.

In the case of news exposure, it is difficult to exaggerate the magnitude of overreporting. Figure 1 illustrates the inaccuracy of self-reported exposure to network news. It compares survey estimates from the 2000 National Annenberg Election Survey to Nielsen estimates of the nightly news audience for the same period. (For details, see note to Figure 1 and Prior, 2009.) According to Nielsen data, between 30 and 35 million people watched the nightly news on an average weekday in 2000. Based on self-reports, that number is between 85 and 110 million. Hence, survey

**FIGURE 1 Comparing Nielsen Ratings and Survey Estimates of the Nightly Network News Audience**



*Note:* The NAES (Romer et al. 2004) was conducted as a rolling cross-sectional design and produced independent daily random samples of U.S. residents for the entire year 2000. Respondents were asked "How many days in the past week did you watch the national network news on TV—by national news, I mean Peter Jennings on ABC, Dan Rather on CBS, Tom Brokaw on NBC, Fox News or UPN News?" Responses to the NAES question are divided by seven and averaged across respondents to calculate the expected daily network news audience. Daily estimates are aggregated to generate weekly averages. Multiplying these percentage estimates by the voting-age population in 2000 produces survey-based estimates of the daily network news audience (averaged by week). Nielsen estimates are weekly averages of the average weekday audience for the ABC, CBS, and NBC evening newscasts (measured in number of viewers). These estimates are available from Associated Press reports through Lexis-Nexis. Because these average audience measures count viewers only in proportion to the portion of the newscast they watched, they underestimate the total weekday audience. But they also miss weekend viewing, which is considerably lower than weekday viewing. Alternative Nielsen measures show that these two biases are relatively small (less than 5 million viewers in each direction even under the most conservative assumptions) and roughly neutralize each other. (For more information, see Prior, 2009.) The trend lines are generated using locally weighted regression on time with a bandwidth of .1.

<sup>1</sup>These studies were identified by searching the three journals' electronic archives (JSTOR, Blackwell-Synergy, Cambridge Journals Online) using the following search terms for the years 1998 to 2007: media exposure, media use, television viewing, television watching, television exposure, news exposure, debate viewing, ad recall, ad exposure, advertising exposure, newspaper reading, campaign exposure, news consumption, and media consumption. Studies that did not use self-reported exposure to mass media in their empirical analysis were not counted.

estimates overstate the size of the network news audience by a factor of three.

## What Causes Inflated Self-Reports of News Exposure?

Survey methodologists have developed a model that specifies what respondents do when they answer a survey question about the frequency of their past behavior (Schwarz and Oyserman 2001; Tourangeau, Rips, and Rasinski 2000; for different versions of the model, see Schwarz 1999). According to this model, respondents have to (1) understand the question, (2) recall the relevant behavior, (3) estimate the frequency of the relevant behavior, (4) map the frequency onto the response alternatives, and (5) report either their candid answer or a socially desirable answer.

Errors at the first or fourth stage are least likely for some common question about news exposure. Respondents appear to understand what we mean by “network news.” In 2000, the American National Election Study (ANES) and the National Annenberg Election Survey (NAES) used different questions about network news exposure. The ANES asked respondents “How many days in the past week did you watch the national network news on TV?” To this question, the NAES added the explanation that “by national news, I mean Peter Jennings on ABC, Dan Rather on CBS, Tom Brokaw on NBC, Fox News, or UPN News.” Mentioning “Fox News” and UPN News even though neither the Fox broadcast network nor UPN has national news programs did not confuse respondents: Despite the different question wordings (and different survey procedures), means and variances are statistically indistinguishable. During the period when both surveys were in the field (September 5 to November 6), the weighted mean response to the ANES question was 3.19 days with a standard deviation of 2.80, compared to a mean of 3.06 and a standard deviation of 2.66 for the NAES. The overreporting in Figure 1 is robust to variations in how “network news” is defined.

The mapping of frequencies onto response options (stage 4) is a serious concern for questions that rely on “vague qualifiers” (Schwarz and Oyserman 2001) such as “sometimes” and “regularly” because respondents may be unsure which response option corresponds to their estimated frequency of exposure (Schwarz et al. 1985). This problem is largely absent for the question format most widely used in the ANES and other prominent political surveys. Asked “how many days in the past week” they watched a particular type of program, respondents report the number of days.

Despite this easy mapping of estimated frequencies into responses, the question still produces seriously inflated reports (see Figure 1).

This study thus focuses on the remaining three stages of the general self-report model. At the second stage, respondents may not recall all episodes of the behavior or incorrectly recall them as having occurred during the reference period (an error called “telescoping”):

When the question pertains to a frequent behavior, respondents are unlikely to have detailed representations of numerous individual episodes of a behavior stored in memory. Instead, the various instances of closely related behavior blend into one global, knowledge-like representation that lacks specific time or location makers . . . As a result, individual episodes of frequent behaviors become indistinguishable and irretrievable. (Schwarz and Oyserman 2001, 136–37)

When respondents believe that they recalled some, but not all episodes of the behavior, they estimate its frequency (stage 3). Their estimation can be based on rules, on various heuristics—for example, on the availability heuristic (Tversky and Kahneman 1973), so easily recalled episodes weigh more heavily—or on an “innate sense of relative or absolute frequency” (Burton and Blair 1991, 52). Importantly, this recall and estimation process can produce systematic overreporting. For example, overreporting occurs when respondents confuse instances of the relevant behavior with instances where they only considered the behavior without actually engaging in it (Belli, Traugott, and Beckmann 2001; Belli et al. 1999). Overreporting is also likely when respondents rely heavily on estimation instead of enumeration because estimation tends to generate higher self-reports (Brown and Sinclair 1999; Burton and Blair 1991).

A simple explanation of overreporting thus centers on lack of effort. According to my *satisficing hypothesis*, overreporting occurs due to satisficing (Krosnick 1991) when respondents are not motivated enough to search their memory thoroughly for instances of news exposure or to come up with a valid inference rule. Burton and Blair (1991), for example, find that instructing business majors to take additional time improved the accuracy with which they recalled course grades and courses taken in other departments.

Even considerable effort may not reduce overreporting if recall and estimation are simply too difficult. According to Schwarz and Oyserman (2001, 137), episodes of news exposure could be “indistinguishable and irretrievable” even for the most motivated respondent. Conrad, Brown, and Cashman conclude that poor recall may force respondents to

base their self-reports on as little as “general impressions” so “[a]ll they can do is convert their impression that frequency is high to a relatively large number” (1998, 352).

According to the *flawed estimation hypothesis*, respondents use inappropriate estimation strategies to infer their news exposure based on limited recall of relevant episodes. Only explicit help with estimating news exposure can improve self-reports. Burton and Blair speculate that accuracy of self-reports may be increased by “offering information about the frequency distribution across the respondent population,” which “may be particularly helpful in situations where respondents become overwhelmed by the estimation task and simply guess” (1991, 77). According to this argument, giving respondents an idea of the size of the audience or encouraging them to think about other people’s viewing habits can improve their inference rules and reduce overreporting.

Finally, overreporting may arise at the fifth stage of the self-report model. The *social desirability hypothesis* holds that overreporting occurs because a respondent feels that his exposure estimate is embarrassingly low and adds a few days of news viewing to look good to himself or the interviewer. Social desirability pressure is the most prominent explanation for turnout overreporting (Belli et al. 1999; Bernstein, Chadha, and Montjoy 2001; Holbrook and Krosnick 2005; Presser 1990). It may inflate reports of news exposure as well because staying informed is also considered indicative of a “good citizen.” Whether social desirability bias causes overreporting of news exposure has never been tested.

## Research Design and Data

To investigate if social desirability bias, satisficing, or flawed estimation cause inflated self-reports of news exposure, I designed a series of survey experiments. They were included in two different general population surveys. Study 1 was a phone survey with a probability sample of U.S. residents 18 and older conducted by the Center for Survey Research at Indiana University. Data for this study were collected through Time-Sharing Experiments for the Social Sciences (TESS), funded by the National Science Foundation. My module was in the field between October 13, 2005 and February 7, 2006 and yielded 916 completed interviews.<sup>2</sup> The module was placed first in the interview.

<sup>2</sup>The response rate (AAPOR response rate 3) for the two modules was 30.2%.

Study 2 was conducted by Knowledge Networks in March and April 2008. Knowledge Networks interviews national probability samples over the Internet by providing a large panel, selected through random digit dialing, with WebTV units and free Internet connections in exchange for taking surveys. The participants for this study constitute a randomly selected subset of the KN panel and are thus close to a random sample of the U.S. adult population.<sup>3</sup> Data from both studies were weighted to reflect Census distributions on gender, age, race, and region.

In Study 1, all respondents participated in three different experiments. First, the List Experiment asked respondents if they performed a variety of behaviors, including watching the news, on the day before the interview. It constitutes a precise test of the social desirability hypothesis. The next experiment covers general television news exposure and evaluates the satisficing hypothesis. Finally, the Anchor Experiment tests the flawed estimation hypothesis for exposure to nightly network news. All assignments to different experimental groups were random and orthogonal to the assignments for the other experiments. Study 2 replicates the Anchor Experiment for a different interview mode and extends it by measuring response time and testing alternative explanations. Unlike studies of turnout overreporting, I cannot directly validate respondents’ self-reported news exposure, so hypotheses are evaluated by comparing the degree of overreporting in the different experimental conditions. Appendix A reports a check of the randomizations for both studies.

## Results

To test the *social desirability hypothesis*, I designed the *List Experiment*. It allows respondents to report their news exposure without revealing it to the interviewer, thus relieving presentational pressures. The experiment relies on a technique that has been used successfully in the past to show social desirability bias. The “item count technique” (Miller 1984) was validated in a number of different contexts, demonstrating social desirability bias in reports of illegal drug use and risky sexual behavior, for example (for a summary of past studies, see Holbrook and Krosnick 2005). Under the name “list experiment,” it has been used in political science to study racial attitudes (e.g.,

<sup>3</sup>The survey completion rate was 61%. The following analyses exclude 28 cases because of completely missing data or response set on an eight-item response grid just preceding the exposure question.

Gilens, Sniderman, and Kuklinski 1998; Kuklinski, Cobb, and Gilens 1997). Holbrook and Krosnick (2005) have recently used the technique to document social desirability bias in turnout reporting.

In this experimental design, two-thirds of the respondents are read a list of behaviors and asked to report how many of those behaviors they performed the day before the interview. In one experimental condition, the list contains the sensitive behavior, in this case watching a news program on television. In a second condition, the list is identical except that the sensitive behavior is not included. The difference between the two list conditions estimates the frequency of the sensitive behavior even though no respondent is directly asked about it. The behaviors on the list are chosen so that respondents will not have performed all or none of the behaviors because in those cases respondents could in fact reveal to the interviewer whether or not they watched the news. The frequency estimate from the list experiment can then be compared to a question that asks the last third of respondents directly whether or not they engaged in the sensitive behavior the day before the interview. If social desirability inflates frequency estimates, the direct estimate should be higher than the estimate inferred from the list conditions.

In the List Experiment, respondents were randomly assigned to one of three conditions, the five-item list which includes television viewing, the four-item list without television viewing, or a direct question about television viewing. One-third of the respondents was asked directly

Did you watch a news program on television yesterday?

A second third was asked the following question:

Here is a list of four things that some people do and some people do not do. Please listen to them and then tell me how many of them you did yesterday. Do not tell me which things you did. Just tell me how many. Here are the four things.

- Had a cup of coffee
- Made a phone call
- Went to the movies
- Took a shower

How many of these things did you do yesterday?

For the final third of respondents, the list included a fifth behavior: “Watched a news program on television.” In response to the direct question, 66% of the respondents reported exposure to a TV news program “yesterday” (N = 310). In contrast, the difference between the mean number of behaviors in the two list conditions was .71 (3.26 for the five-

item list [N = 318] minus 2.55 for the four-item list [N = 282]),<sup>4</sup> indicating that 71% of the respondents watched a news program the day before the interview. The difference between the direct and indirect estimate is insignificant and in the wrong direction, thus failing to support the social desirability hypothesis.

Using the same technique to study turnout overreporting, Holbrook and Krosnick (2005) find that social desirability affects turnout self-reports in phone interviews, but not when respondents answer the questions on their computer screens. The present study, in contrast, finds no indication of social desirability bias in self-reports of news exposure precisely for the interview mode—telephone—that exacerbated social desirability bias for turnout. Lack of statistical power cannot explain the failure to find significant results because the effect estimates were in the wrong direction. Social desirability bias does not appear to affect overreporting of news exposure. This conclusion is consistent with results of an experiment in the 1989 ANES pilot study testing different wordings of the news exposure question. Respondents who were asked about their news exposure “in a typical week” reported higher news exposure than respondents asked about the “past week.” But they also reported higher exposure to nighttime television entertainment programs (Price 1993), a behavior that should not be influenced by considerations of social desirability in the same way as news exposure.

As a test of the *satisficing hypothesis*, the *15-Seconds Experiment* was designed to determine if respondents misreport their exposure to television news because they do not devote sufficient effort to recalling instances of news exposure and estimating a frequency. Drawing on a design employed by survey methodologists in the past (Burton and Blair 1991), the treatment in the 15-Seconds Experiment induces respondents to think more carefully about an exposure question than they normally would. The treatment group received the following instructions:

The next question is very important. After I finish reading the question, I would like you to spend at least 15 seconds thinking about it. I will let you know when the 15 seconds are up. If you wish to take more time, just let me know. Okay?

Respondents in the control group did not hear any introduction before the exposure question:

<sup>4</sup>The list of behaviors was pretested and worked as intended. Only 2.8% of the respondents in the two list conditions reported engaging in either the minimum or the maximum number of behaviors.

In a typical week, how many days do you watch the news on television?

The treatment did not significantly affect reported exposure. Respondents in the treatment group reported 4.9 days of news exposure, compared to 4.7 days in the control group. The difference remains insignificant when political interest and demographics are controlled for.<sup>5</sup>

For a second test of the satisficing hypothesis, I used a *Follow-Up Challenge* to give respondents an opportunity to reconsider and update their answers. Using a similar procedure, Burton and Blair (1991, 63) find that additional response time improved the accuracy of self-reports for some behaviors. Following their report of general TV news exposure, respondents were alerted to the difficulty of estimating news exposure and given a chance to change their answer:

Just to be sure because people often find it difficult to estimate their news exposure correctly, and thinking about it very carefully, how many days in a typical week do you watch the news on television?

Only 13% of the respondents provided a different answer in response to the follow-up question. This result confirms previous findings that demonstrate fairly high reliability of self-reported news exposure (Price 1993). Furthermore, almost half of the respondents who provided a different second self-report increased their estimate. Hence, even in the few instances when additional consideration led respondents to change their initial answer, overreporting remained at the same level as before.

In sum, simply encouraging people to try harder does not reduce overreporting of news exposure, according to two independent tests. In Burton and Blair's study, the self-reported frequency of ATM withdrawals and check writing did not become more accurate when respondents had extra time. They conclude that "a response-time manipulation . . . is only effective at improving response accuracy when the episodes in question are readily accessible in a survey context" (1991, 75). For respondents who do not store individual episodes of low-salience behaviors such as ATM withdrawals or news viewing in memory, the estimation problem seems to be so difficult that greater effort alone does not help.

Rejection of the social desirability and satisficing hypotheses suggests that people inadvertently over-

estimate their news exposure despite their best efforts to respond accurately. To test this *flawed estimation hypothesis*, I designed the *Anchor Experiment*. It follows survey methodologists' recommendations to point respondents towards effective estimation rules. Information about the frequency of the behavior in the population offers respondents a reference point for their estimates and encourages them to consider whether or not this reference point applies to them (Burton and Blair, 1991, 77). Inviting comparisons with other people can also help estimation (Schwarz et al. 1985). As part of the Anchor Experiment in both studies, respondents were asked how often they watched network news:

The next question is about the nightly national network news on CBS, ABC and NBC.

This is different from local news shows about the area where you live and from cable news channels such as CNN and Fox News Channel.

How many days in the past week did you watch the national network news on television?

The control group heard only this question. In three treatment groups, different (randomly assigned) introductory statements preceded the question. They alerted respondents that audiences may change over time, thus cautioning them about inference rules based on impressions from the more distant past. One subset of respondents was told that

Television news audiences have declined a lot lately. Few Americans watch the national network news on a typical weekday evening.

In order to determine if a concrete population frequency has a different effect than vague population information, a second subset was told that

Television news audiences have declined a lot lately. Less than one out of every ten Americans watches the national network news on a typical weekday evening.<sup>6</sup>

Both anchors should lower reported exposure because they give respondents some information on the actual size of the network news audience and explicitly mention other Americans to encourage comparison. But the competing social desirability hypothesis offers exactly the same prediction: low frequencies should make it more acceptable to report low news exposure. The social desirability and flawed estimation hypotheses yield different predictions only for high-population

<sup>5</sup>Subjects did apparently take more time to think about their answer. In Study 1, response time to individual questions was not measured, but the overall length of my survey module was. Subjects in the 15-Seconds condition took half a minute longer to complete the interview than those in the control group.

<sup>6</sup>In the TESS study, the second sentence was "Only one out of every ten Americans watches the national network news on a typical weekday evening." Counting Americans under 18, fewer than 10% of Americans watch network news on average.

frequencies. Hence, a third anchor was assigned to another subset of respondents:

With all that's going on in the world these days, many Americans watch the national network news on a typical weekday evening.

According to the social desirability hypothesis, this anchor should *increase* social desirability pressures and therefore inflate self-reports further. In contrast, the flawed estimation hypothesis predicts reduced overreporting because “[p]roviding respondents with information about the average person’s frequency is likely to encourage estimation based on the respondent’s sense of how he or she compares with the norm” (Burton and Blair 1991, 52–53). Respondents who do not watch a lot of news determine that they fit the norm when they hear that “few Americans” (or “1 out of every 10 Americans”) watch the news and report the low estimate. When they hear that “many Americans” watch the news, they realize that they *do not* fulfill this norm and *also lower their estimates*. The “many Americans” condition thus offers a direct test of the flawed estimation hypothesis against the social desirability hypothesis.

Respondents in all treatment conditions took significantly longer to complete the question than control subjects, according to response time measures included in Study 2. Median completion time was 21 seconds in the control group, compared to 29 seconds in the *Few Americans* treatment, 28 seconds in the *1 in 10* treatment, and 25 seconds in the *Many Americans* treatment.

All three anchors lowered self-reported news exposure. Table 1 shows regression estimates of the experimental effects holding constant a series of demographic and attitudinal factors. In Study 1, respondents reported an average of 2.6 days of network news exposure in the control group. News exposure was reduced by .4 days in the *Many Americans* condition, .5 days in the *Few Americans* condition, and .6 days in the *1 in 10* condition. The treatments thus reduced reported news exposure by between 16 and 22%.

In Study 2, overall treatment effects are smaller, reaching statistical significance only in the *Few Americans* condition (Table 1, column 2). A principal difference between the two studies is the interview mode. In the TESS study, respondents were read the question by the interviewer. In the KN study, respondents completed the survey question on a computer screen. Some of them did so very quickly (10%, for example, in 11 seconds or less). Respondents with short completion times may not have read the question carefully enough to be affected by experimental

manipulation. Figure 2 graphs the experimental effects as a function of screen completion time. The plot shows residual news exposure after the main effects of demographic predictors are removed. The gaps between the residuals in the control group and in the treatment groups indicate the experimental effects. The area between the vertical lines marks the range between the 5<sup>th</sup> and 95<sup>th</sup> percentiles of completion time. Treatment effects are small and inconsistent for respondents with short completion times, but increase for respondents who took longer.

The last two columns in Table 1 show experimental effects below and above the median completion time. The median split is implemented within each experimental group, so that the slow half of respondents in the control group is compared to the slow half of respondents in the treatment groups. None of the anchors made a significant difference for the respondents who completed the screen quickly. Among respondents who considered the question carefully, however, significant effects emerge. The *Few Americans* and *1 in 10* treatments both lowered self-reports by almost .7 days. The effect of the *Many Americans* treatment of .4 days is marginally significant ( $p = .10$  in a one-tailed test).

**How does the “many Americans” anchor lower self-reports?** Although in Study 2 the *Many Americans* treatment reduced self-reports only for respondents who took sufficient time, the two studies together reject the social desirability hypothesis, which predicted higher self-reports in the *Many* condition than in the control group. Consistent with the *flawed estimation hypothesis*, even a (vague) high population frequency reduced reported news exposure. Burton and Blair (1991) anticipated this result by suggesting that providing population frequencies encourages comparisons with others and thereby improves estimation. To verify this mechanism, Study 2 asked respondents right after the *Anchor Experiment* if they watched network news “more often, less often, or about as often as most other Americans.” Among respondents above the median completion time, 10% thought they watched news more often than most others in the *Many Americans* condition, compared to 16% in the control group. The percentage who said they watched less than others remained the same. Alerting respondents that “many Americans” watch the news thus helped at least some of them infer that their own news consumption is relatively low.

In the *1 in 10* and *Few Americans* conditions, the slower half of respondents also lowered their estimates of how their news consumption compares to others. Between 50 and 51% of them reported

**TABLE 1 The Impact of Inference Rules on Self-Reports of Nightly Network News Exposure (The Anchor Experiment)**

	Study 1 (TESS)	Study 2 (Knowledge Networks)		
		All respondents	Time < median	Time > median
<i>Experimental Conditions</i>				
“few Americans”	-.46** (.20)	-.38* (.20)	-.20 (.27)	-.65** (.30)
“1 out of every 10 Americans”	-.57** (.20)	-.20 (.19)	.26 (.26)	-.66** (.29)
“many Americans”	-.41** (.17)	-.17 (.20)	.09 (.26)	-.38 (.30)
“some watch . . . others don’t”		-.19 (.18)	-.06 (.25)	-.34 (.28)
“audiences have declined”		-.02 (.20)	.19 (.27)	-.24 (.30)
“network news draw large audiences”		.13 (.19)	.57** (.25)	-.30 (.29)
<i>Control Variables</i>				
Follows politics “most of the time”	.95** (.15)	1.18** (.13)	1.05** (.18)	1.31** (.19)
College graduate	-.41** (.16)	-.32** (.13)	-.45** (.17)	-.07 (.23)
Age <sup>a)</sup>	.05** (.004)	.04** (.004)	.05** (.006)	.03** (.01)
Female	-.10 (.14)	.24** (.11)	.38** (.16)	.15 (.17)
Minority	.23 (.15)	.57** (.13)	.57** (.18)	.56** (.19)
Employed	-.41** (.004)	-.33** (.12)	-.16 (.17)	-.40** (.18)
Internet connection speed		-.25** (.12)	.02 (.18)	-.42** (.18)
Conservative ideology	-.58** (.17)	-.44** (.14)	-.37* (.20)	-.53** (.21)
Liberal ideology	.29 (.20)	-.16 (.17)	-.11 (.22)	-.24 (.26)
Intercept	2.46** (.20)	2.44** (.18)	1.83** (.27)	2.89** (.26)
Adjusted R <sup>2</sup>	.25	.20	.17	.18
N	901	1531	777	754

\*\**p* < .05, \* *p* < .10 (two-tailed) <sup>a)</sup> deviation from sample mean.

*Note:* Cells show OLS coefficients with standard errors in parentheses. The regressions also include a dummy variable for respondents who did not answer the ideology question. The two rightmost columns break up KN respondents by completion time for the network news exposure screen.

watching less news than others, compared to only 44% in the control group. (The percentage who reported watching more than others remained almost unchanged.) The results suggest that all three anchors stimulated comparisons with other people that helped people reach a more realistic estimate of their own news exposure.

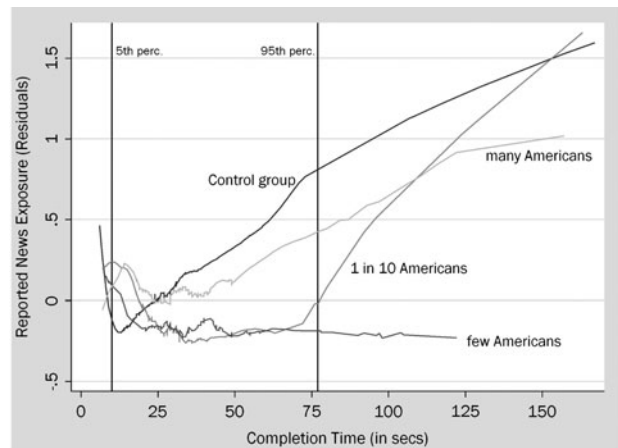
**Do population frequencies or references to others reduce overreporting?** All three anchors discussed so far combine information about the population frequency and a reference to other Americans. In order to determine if both elements are necessary to reduce overreporting, Study 2 included three other anchor conditions:

[Anchor 4: ] Television news audiences have declined a lot lately.

[Anchor 5: ] With all that’s going on in the world these days, the national network news draw large audiences.

[Anchor 6: ] With all that’s going on in people’s lives these days, some watch the national network news on a typical evening, while others don’t.

**FIGURE 2 Self-Reported News Exposure by Response Time, Anchor Experiment (Study 2)**



*Note:* Residuals were calculated by regressing self-reported news exposure on political interest, education, age, gender, minority status, income, Internet connection speed, tenure on the KN panel, and participation in concurrent political panel studies.



Anchor 4 begins exactly as the *1 in 10* and the *Few Americans* anchor, but omits the second sentence which explicitly refers to other people. Anchor 5 resembles the *Many Americans* anchor, but mentions only “large audiences” instead of “many Americans watch[ing]” news. As Table 1 shows, Anchors 4 and 5 did not significantly reduce overreporting. (In fact, Anchor 4 caused greater overreporting among fast respondents.) Neither did these treatments increase completion times as much as the other anchors. Anchor 6 refers to other people, but does not provide a population frequency. It lowers self-reports, but just misses conventional levels of statistical significance ( $p = .15$  for the entire sample;  $p = .11$  for the slower half; both one-tailed). These results suggest that only the combination of population frequency and explicit reference to other people remedies flawed estimation of news exposure.

### Who Overreports?

For two reasons, it is important to examine who lowers their reported news exposure most when offered help with the estimation. First, if everybody overreports to the same extent, the resulting errors will not distort estimates of media effects. If, in contrast, overreporting varies with other variables of interest, inferences will be biased (see Brady 1986; Katz 2000). Second, larger experimental effects among particular types of respondents can add circumstantial evidence that the anchors generate more accurate exposure estimates.

The models in Table 2 present separate regressions of reported news exposure on several demographic and attitudinal variables for the control group and the three effective anchor conditions. (The *Few Americans* and the *1 in 10* conditions are pooled to maximize cases as differences between the two conditions were negligible.) Study 2 included a separate question about respondents’ primary news source and a vocabulary test, which was adopted from the General Social Survey and is often interpreted as a measure of general cognitive abilities.<sup>7</sup> The effects of

both variables bolster the validity of the modified exposure question. Respondents who name television as their primary news source report more network news exposure than others. But this effect is more than twice as big (and statistically different) in the treatment conditions compared to the control group. The modified exposure questions match more closely what respondents report elsewhere in the interview.

Respondents with lower cognitive abilities report higher-than-average news exposure in the control group, but not in the treatment conditions. This difference, too, is statistically significant and consistent across treatments. Overreporting declines precisely among respondents—those with lower cognitive abilities—who should benefit the most from a treatment that reduces the difficulty of estimating one’s own news exposure.

Vote validation studies have shown that more educated and more politically involved respondents are more likely to falsely report that they voted (Belli, Traugott, and Beckmann 2001; Presser 1984; Presser and Traugott 1992; Silver, Anderson, and Abramson 1986). Politically interested college graduates are also most likely to lower their exposure reports in the anchor conditions. As Table 2 shows, in both studies the interaction between the two variables is significantly lower in the treatment conditions than in the control group. To facilitate interpretation, Figure 3 graphs predicted exposure reports in the different conditions as a function of interest and education (for a white, ideologically moderate, employed women). Experimental effects are small among respondents who are not very politically interested or do not have a college degree. Among politically interested college graduates, however, the anchors lower self-reported exposure by between 24 and 55%.

Why would more politically interested, educated respondents estimate their exposure less accurately than others? Subgroup analysis of the *List Experiment* yields no evidence for social desirability bias in this group. Instead, interested college graduates seem to rely too heavily on their generally high political involvement when estimating network news exposure without help. Even when they recall only a few episodes of news exposure, they may infer frequent exposure from their considerable interest in (and knowledge of) politics. Documenting a similar case of faulty reasoning, Bishop, Oldendick, and Tuchfarber (1984) have shown that survey respondents report “following what’s going on in government and public affairs” considerably less often when they are asked about it after several difficult political knowledge questions. People’s estimates of their own political

<sup>7</sup> Respondents were first asked, “From which of these media do you typically get most of your news?” Those who selected TV received a score of 2. Those who did not were then asked, “Do you regularly get news from any other medium?” and scored 1 if they selected TV. All others received a score of 0. In the vocabulary test, respondents were asked to select synonyms for eight different words (see Verba, Schlozman, and Brady 1995, 561–62). The variable is scored 0–8.

TABLE 2 Predictors of Self-Reported Network News Exposure in the Anchor Experiment

	<i>Experimental Condition</i>		
	Control Group	“few/1 out of 10 Americans”	“many Americans”
<b>Study 1 (TESS)</b>			
Follows politics “most of the time”	.86** (.31)	1.50** (.30)	1.07** (.29)
College graduate	-.31 (.48)	.10 (.40)	.13 (.38)
Follows politics × College Grad.	.37 (.61)	<b>-.99* (.52)</b>	<b>-1.21** (.51)</b>
Age <sup>a)</sup>	.06** (.01)	<b>.04** (.01)</b>	.05** (.01)
Female	-.09 (.26)	.12 (.23)	-.48** (.24)
Minority	.14 (.28)	-.02 (.25)	.72** (.26)
Employed	-.21 (.28)	-.68** (.25)	-.41* (.24)
Conservative ideology	-.72** (.34)	-.71** (.28)	-.65** (.27)
Liberal ideology	.78** (.38)	.15 (.33)	<b>-.04 (.31)</b>
Intercept	2.28** (.33)	1.85** (.28)	2.19** (.30)
Adjusted R <sup>2</sup>	.27	.25	.23
N	279	321	301
<b>Study 2 (Knowledge Networks)</b>			
Primary news source TV	.37** (.15)	<b>.77** (.15)</b>	<b>.85** (.21)</b>
Cognitive ability <sup>a)</sup>	-.28** (.06)	<b>-.07 (.07)</b>	<b>-.07 (.10)</b>
Follows politics “most of the time”	1.03** (.28)	<b>1.89** (.32)</b>	1.11** (.49)
College graduate	.04 (.35)	.50 (.33)	.64 (.47)
Follows politics × College Grad.	.16 (.52)	<b>-1.55** (.49)</b>	<b>-1.59** (.75)</b>
Age <sup>a)</sup>	.05** (.01)	.03** (.01)	.04** (.01)
Female	-.27 (.22)	<b>.33 (.22)</b>	<b>.55* (.32)</b>
Minority	.49* (.25)	.37 (.27)	<b>-.79* (.41)</b>
Employed	-.37 (.25)	-.28 (.26)	-.55 (.36)
Internet connection speed	-.18 (.24)	.06 (.25)	.22 (.36)
Conservative ideology	-.11 (.27)	-.53* (.30)	-.12 (.43)
Liberal ideology	-.36 (.33)	<b>.27 (.32)</b>	.28 (.43)
Intercept	2.07** (.38)	<b>.65* (.37)</b>	<b>.88* (.53)</b>
Adjusted R <sup>2</sup>	.28	.27	.26
N	366	355	183

\*\**p* < .05, \* *p* < .10 (two-tailed) <sup>a)</sup> deviation from sample mean.

Note: Cells show OLS coefficients with standard errors in parentheses. The regressions also include a dummy variable for respondents who did not answer the ideology question. **Bolded coefficients** are different from the same coefficient in the control group at *p* < .10 (two-tailed).

involvement are affected by how knowledgeable they feel. Their estimates of news exposure might be affected in a similar way unless the question draws attention to the actual size of the news audience and thereby corrects for an otherwise misleading inference rule.

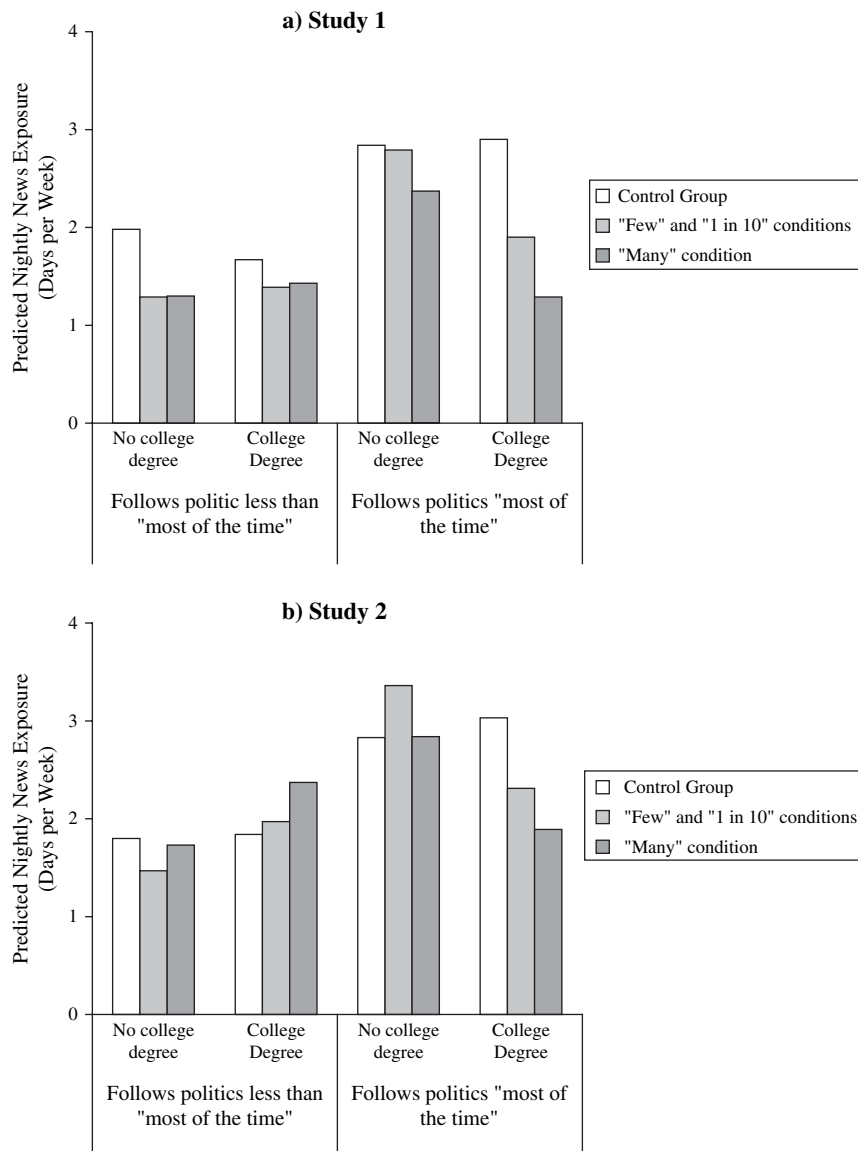
With remarkable consistency across the two studies, the anchor conditions indicate that network news attracts predominantly less educated but politically interested viewers. This finding squares with the reputation of television as an “easy” (Salomon 1984) news medium that attracts a less educated audience. College graduates with an interest in politics, on the other hand, look like frequent viewers of television

news on the traditional exposure measures, but may in fact get their news elsewhere.

### The Consequences of Reduced Overreporting for Conclusions about Political Behavior

The validity of inferences about the effects of news exposure depends on the validity of the exposure measure. To determine if measures of news exposure that produce less overreporting in the aggregate also change conclusions about the relationship between

FIGURE 3 Predicted Network News Exposure, by Education and Political Interest (Anchor Experiment)



Note: Bars are predicted values from regression models in Table 2 for a white, ideologically moderate, employed women (who gets some news from TV and has dialup Internet access).

news exposure and political knowledge, I use the exposure measures in the different experimental conditions as independent variables in a model of political knowledge. This cross-sectional design does not allow any causal inferences about learning from news exposure (see Neuman, Just, and Crigler (1992) for excellent experimental research on this topic), but it establishes how informed the self-reported news audience is. The dependent variable is an additive political knowledge index that ranges from 0 to 12. Appendix B lists the questions and describes the scoring.

Results in Table 3 show that self-reported news exposure in the control group is unrelated to knowledge when controls for demographic factors are included. This is consistent with past studies which have found weak relationships between self-reported news exposure and political knowledge when political interest, attention, or demographics are held constant (e.g., Chang and Krosnick 2002; Price and Zaller 1993). To assess the effect of news exposure in the anchor conditions, the model includes interactions between exposure and the effective experimental

**TABLE 3** The Effect of Self-Reported Network News Exposure on Political Knowledge, by Experimental Condition (Anchor Experiment, Study 2)

	DV: Political Knowledge (0–12)
“1 in 10”/ “Few Americans”	-.85** (.27)
“Many Americans”	-.63* (.33)
“Some watch . . . others don’t”	-.94** (.30)
Reported Network News Exposure	-.012 (.06)
Exposure × “1 in 10”/ “Few”	.13* (.08)
Exposure × “Many”	.18* (.10)
Exposure × “Some watch . . . ”	.22** (.09)
College Graduate	2.10** (.18)
Age <sup>a)</sup>	.03** (.005)
Female	-.92** (.15)
Minority	-1.20** (.18)
Employed	.02 (.17)
Internet connection speed	.84** (.17)
Intercept	3.94** (.44)
Adjusted R <sup>2</sup>	.27
N	1139

\*\* $p < .05$ , \* $p < .10$  (two-tailed) <sup>a)</sup> deviation from sample mean. Note: Cell entries are OLS coefficients and standard errors in parentheses. The model also includes indicator variables for several orthogonal experimental treatments related to the knowledge questions (see appendix).

conditions. The relationship in the *Some Watch . . . Others Don’t* condition is tested as well because it encourages comparison with other people and caused a marginal decline in reported exposure.

The way we measure exposure has noteworthy implications for the relationship between news exposure and political knowledge. When overreporting is reduced by an effective anchor, reports of network news exposure are significantly related to political knowledge. All three interactions between treatment and reported exposure are positive and statistically significant. Their magnitudes imply that the difference between a nonviewer and someone who watches network news every day amounts to about one point on the knowledge index—about the same as the gender gap in this analysis and half of the effect of a college degree.

Previous studies have compared the impact of different media exposure measures on factual information to assess their validity (e.g., Chang and Krosnick 2002; Price and Zaller 1993). By that standard, the stronger relationship between exposure and knowledge in the anchor conditions attests to better measurement. The standard is flawed, however, because the true relationship between exposure and factual information is unknown. It is not clear

*ex ante* if a stronger or a weaker relationship between the new exposure questions and knowledge would validate the exposure measures. Instead, the results in Table 3 illustrate the importance of measurement for conclusions about media effects. The null findings in the control group (and in many other survey-based studies) do *not* mean that network news viewers are as uninformed about politics as the rest of the population. Rather, they reflect flawed measurement of network news exposure. Better measures reveal a more informed network news audience.

## Conclusion

Overreporting of news exposure is a serious problem for survey research. Aggregate survey-based estimates of news exposure are several times higher than independent assessments that do not rely on self-reports. Worse, overreport bias is not constant across respondents (Prior, 2009). This study reports the first systematic examination of why many people overstate their exposure to television news. A set of experiments embedded in two different representative surveys of U.S. residents indicates that overreporting results from unrealistic demands on respondents’ memory, not their motivation to misrepresent or provide superficial answers. The primary cause of inflated self-reports is imperfect recall coupled with the use of flawed inference rules. Most respondents are incapable of recalling most or all episodes of news exposure, so they estimate their exposure, a strategy that tends to generate higher self-reports than enumeration (Brown and Sinclair 1999; Burton and Blair 1991). To lower reports of news exposure, surveys should offer respondents help with the estimation.

Survey methodologists have long known that behavioral self-reports can overtax respondents’ memory. In response to questions about frequent activities of low salience, “their answers are likely to be based on some fragmented recall and the applications of inference rules to compute a frequency estimate” (Schwarz 1999, 97). Commenting on inflated self-reports of news exposure, Price and Zaller characterize the problem less charitably: “Answers to survey questions asking [respondents] to report rates of media use may thus depend heavily upon guesswork” (1993, 136). This “guesswork” is not random, however. It matters what kind of inference rules respondents employ. Survey researchers have recommended providing population averages in order to encourage comparison with other people (Burton and Blair 1991; Schwarz et al. 1985). Implemented in

the *Anchor Experiment*, this strategy reduced overreporting significantly in two different studies.

Exposure questions that reduce aggregate overreporting revise several substantive conclusions about who watches the news and to what effect. Whereas more politically interested respondents report higher exposure in the traditional design, self-reports among respondents who are both interested and well-educated drop when the survey helps them with the estimation. Less educated but politically interested Americans emerge as the prime audience for network news. Results also suggest that network news exposure is more strongly related to political knowledge than traditional exposure questions indicate.

The rejection of the social desirability and satisficing hypotheses indicates that respondents' motivation—either to look good or to provide accurate answers—is not the main cause of overreporting. The *List Experiment*, a technique that allows respondents to report a behavior without revealing it directly and has been successfully used to demonstrate social desirability bias for many behaviors, did not generate lower self-reports. One of the treatments in the *Anchor Experiment* informed respondents that “many Americans” watch the news and thus increased the social pressure to report high exposure. Yet, like the other two anchors, this one, too, lowered self-reports. Two independent tests thus produced no support for the social desirability hypothesis.

Absence of social desirability pressures on self-reported news exposure is not necessarily inconsistent with studies that find social desirability bias in self-reported turnout (Belli et al. 1999; Bernstein, Chadha, and Montjoy 2001; Holbrook and Krosnick 2005; Presser 1990). There is no reason to assume that overreporting of turnout and news exposure must be caused by the same factors. Even with respect to turnout, however, some evidence hints at other sources of error. Comparing separate random samples interviewed in the three months after the 1996 presidential election, Belli et al. (1999) show that turnout overreporting increases with time—which they find difficult to reconcile with deliberate misreporting since “if anything, one would expect conscious lying would occur most often shortly following the election, when remembering one's voting behavior would be most certain” (1999, 105). Abelson, Loftus, and Greenwald (1992) find a decline in the accuracy of vote reports over time because some respondents become more likely to falsely report that they voted, while others become more likely to falsely report that they did not vote. The latter trend appears to be inconsistent with social desirability bias. Lower accuracy among both

voters and nonvoters suggests that memory error contributes to misreporting of turnout as well.

Overreporting of news exposure appears not to be caused by satisficing. The *15-Seconds Experiment* and the *Follow-Up Challenge*, designs previously employed to lower self-reports by encouraging greater cognitive effort (Burton and Blair 1991), did not affect responses. Respondents are doing their best to provide genuine estimates of their news exposure, but are frustrated by the difficult estimation task. Despite the absence of direct evidence for the satisficing hypothesis, further investigation is warranted because of a concern with self-administered online surveys that emerged in Study 2. Respondents who completed the exposure question quickly were not affected by the experimental manipulations. Fast responses might reflect easy estimation. Respondents who never watch network news (or always do) face a simpler task than irregular viewers—and may not require, or benefit from, help with the estimation. But the absence of experimental effects at fast response times could also indicate satisficing. In that case, effective questions should encourage both appropriate inference rules and sufficient effort.

Despite the clear recommendations that emerge from this study, it represents only the beginning of a research agenda to address the problem of inaccurate self-reports for media effects research. Most importantly, it could not directly validate people's true news exposure. The treatments in the *Anchor Experiment* disproportionately reduced reported exposure among respondents with low cognitive abilities and respondents with other primary news sources than television, but this is at best indirect evidence for the validity of the modified exposure measures. A comparison with unobtrusively recorded behavioral data could produce more definitive evidence. Secondly, this study only evaluated the effects of news exposure on one dependent variable in a cross-sectional context. Further tests are needed to establish how much our conclusions about media effects change when more accurate exposure measures are used. Despite these limitations, the stakes are clear: without more accurate self-reports of media exposure, survey research will not be able to examine media effects rigorously.

Even though one justification for experiments in political science is precisely a concern about the validity of self-reports (see, e.g., Ansolabehere, Iyengar, and Simon 1999), it is essential to assess exposure to understand the real-world relevance of media effects found in experiments. The explosion of media choice (e.g., Prior 2007) makes it more dubious to simply assume that experimental subjects will be exposed to

the treatment outside the experimental context. In the extreme, experimental effects may occur entirely among subjects who do not receive the message in the real world. Innovative experimental studies address this concern by building message selection into the design (e.g., Arceneaux and Johnson 2007; Iyengar and Hahn 2007). Even this innovation cannot establish the frequency of exposure (and thus the cumulative effect) and necessarily sacrifices realism. Survey-based measures of news exposure are needed in the future, so that survey research can contribute its share to methodological pluralism in the study of media effects. For this contribution to be meaningful, political scientists and communication scholars cannot ignore the egregious overreporting of news exposure. In this study, I have taken a modest first step by identifying causes of overreporting and devising survey procedures to reduce its impact.

### Acknowledgments

Data for Study 1 were collected through Time-Sharing Experiments for the Social Sciences (TESS), funded by the National Science Foundation. I am grateful to Diana Mutz at TESS and Heather Terhune and John Kennedy of the Center for Survey Research at the Indiana University for their help in implementing this project. Thanks to Poom Nukulkiy and Mike Dennis at Knowledge Networks for programming and carrying out Study 2. For helpful comments on earlier versions of this paper, I would like to thank Larry Bartels, Kelly Garrett, Marty Gilens, Allyson Holbrook, Jon Krosnick, Tali Mendelberg, Russ Neuman, and Talia Stroud. Thanks to Michael Noveck for his research assistance.

### Appendix A Randomization Check

To verify randomization, I regressed the random variables for each of the experiments on a set of demographic and attitudinal variables (education, age, gender, race, employment status, marital status, political interest, ideology, week of the interview (Study 1), and speed of Internet connection (Study 2)). In all four cases, the null model cannot be rejected, indicating successful randomization. Eight out of 131 coefficients are significant at the 95% level, which is close to the 6.6 expected by chance. To account for chance differences in demographics or attitudes, all analyses report results with controls.

TABLE A1

Experiment	# of conditions	Likelihood ratio test against null model	Number of coefficients with $p < .05$
<i>Study 1</i>			
List Exp.	3	$\chi^2 (18) = 19.7,$ $p = .35$	1 out of 18
15 Seconds Exp.	2	$\chi^2 (9) = 8.2,$ $p = .51$	0 out of 9
Anchor Exp.	4	$\chi^2 (27) = 34.3,$ $p = .16$	2 out of 27
<i>Study 2</i>			
Anchor Exp.	7	$\chi^2 (72) = 85.5,$ $p = .13$	5 out of 77

Note: Models are multinomial logit models if the number of conditions exceeds 2, binary logit models otherwise.

### Appendix B Measuring Political Knowledge

The political knowledge index was created by summing correct answers to the following questions:

1. Who is the current secretary of defense? (Donald Rumsfeld, John Ashcroft, Robert Gates, Colin Powell)
2. Who is the current Speaker of the U.S. House of Representatives? (Nancy Pelosi, Dana Perino, Barbara Boxer, Elizabeth Edwards)
3. Who is the Chief Justice on the U.S. Supreme Court? (William Rehnquist, John Roberts, Antonin Scalia, Samuel Alito)
4. Who is the President of Iran? (Mahmoud Ahmadi-najad, Nouri al-Maliki, Hamid Karzai, Pervez Musharraf)
5. What office is currently held by Condoleezza ("Condi") Rice [the person shown in this picture]? (Director of the Central Intelligence Agency, National Security Adviser, Secretary of State, White House Chief of Staff)
6. What position is currently held by Ben Bernanke [the person shown in this picture]? (Treasury Secretary, Chairman of the Federal Reserve, Secretary of Homeland Security, Chairman of the National Economic Council)
7. What position is currently held by Michael Mukasey [the person shown in this picture]? (U.S. Attorney General, President's General Counsel, Senate Majority Leader, Supreme Court Justice)
8. What position is currently held by Nicolas Sarkozy [the person shown in this picture]? (President of France, Foreign Minister of the European Union, Prime Minister of Turkey, UN Secretary General)

- 9–10. Of these four politicians, who is the most conservative? And who is the most liberal? (Newt Gingrich, Arnold Schwarzenegger, Joe Lieberman, John Kerry)
11. Of the 100 members of the U.S. Senate, how many are members of the Democratic Party?

All questions except #11 were multiple-choice (response options are shown in parentheses). Question 11 was scored as follows: 2 points for “51”, 1.5 points for 52–55, 1 point for 49, 50, and 56–59, and 0.5 points for 40–48 and 60–100. This coding scheme rewards partial knowledge (e.g., which party has a majority and whether the majority is veto-proof). The resulting index ranges from 0 to 12 with a mean of 5.5 and a standard deviation of 3.0.

As part of a separate study, the response options for questions 1–10 were randomly shown either as photographs, names, or photographs and names. Also based on random assignment, half the questions included an explicit “Don’t Know” option. Random assignment to these conditions was independent of the randomization in the Anchor Experiment. The coefficients in Table 3 therefore indicate average effects across conditions. In order to adjust for differences in the de-facto difficulty of knowledge questions as a function of experimental conditions (e.g., better performance without a “Don’t Know” option), the model in Table 3 includes dummy variables for all independent randomizations.

*Manuscript submitted 13 August 2007*

*Manuscript accepted for publication 5 October 2008*

## References

- Abelson, Robert P., Elizabeth F. Loftus, and Anthony Greenwald. 1992. “Attempts to Improve the Accuracy of Self-Reports of Voting.” In *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*, ed. Judith M. Tanur. New York: Russell Sage Foundation, 138–53.
- Ansolabehere, Stephen, Shanto Iyengar, and Adam Simon. 1999. “Replicating Experiments Using Aggregate and Survey Data: The Case of Negative Advertising and Turnout.” *American Political Science Review* 93 (4): 901–10.
- Arceneaux, Kevin, and Martin Johnson. 2007. “Channel Surfing: Does Choice Reduce Videomalaise?” Presented at the Annual Meeting of the Midwest Political Science Association.
- Bartels, Larry M. 1993. “Messages Received: The Political Impact of Media Exposure.” *American Political Science Review* 87: 267–85.
- Belli, Robert F., Michael W. Traugott, and Matthew N. Beckmann. 2001. “What Leads to Voting Overreports? Contrasts of Overreporters to Validated Voters and Admitted Nonvoters in the American National Election Studies.” *Journal of Official Statistics* 17 (4): 479–98.
- Belli, Robert F., Michael W. Traugott, Margaret Young, and Katherine A. McGonagle. 1999. “Reducing Vote Overreporting in Surveys: Social Desirability, Memory Failure, and Source Monitoring.” *Public Opinion Quarterly* 63 (1): 90–108.
- Bernstein, Robert, Anita Chadha, and Robert Montjoy. 2001. “Overreporting Voting: Why It Happens and Why It Matters.” *Public Opinion Quarterly* 65 (1): 22–44.
- Bishop, George F., Robert W. Oldendick, and Alfred Tuchfarber. 1984. “What Must My Interest in Politics Be If I Just Told You ‘I Don’t Know?’” *Public Opinion Quarterly* 48 (2): 510–19.
- Brady, Henry E. 1986. “The Perils of Survey Research: Interpersonally Incomparable Responses.” *Political Methodology* 11 (3–4): 269–91.
- Brown, Norman R., and Robert C. Sinclair. 1999. “Estimating Number of Lifetime Sexual Partners: Men and Women Do It Differently.” *Journal of Sex Research* 36 (3): 292–97.
- Burton, Scot, and Edward Blair. 1991. “Task Conditions, Response Formulation Processes, and Response Accuracy for Behavioral Frequency Questions in Surveys.” *Public Opinion Quarterly* 55 (1): 50–79.
- Chang, LinChiat, and Jon A. Krosnick. 2002. “Measuring the Frequency of Regular Behaviors: Comparing the ‘Typical Week’ to the ‘Past Week.’” *Sociological Methodology* 33: 55–80.
- Conrad, Frederick G., Norman R. Brown, and Erin R. Cashman. 1998. “Strategies for Estimating Behavioral Frequency in Survey Interviews.” *Memory* 6 (4): 339–66.
- Gerber, Alan S., James G. Gimpel, Donald P. Green, and Daron R. Shaw. 2007. “The Influence of Television and Radio Advertising on Candidate Evaluations: Results from a Large Scale Randomized Experiment.” Presented at the Annual Meeting of the Midwest Political Science Association.
- Gerber, Alan S., and Donald P. Green. 2000. “The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment.” *American Political Science Review* 94 (3): 653–63.
- Gilens, Martin, Paul M. Sniderman, and James H. Kuklinski. 1998. “Affirmative Action and the Politics of Realignment.” *British Journal of Political Science* 28 (1): 159–83.
- Holbrook, Allyson L., and Jon A. Krosnick. 2005. *Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count and Randomized Response Technique*. University of Illinois at Chicago. Unpublished manuscript.
- Iyengar, Shanto. 1991. *Is Anyone Responsible? How Television Frames Political Issues*. Chicago: University of Chicago Press.
- Iyengar, Shanto, and Kyu S. Hahn. 2007. “Red Media, Blue Media: Evidence of Ideological Polarization in Media Use.” Presented at the Annual Meeting of the International Communication Association, San Francisco.
- Iyengar, Shanto, and Donald R. Kinder. 1987. *News That Matters: Television and American Opinion, American Politics and Political Economy*. Chicago: University of Chicago Press.
- Johnston, Richard, Michael G. Hagen, and Kathleen Hall Jamieson. 2004. *The 2000 Presidential Election and the Foundations of Party Politics*. Cambridge: Cambridge University Press.
- Katz, Jonathan N. 2000. *Correcting for Survey Misreports Using Auxiliary Information*. Pasadena: California Institute of Technology.
- Krosnick, Jon A. 1991. “Response Strategies for Coping with Cognitive Demands of Attitude Measures in Surveys.” *Applied Cognitive Psychology* 5: 213–36.

- Kuklinski, James H., Michael D. Cobb, and Martin Gilens. 1997. "Racial Attitudes and the 'New South'." *Journal of Politics* 59 (2): 323–49.
- Lazarsfeld, Paul F., and Bernard Berelson. 1944. *The People's Choice; How the Voter Makes up His Mind in a Presidential Campaign*. New York: Columbia University Press.
- Miller, Judith D. 1984. "A New Survey Technique for Studying Deviant Behavior." Ph.D. diss., George Washington University.
- Neuman, W. Russell, Marion R. Just, and Ann Crigler, N. 1992. *Common Knowledge: News and the Construction of Political Meaning*. Chicago: University of Chicago Press.
- Patterson, Thomas E., and Robert D. McClure. 1976. *The Unseeing Eye: The Myth of Television Power in National Politics*. New York: Putnam.
- Presser, Stanley. 1984. "Is Inaccuracy on Factual Survey Items Item-Specific or Respondent-Specific?" *Public Opinion Quarterly* 48 (1): 344–55.
- Presser, Stanley. 1990. "Can Changes in Context Reduce Vote Overreporting in Surveys?" *Public Opinion Quarterly* 54: 586–93.
- Presser, Stanley, and Michael Traugott. 1992. "Little White Lies and Social Science Models: Correlated Response Errors in a Panel Study of Voting." *Public Opinion Quarterly* 56 (1): 77–86.
- Price, Vincent. 1993. "The Impact of Varying Reference Periods in Survey Questions about Media Use." *Journalism Quarterly* 70(3): 615–27.
- Price, Vincent, and John Zaller. 1993. "Who Gets the News? Alternative Measures of News Reception and Their Implications for Research." *Public Opinion Quarterly* 57 (2): 133–64.
- Prior, Markus. 2007. *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. New York: Cambridge University Press.
- Prior, Markus. 2009. "The Immensely Inflated News Audience: Assessing Bias in Self-Reported News Exposure." *Public Opinion Quarterly* 73 (1): 130–143.
- Romer, Daniel, Kate Kenski, Paul Waldman, Christopher Adasiewicz, and Kathleen Hall Jamieson. 2004. *Capturing Campaign Dynamics*. New York: Oxford University Press.
- Salomon, Gavriel. 1984. "Television Is 'Easy' and Print Is 'Tough': The Differential Investment of Mental Effort in Learning as a Function of Perceptions and Attributions." *Journal of Educational Psychology* 76 (4): 647–58.
- Schwarz, Norbert. 1999. "Self-Reports: How the Questions Shape the Answers." *American Psychologist* 54 (2): 93–105.
- Schwarz, Norbert, Hans-J. Hippler, Brigitte Deutsch, and Fritz Strack. 1985. "Response Scales: Effects of Category Range on Reported Behavior and Comparative Judgments." *Public Opinion Quarterly* 49 (3): 388–95.
- Schwarz, Norbert, and Daphna Oyserman. 2001. "Asking Questions about Behavior: Cognition, Communication and Questionnaire Construction." *American Journal of Evaluation* 22 (2): 127–60.
- Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80 (2): 613–24.
- Tourangeau, Roger, Lance J. Rips, and Kenneth, Rasinski. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.
- Tversky, Amos, and Daniel Kahneman. 1973. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology* (5): 207–32.
- Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1995. *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge, MA: Harvard University Press.

Markus Prior is assistant professor of politics and public affairs, Princeton University, Princeton, NJ 08544.