

 aptara <small>The Content Transformation Company</small>	AJPS	ajps'600	Dispatch: April 28, 2012	CE: N/A
	Journal	MSP No.	No. of pages: 12	PE: Martha

# Televised Exposure to Politics: New Measures for a Fragmented Media Environment

**Susanna Dilliplane** University of Pennsylvania  
**Seth K. Goldman** University of Pennsylvania  
**Diana C. Mutz** University of Pennsylvania

*For many research purposes, scholars need reliable and valid survey measures of the extent to which people have been exposed to various kinds of political content in mass media. Nonetheless, good measures of media exposure, and of exposure to political television in particular, have proven elusive. Increasingly fragmented audiences for political television have only made this problem more severe. To address these concerns, we propose a new way of measuring exposure to political television, and evaluate its reliability and predictive validity using three waves of nationally representative panel data collected during the 2008 presidential campaign. We find that people can reliably report the specific television programs they watch regularly, and that these measures predict change over time in knowledge of candidate issue positions, a much higher standard of predictive validity than any other measure has met to date.*

Measuring political media exposure is a perennial source of difficulty for scholars. Problems with measurement have serious implications for the state of research involving media exposure and have led to “one of the most notable embarrassments of modern social science” (Bartels 1993: 267). To study many contemporary political phenomena, researchers need reliable and valid measures of the extent to which people have been exposed to political content in mass media, yet measures of this kind have proven elusive. In this study we propose a new measure of exposure to political television and draw on a large representative panel sample in order to test the reliability and predictive validity of this approach.<sup>1</sup>

We begin by reviewing the limitations of traditional measures of televised exposure to politics. We then analyze the reliability and validity of an alternative measure of exposure, the program list technique, assessing its utility for a variety of purposes common among scholars interested in political communication. Finally, we discuss the advantages and disadvantages of these measures for

understanding how people are influenced by their media environments.

Overall, we find that these new measures do a superior job of capturing exposure to the varied political content available in today’s fragmented media environment. Notably, these measures exhibit high levels of true-score reliability and predict within-person change over time in campaign-specific knowledge and participation. Moreover, although our focus is on television, this approach has the potential to be used in measuring exposure to other media as well.

## Traditional Approaches

In the context of representative population surveys, the traditional approach to measuring political television exposure has been to ask respondents to self-report how many days per week or hours per day they spend watching news on television in an average week or in the past

---

Susanna Dilliplane Annenberg School for Communication, University of Pennsylvania 3620 Walnut Street Philadelphia, PA 19104 sdilliplane@asc.upenn.edu. Seth K. Goldman Annenberg School for Communication, University of Pennsylvania, 3620 Walnut Street, Philadelphia, PA 19104, sgoldman@asc.upenn.edu. Diana C. Mutz University of Pennsylvania, Department of Political Science, 208 S. 37th Street, Room 217, Philadelphia, PA 19104-6215, dmutz@asc.upenn.edu.

<sup>1</sup>The data used in the study are publicly available through the Annenberg Public Policy Center’s website at <http://www.annenbergpublicpolicycenter.org/NewsDetails.aspx?myId=263>.

*American Journal of Political Science*, Vol. 0, No. 0, xxx 2012, Pp. 1–13

©2012, Midwest Political Science Association

DOI: 10.1111/j.1540-5907.2012.00600.x

week. These approaches have several problems in today's fragmented media environment. Due to huge increases in the types of available political programs, even the most motivated respondents are likely to disagree on what counts as "news," "programs about the campaign," "political television," or "news and public affairs." Today, political content appears on a wide variety of programs, including daytime and evening programs, general interest talk shows, opinion programs, satire, and more (Williams and Delli Carpini 2011). Unfortunately, when respondents report watching an hour of "political television," researchers have no way of knowing how much or what kind of content viewers have seen, and thus cannot test hypotheses connecting the content of exposure to specific consequences.

Traditional exposure measures have also been heavily criticized for lacking reliability, validity, or both as a result of the excessive cognitive burden they place on survey respondents (e.g., Price and Zaller 1993). A reliable estimate of hours (or even days) watching politics (or news) in a typical week could involve a prolonged multi-stage recall process (Schwarz and Oyserman 2001). Given the tendency to answer quickly, respondents likely rely on shortcuts to come up with off-the-cuff estimates, thus reducing exposure measures to little more than self-assessed levels of political interest (Prior 2009b). In the next sections, we consider these criticisms more closely, assessing what is known about the reliability and validity of traditional exposure measures.

## Reliability

How reliable are traditional survey questions about political television exposure? The received wisdom is that media measures in general have low reliability. In reality, the extremely small literature on this topic suggests contradictory conclusions because studies often refer to entirely different types of reliability that are not comparable, nor equally important. Positive assessments of reliability generally come from cross-sectional studies that assess reliability with a measure of internal (inter-item) consistency, such as Cronbach's alpha. Several questions, all designed to tap television exposure to the campaign, albeit in slightly different ways, are combined into a single index of political television exposure. Scholars often claim high levels of reliability for measures of exposure using this approach.<sup>2</sup>

<sup>2</sup>For example, the 1996 American National Election Study included four items: (a) number of days watching local news, (b) number of days watching national TV news, (c) amount of attention to cam-

However, many concerned about the low reliability of media exposure measures typically have another conception of reliability in mind; that is, the relationship between the score that scholars observe and the underlying "true score" that is unobserved. True-score reliabilities assume that there is an underlying hidden variable that one is trying to measure via a particular question or questions (Lord and Novick 1968, p. 61). Even though the hidden variable cannot be directly observed, it can be inferred when the measures are repeated in three or more waves of panel data (Heise 1969, Wiley and Wiley 1970). Unlike reliability assessments such as Cronbach's alpha, true-score reliabilities separate measurement error from change in underlying true-scores, thus providing a measure of reliability that is *independent of stability* (Heise 1969).

Moreover, true-score reliability is a much higher standard than internal consistency because two measures that are both close to a given true score (i.e., valid) must also be close to each other (i.e., reliable) (see Zaller 2002). Conversely, two measures that are close to each other (i.e., reliable) need not necessarily be close to the true score (i.e., valid). So although internal consistency and true-score reliabilities have the same general goal of better measurement, they evaluate fundamentally different things and cannot be directly compared.

Unfortunately, true-score reliabilities have rarely been calculated in the literature on media exposure because this assessment requires three or more waves of panel data, which are seldom available. We located only one published study examining the true-score reliability of political television exposure. Using a three-wave sample from the 1980 American National Election Study Panel, Bartels (1993) found a true-score reliability of .75 for the question, "How often do you watch the national network news on early evening TV – every evening, 3 or 4 times a week, once or twice a week, or less often?"<sup>3</sup> Although Bartels' findings are often cited as evidence of the *low* reliability of media exposure measures, he actually found lower levels of reliability for common measures of other key political constructs, such as presidential job approval and issue preferences. Nonetheless, individuals' capacity to assess their exposure to political content on

paigned stories on local TV, and (d) amount of attention to campaign stories on national TV. The four items created a scale with a Cronbach's alpha of .79, which sounds like a relatively high reliability to most scholars (Zaller 2002). Other efforts have reported similarly high alphas based on multi-item scales collected at a single point in time (e.g., Price 1993).

<sup>3</sup>A study using a three-wave panel reported that television news exposure was highly stable over a two-year period but did not report true-score reliabilities (Chaffee and Schluender (1986).

television has probably declined since 1980 due to the numerous and diverse sources in the contemporary media environment.

## Validity

Concerns about the low reliability of media exposure measures may be exaggerated, but complaints about their validity are even more widespread. As Zaller (2002) suggests, it is difficult to take seriously a measure of news exposure that has a very weak relationship with what ought to be its consequence – that is, political knowledge. And even when exposure does correlate with knowledge, cross-sectional survey data have made it difficult to make firm claims about causal relationships.

Evidence of low criterion validity has led some scholars to advocate abandoning measures of exposure altogether in favor of a textbook political knowledge index (Price and Zaller 1993). Although useful as a surrogate for some purposes, a textbook political knowledge index usually taps the kind of civics knowledge gained through schooling rather than through ongoing exposure to political media. Even a measure of current events knowledge would still conflate what people are exposed to with what people retain from that exposure (Althaus and Tewksbury 2007, 2009). Evidence of online information processing suggests that people are exposed to a great deal of content that influences their views but is not necessarily recalled (Lodge and Stroh 1993).

The gold standard for assessing the validity of media exposure is how well a measure predicts political knowledge *gain*, but in practice, tests of validity are often reduced to how well a measure correlates with political knowledge at a single point in time. Such cross-sectional analyses obviously cannot capture over time *changes* in knowledge. Even when cross-sectional analyses include a variety of control variables, they remain vulnerable to spuriousness from unmeasured and/or unobservable confounders, not to mention reverse causality. Simply put, evaluation of how well exposure measures predict knowledge *gain* requires panel data.

Panel data are required to assess both true-score reliability and to conduct the best analysis of predictive validity. Accordingly, in this study we use panel data to evaluate the reliability and validity of a new approach to measuring political television exposure. Thus, the analyses presented below provide tests of reliability and validity that are both appropriate and rigorous, as well as an improvement over most previous assessments of reliability and validity.

## Research Design

We use data from the 2008 National Annenberg Election Study (NAES), a five-wave panel collected online by Knowledge Networks of Menlo Park, CA. A large, nationally representative probability sample of respondents was initially recruited through random digit dialing, and then upon empanelment was given Internet access if needed (either via personal computer or WebTV). The study was fielded beginning October of 2007, before the primary season began, through the inauguration in January of 2009, for a span of over 15 months.<sup>4</sup> Just over 12,000 people completed all three of the waves that included the new television exposure measures, Waves 2, 4, and 5. Respondents received no other political surveys throughout the period of the study in order to minimize the possibility of attrition based on political interest. Although some attrition occurred, the demographic composition of the re-interviewed sample differed little from that of the sample of respondents interviewed in wave 1, consistent with prior studies of attrition using Internet panels (see Chang and Krosnick 2009).<sup>5</sup>

## The Program List Technique

Because the NAES panel was executed via Internet, it was possible to present respondents with lists of individual television programs and to ask them to check off the ones they watched regularly.<sup>6</sup> Specific programs were chosen for inclusion based on Nielsen ratings in 2007, when the panel began. The lists included heavily political programs as well as those that only occasionally touched on this

<sup>4</sup>Wave 1 took place from October 2, 2007 – December 31, 2007; Wave 2 from January 1, 2008 – March 31, 2008, Wave 3 from April 2, 2008 – August 28, 2008, Wave 4 from August 29, 2008 – November 4, 2008, and Wave 5 from November 5, 2008 – January 31, 2009.

<sup>5</sup>Using panel weights to correct demographic representativeness did not change the results.

<sup>6</sup>Respondents were first asked, "From which of the following sources have you heard anything about the presidential campaign?" Respondents checked all of the categories (presented in randomized order) that applied to them: (1) television news programs (morning or evening); (2) newspapers; (3) television talk shows, public affairs or news analysis programs; (4) Internet sites, chat rooms or blogs; (5) radio news or radio talk shows; (6) news magazines; and (7) have not heard anything about the presidential campaign. Respondents who selected categories (1) or (3) were then shown a screen with a list of TV programs asking, "Which of the following programs do you watch regularly on television? Please check any that you watch at least once a month." Later screens were sprinkled throughout the survey in random order to avoid boredom from checking off programs from multiple screens in a row.

topic (see Table A1, Appendix A for a full list). A total of 49 politically-relevant television programs appeared on four screens sprinkled throughout the survey. Each screen included 12–15 programs, along with some fiction-based entertainment programs (e.g., *Big Love*)<sup>7</sup> and a “None of the above” option – all with clickable check boxes beside them.

The program list approach advances two goals.<sup>8</sup> First, it decreases the cognitive demands placed on respondents. When viewers watch TV, they likely think in terms of the programs they watch rather than in time units (e.g., hours or days) or in terms of researcher-defined categories of political programs. Thus, it should be easier for respondents to recognize and report regular viewership of specific programs than to mentally tally summary amounts of time devoted to abstract categories of programming.

Second, this approach promotes content validity by more accurately incorporating the relevant domain of exposure. For the many reasons outlined above, merely asking about exposure to “national news” is losing favor in political surveys. For instance, the 2004 European Election Study (EES) included open-ended questions asking respondents to name up to five channels or television news programs that they watch regularly, and the 2009 EES asked respondents how many days in a typical week they watch each of two specific programs. Other surveys such as the NAES rolling cross-sectional telephone surveys in 2004 and 2008 have likewise asked about viewership of specific programs. The program list technique has the virtue of listing all widely viewed programs with political content, thus eliminating the problem of having each respondent individually decide which programs should count as “political” programs.

In addition, the program list approach encompasses real-time viewing as well as delayed viewing, and exposure via a television, computer screen, or handheld device. In this way, it incorporates the multiple modes of viewing currently possible, and anticipates the future when mode of delivery is largely independent of content.

We restricted our measures to programs viewed regularly because the kinds of accumulative (as opposed to single exposure) effects of media that most scholars study

<sup>7</sup>The reason for including apolitical programs was to retain the interest and attention of more apolitical respondents who might otherwise not read through the entire list.

<sup>8</sup>Though new to the study of political communication, a program list-based approach to measuring media exposure has been used in other domains of research, such as studies of exposure to violence on TV (e.g., Huesmann, Lagerspetz, and Eron 1984; Huesmann, Moise-Titus, Podolski, and Eron 2003).

should register primarily as a function of regular viewing rather than an isolated chance exposure. And although what counts as “regular” might be different for a show that airs once a week versus one that airs every night, by defining this for respondents as “at least once a month” for all programs, we preserved a degree of consistency for ease of answering.

Using the program lists, we constructed four indicators of political television exposure. *Total Number of Political TV Programs* is the sum of politically relevant programs each respondent checked off. This indicator is admittedly crude. Some of these programs air daily while others air only once a week; some are hour-long programs, whereas others last thirty minutes; and some are heavily focused on election coverage, whereas others feature campaign-related content only occasionally.

A more refined indicator, *Total TV Weighted by Days per Week*, relies on the same raw information, but weights each program by the number of days per week it aired. For example, regular viewing of a program airing 5 days per week is weighted 5 times more than regular viewing of a weekly program. A third indicator, *Total TV Weighted by Minutes per Week*, uses a similar strategy, but is weighted by the total number of minutes per week that a program aired.

A third strategy oriented toward refining this indicator relies on assessments of the extent to which each program included campaign content. Two independent coders categorized each political program’s content as heavily oriented toward campaigns and politics (3), somewhat oriented toward campaigns and politics (2), or only occasionally oriented toward campaigns and politics (1). These categories proved reliable, with 92 percent agreement between coders.<sup>9</sup> The programs were weighted by this information to create the final indicator: *TV Weighted by Level of Campaign Content*. To promote rough comparability of scales, each was rescaled to 0–1.<sup>10</sup>

## Results

First, we evaluate the true-score reliabilities of these measures using the Heise (1969) method for three-wave panel

<sup>9</sup>Discrepancies were resolved by discussion and additional viewing.

<sup>10</sup>One cannot, technically speaking, compare coefficients for indicators that do not use identical scales (King 1986). Of course, it is unclear even with time-based measures what it means to compare an hour of television viewing with an hour of newspaper reading. However, by rescaling the indicators to the same range, we make it easier for readers to avoid being misled by apparent differences in the size of coefficients that are due entirely to scales.

**TABLE 1 True-Score Reliability of Political Television Exposure Measures**

	Reliability
<i>Total Number of Political TV Programs</i>	.83
<i>Total TV in Days per Week</i>	.81
<i>Total TV in Minutes per Week</i>	.79
<i>Total TV Weighted by Level of Campaign Content</i>	.84
<i>Individual Political TV Programs (averaged across all 49 programs)</i>	.88
<i>Hours watched TV last night between 6 and 11</i>	.52
<i>How often do you watch national network news? (from Bartels 1993)</i>	.75

*Note:* The table presents Heise reliability estimates derived by taking the product of the raw correlations between waves 2 and 4 and between waves 4 and 5, and then dividing by the raw correlation between waves 2 and 5 (for the raw correlations, see Table A2 in Appendix A). The sample size for all of the calculations is 12,081, which includes all respondents who answered the media exposure measures in waves 2, 4, and 5.

data.<sup>11</sup> Second, we provide tests of predictive validity using fixed effects regression to examine whether political TV viewing predicts within-person *change* in levels of knowledge about candidate issue positions and campaign participation. And third, we provide tests of discriminant validity focusing on visual candidate knowledge (see Appendix B for details on question-wording, coding, and reliability of the outcome measures).

### Reliability

Table 1 reports the true-score reliabilities of the exposure indicators using the Heise (1969) method, which estimates reliability *independent of* change (or stability) in the underlying true-scores.<sup>12</sup> As shown in Table 1, the four indicators of political TV exposure are roughly equally reliable, with *Total TV Weighted by Level of Campaign Content* exhibiting the highest true-score reliability. Although exposure to individual TV programs is only occasionally

<sup>11</sup>This method has the advantage of simplicity and transparency, offering just one estimate of reliability for any given three-wave panel. Its disadvantage is the assumption that the reliability of the measured scores remains stable over time.

<sup>12</sup>This method takes the product of the correlation between waves 2 and 4 and between waves 4 and 5, and then divides by the correlation between waves 2 and 5. Wiley and Wiley (1970) proposed an alternative to the Heise method based on structural equation modeling, but it relies on additional assumptions to support the structural equation models. As a result, it is less robust than the Heise approach.

of interest, it is worth noting that respondents reliably identified watching individual TV programs as well. As shown in the bottom row in Table 1, the average true-score reliability across all 49 political programs is .88. Overall, these true-score reliabilities are as good as or better than the one previously reported for a traditional measure of political television exposure (Bartels 1993).

Although stability is not a desired quality of measurement, we also find that viewing political television is a highly stable behavior, after correcting for reliability. The between-wave stability coefficients for all of the measures constructed are above .9, indicating that watching political television is a highly habitual behavior. On the one hand, researchers relying on measures of television viewing from a single point in time may find this very reassuring because it means that a cross-sectional measure is not likely to be all that different from one that was measured a year earlier. On the other hand, this high level of stability means that there is little within-person variance over time, thus making it more difficult to relate the effects of change in exposure to changes in knowledge or behavior. Nonetheless, in situations in which new information is introduced into the media environment, as is the case with candidate-related information during elections, there is plenty of reason to expect that individuals with higher levels of stable, ongoing exposure to political television will, as a result, increase more in knowledge than those with lower stable exposure.

### Predictive Validity

The central remaining task is to assess the validity of these measures. A common critique of self-report measures is that they represent how respondents think about themselves rather than their amounts of actual exposure. Respondents might check off many political programs from this list, not because they watch so many programs, but because they think of themselves as political people. In our analyses, however, respondents did not simply report overall amounts of political viewing; they reliably reported watching the *same* political programs. Across all three waves, respondents were highly consistent in reporting whether they watched a given program, not just in reporting the same total number of programs (see Table A1). On average, across all 49 programs, 87 percent of respondents consistently reported watching (or not watching) a given program in all three waves. The fact that many people select the *same* programs in three waves spanning almost a full year makes it seem unlikely that they are just selecting a number of programs in order

**TABLE 2 The Impact of Within-Person Change in Political TV Exposure on Within-Person Change in Knowledge of Candidate Issue Positions (Fixed Effects Panel Analysis)**

Total Number of Political TV Programs	.54*** (.15)			
Total TV Weighted by Level of Campaign Content		.60*** (.14)		
Total TV in Days per Week			.46** (.13)	
Total TV in Minutes per Week				.36** (.11)
Wave	1.14*** (.01)	1.14*** (.01)	1.14*** (.01)	1.14*** (.01)
Constant	.37*** (.02)	.36*** (.02)	.38*** (.02)	.39*** (.02)
Sample Size	10,986	10,986	10,986	10,986

Note: Unstandardized coefficients from linear fixed effects regression models are shown, with standard errors in parentheses. Candidate knowledge ranges from 0–4. All of the independent variables range from 0–1. \*\*\*p < .001, \*\*p < .01, \*p < .05

to appear more politically involved and consistent with their self-image.

A more rigorous test than face or content validity is predictive validity – i.e., whether political television exposure significantly predicts over time *change* in relevant outcomes. For this evaluation of construct validity, we use the most widely accepted outcome of exposure, political knowledge, as well as a second outcome sometimes used to validate media exposure measures, campaign participation.

We capitalize on repeated measures to evaluate whether political TV exposure can significantly predict over-time change in *Knowledge of Candidate Issue Positions* and *Campaign Participation*.<sup>13</sup> We find the same results using both outcome measures, though for the sake of brevity, we present only the findings for knowledge gain (see Appendix C for the findings using campaign participation).

For each analysis of change over time, we use fixed effects models of within-person change. The advantages of this approach for disentangling preexisting individual differences from actual change are tremendous. With panel data, fixed effects regression uses only within-person variance, comparing each individual to him/herself at an earlier point in time. As a result, the constant effects of stable characteristics, whether observed or unobserved, cannot produce spurious associations between political television exposure and the dependent variables (Allison 2009). Variables such as education, income, age, ongoing

political interest, party affiliation and the like all fall out of these models because they are constants over time. By contrast, “in individual-level cross-sectional studies, differences in opinions between those exposed to the media and those who remain unexposed may simply reflect pre-existing differences between the two groups in political attitudes or characteristics” (Bartels 1993: 267). In addition, by including a dummy variable for wave in each equation, we efficiently capture the average effect of all other time-varying influences. For example, to the extent that interest rises across the board during the campaign, the wave variable captures this change. This technique arguably provides the most stringent causal test possible outside of an experimental setting (Allison 2009).

We first evaluate whether *change* in political TV exposure significantly predicts *change* in *Knowledge of Candidate Issue Positions*. Table 2 shows four fixed effects regression models, one for each of the exposure indicators. Within-person increases in each of the indicators significantly predict within-person increases in knowledge. Importantly, Table 2 demonstrates that over time increases in exposure to political television significantly predict knowledge *gain*, not simply knowledge *level*. *Total TV Weighted by Level of Campaign Content* (.60) and *Total Number of Political TV Programs* (.54) are the strongest predictors. As indicated by the significant *Wave* coefficients, knowledge increased over time during the campaign for other reasons as well, just as one would expect.

Although these results provide strong causal evidence, the limited amount of within-person variation in exposure over time may have resulted in unreliable estimates of effect size (Allison 2009). It is possible that we

<sup>13</sup>Knowledge levels increased significantly, albeit modestly, during the campaign. Mean change between wave 1 and 4 was 1.14 on a scale of 0–4; for participation, mean change between wave 3 and 5 was .22 on a scale of 0–3.

**TABLE 3 The Impact of *Stable* Levels of Political TV Exposure on Within-Person Change in Knowledge of Candidate Issue Positions (Fixed Effects Panel Analysis)**

Total Number of Political TV Programs X Wave	1.06*** (.10)			
Total TV Weighted by Level of Campaign Content X Wave		1.13*** (.10)		
Total TV in Days per Week X Wave			1.02*** (.10)	
Total TV in Minutes per Week X Wave				.94*** (.09)
Wave	1.00*** (.02)	.99*** (.02)	1.00*** (.02)	1.01*** (.02)
Constant	.44*** (.01)	.44*** (.01)	.44*** (.01)	.44*** (.01)
Sample Size	10,986	10,986	10,986	10,986

Note: Unstandardized coefficients from linear fixed effects regression models are shown, with standard errors in parentheses. Candidate knowledge ranges from 0–4. All of the independent variables range from 0–1. \*\*\*p < .001, \*\*p < .01, \*p < .05

are underestimating the impact of exposure by ignoring the potential effects of *stable* levels of exposure, which the models in Table 2 would not capture. Because campaign coverage changes the media environment during an election, individuals can be exposed to an influx of new information about the candidates as a result of changes in *program content*, not just changes in individual viewing habits. In this way, stable exposure levels may be expected to produce knowledge gain over time. Individuals with high levels of political television exposure could reasonably be expected to show larger gains in candidate issue knowledge than those with low levels of exposure.

To examine this possibility, we conduct an additional test of predictive validity, this time treating political TV exposure as a stable individual characteristic. For each indicator, we calculate the average level of exposure across the three waves, and then use this stable variable to predict *change* in the dependent variables. Although fixed effects models automatically control for the *constant* effects of all stable characteristics — thus making it impossible to include a stable variable in the model — we can use the interaction between political TV exposure and Wave to investigate whether the rate of increase in knowledge varied by level of stable political television exposure. In other words, we expect to find a positive *interaction* between Wave and the indicators of stable exposure, which would suggest that higher levels of exposure produce greater overtime increases in knowledge.

As shown in Table 3, the interactions consistently and significantly predict change in *Knowledge of Candidate Issue Positions*. Those with higher levels of stable political television exposure gain more knowledge over time than those with lower levels of exposure. Moreover, the coeffi-

cients are much larger than those in Table 2 for the very same measures, indicating stronger evidence of influence. And again, *Total TV Weighted by Level of Campaign Content* and *Total Number of Political TV Programs* are the strongest predictors.

As shown in Tables C1 and C2 of Appendix C, when we executed the same analyses as in Tables 2 and 3, but this time using *Campaign Participation* as the dependent variable, we obtained parallel results. The same indicators of TV exposure stood out as significant predictors of participation whether the analysis used change in exposure (Table C1) or stable levels of exposure (Table C2) to predict change in participation over time.

## Discriminant Validity

The analyses thus far consistently support the predictive validity of these indicators of political television exposure. Nonetheless, some may question whether these measures have effectively isolated the impact of television as opposed to other media. To alleviate these concerns, we replicated the analyses of predictive validity (the knowledge models in Tables 2 and 3, as well as the campaign participation models in Tables C1 and C2) with controls for exposure to the campaign via newspapers, newsmagazines, political radio shows, and the Internet. We also included a measure of the total number of hours respondents reported watching television during the previous evening to capture the potential effects of exposure to campaign advertising (see Appendix B for question wording). Including these control variables only slightly reduced the size of the coefficients for political television

**TABLE 4 The Impact of Stable Political TV Exposure on Visual Candidate Knowledge**

Political TV Variables:				
Total Number of Political TV Programs	.36*** (.02)			
Total TV Weighted by Level of Campaign Content		.36*** (.02)		
Total TV in Days per Week			.33*** (.02)	
Total TV in Minutes per Week				.29*** (.02)
Control Variables:				
Hours of TV Watched Last Night between 6 and 11	.06*** (.01)	.06*** (.01)	.06*** (.01)	.06*** (.01)
Newspapers	.04*** (.01)	.05*** (.01)	.05*** (.01)	.05*** (.01)
Newsmagazines	.02* (.01)	.02# (.01)	.02* (.01)	.02* (.01)
Political Radio	.10** (.03)	.09** (.03)	.10** (.03)	.10*** (.03)
Internet	.04*** (.01)	.04*** (.01)	.04*** (.01)	.04*** (.01)
Political interest	.15*** (.01)	.15*** (.01)	.15*** (.01)	.15*** (.01)
Strength of party identification	.04*** (.01)	.04*** (.01)	.04*** (.01)	.04*** (.01)
Strength of ideology	.02* (.01)	.02* (.01)	.02* (.01)	.02* (.01)
Constant	.26*** (.02)	.27*** (.02)	.26*** (.02)	.26*** (.02)
Adj. R <sup>2</sup>	.27	.27	.26	.27
Sample Size	8,060	8,060	8,060	8,060

Note: Unstandardized OLS regression coefficients are shown, with standard errors in parentheses. All of the independent variables range from 0–1. Race, gender, age and family income are included in these models. The increment to R-square uniquely associated with the television variable consistently exceeds that of other media or interpersonal communication variables. \*\*\*p < .001, \*\*p < .01, \*p < .05, #p < .10

exposure, which always remained statistically significant (see Tables C3, C4, C5, and C6). These findings not only help us eliminate concurrent effects arising from exposure to other media, but also support the discriminant validity of these indicators.

As a second test of whether we have successfully discriminated political television from other media exposure, we examine the impact of exposure to political television and other media on respondents' recognition of candidate faces, a form of knowledge that is especially likely to be transmitted by television. In wave 1, respondents were shown images of "some people in the news" (Hillary Clinton, John Edwards, Rudy Giuliani, Mitt Romney, Barack Obama, and Fred Thompson) and asked to select the correct name of each person

from a list of 15 names. We label the proportion of correct answers for each respondent *Visual Candidate Knowledge*.

Table 4 shows OLS regression models using the various indicators of television exposure to predict *Visual Candidate Knowledge*. Because this particular analysis is cross-sectional, unlike the previous analyses, we control for potential confounders such as political interest, strength of party identification, education and so forth, along with exposure to other media. Although the case for a causal relationship cannot be as strong as in the fixed effects panel analyses, it is worth noting that the same two versions of the independent variable (*Total Number of Political TV Programs* and *Total TV Weighted by Level of Campaign Content*) are once again the best predictors, providing a highly consistent result.



Moreover, including measures of exposure to the presidential campaign through newspapers, news-magazines, political radio, and the Internet does not eliminate the effects of political television exposure on *Visual Candidate Knowledge*. To examine the unique contribution of political television exposure to visual candidate knowledge, we evaluate the increment to  $R^2$  provided by the inclusion of each television indicator after already including all of the other media exposure measures as well as control variables. We conduct the same analysis for each of the other types of media, first including all other variables in the model except the medium of interest, and then evaluating the unique contribution to visual knowledge of each individual medium. The  $R^2$  change associated with political TV programs is almost identical in all four regressions shown in Table 4, hovering at .02. In contrast, when all other measures including the television measures are included first, the greatest increment to  $R^2$  associated with any of the other media is .004 (for Internet news exposure). As expected, political television exposure distinguishes itself as the best predictor of visual candidate knowledge. Because use of one medium for campaign-related information tends to be correlated with use of another, it is difficult for most media exposure measures to discriminate at this level.

Ideally, one might like some external standard to establish the validity of viewing measures. However, the indicator people turn to most often—Nielsen ratings—raises more questions than it answers. The proprietary nature of Nielsen's methodology means that scholars have *never* had the access needed to assess its reliability or validity, and the few analyses that have been conducted cast serious doubt on this measure (Napoli 2003, Milavsky 1992). Far from a passive method of assessing television exposure, the people-meter system is another imperfect form of self-report with a sample of unknown quantity. Because our media measures tap regular exposure, we cannot directly compare these estimates to Nielsen ratings of viewing on a given evening, except in the crude form of rank-ordering the programs by their popularity. Although we do find that our measures are highly correlated with Nielsen's estimates,<sup>14</sup> we caution against relying on evidence of this kind (cf. Prior 2009a).

<sup>14</sup>Using program as the unit of analysis, we created a variable representing the percent of the NAES sample viewing each program, and another variable representing the Nielsen rating (from September 2007, when the panel began) for each program that could be matched to it. The Pearson correlation between these two variables was .86 ( $p < .001$ ). We also ranked the Nielsen ratings of programs from high to low and ranked the programs based on survey percentage viewing. Spearman's rho was .76 ( $p < .001$ ) for this rank-ordered association.

Given that the centerpiece of our assessment of validity is the measure's ability to predict over-time change in criterion variables at the individual level, it is important to ask whether other factors present a threat to the causal inferences we make. In most observational designs, other individual characteristics could potentially produce spurious relationships. However, because fixed effects regression uses strictly within-person variation over time, and not between-person variation, individual characteristics such as high levels of education could not have produced spurious associations. This represents a huge improvement over most observational designs, which presume that one has measured and controlled for *all* spurious causes of association.<sup>15, 16</sup>

Perhaps the most commonly raised concern is that measures of media exposure are simply manifestations of political interest. This is a serious problem in cross-sectional studies because interested people and uninterested people differ in many ways, including levels of knowledge and participation. But fixed effects regression eliminates the main effects of all individual characteristics, including initial differences in interest. Because fixed effects regression does not eliminate spuriousness arising from variables that could change over time, such as political interest, all of the models include variables representing each wave of the survey to control for the sum-total effects of all other variables that changed during the campaign. Further, we replicated the analyses controlling for possible differential changes over time in individual political interest.<sup>17</sup> This is an especially stringent test given the likelihood that change over time in political television exposure produces change in political interest.<sup>18</sup>

<sup>15</sup>Fixed effects regression controls for spuriousness far better than lagged dependent variable models (i.e., cross-lagged panel models), which have a variety of problems (for details see Achen 2000; Allison 1990, 2009). Moreover, lagged dependent variable models do not assess within-person change in relevant outcomes, but rather, change over time in the rank-order of different people; put another way, such models still rely on between-person variance.

<sup>16</sup>Although fixed effects regression eliminates the *constant* effects of individual characteristics, the impact of those individual characteristics could *vary* over time. We have no theoretical reason to expect the impact of individual characteristics to change over time, but we nonetheless replicated all of the fixed effects analyses with the addition of interactions between Wave and education, age, gender, income, race, strength of ideology, strength of party identification, and political interest, thus capturing the potentially changing impact of these characteristics. Neither the pattern nor the statistical strength of our previous findings changed.

<sup>17</sup>Specifically, in the analyses treating media exposure as stable we included interactions between political interest and Wave, and in the in the analyses treating media exposure as changing we included change in political interest.

<sup>18</sup>Some might argue that interpersonal discussion, rather than media exposure, produced the gains in knowledge and participation.

While spuriousness appears highly unlikely, reverse causation cannot be ruled out. It is possible that some other source increased some respondents' knowledge of candidate issue positions, and this, in turn, produced increases in political television viewing. Likewise, it is possible that participation increased for other, unknown reasons, and this led to a desire to watch more political television. In neither case, however, is such an outside force particularly likely; one is hard-pressed to think of a force that could accomplish this *beyond those already ruled out above*. Politically knowledgeable and participatory people obviously watch more political television than other people, but fixed effects analyses ignore these between-person differences and focus strictly on within-person change over time.

## Prospects For Future Implementation

Thus far we have demonstrated that the program list technique for measuring political television exposure has strong true-score reliability, along with good content validity, predictive validity, and discriminant validity. We have emphasized these standards for measurement quality because they represent the most difficult hurdles. The indicators derived from the program list technique also demonstrate the same strengths that traditional measures of exposure have, such as good convergent validity (in their correlations with all of the usual suspects, such as political interest and education) and good concurrent validity (in their cross-sectional relationships with other measures of political media use). Moreover, although we have only demonstrated this approach's feasibility in measuring television exposure, its utility likely extends to measurement of exposure to other types of media, such as radio programs, newspapers, websites, and so forth.

As with most approaches to measurement, the program list measure has its tradeoffs. By limiting our measures to regular exposure, which is a low bar as defined

Yet this theory, known as the two-step flow of communication, posits that some people are exposed to political media, and that these "opinion leaders" spread the message to many other people (Katz and Lazarsfeld 1955). In other words, interpersonal discussion is not a spurious influence, but rather a mediator of mass media influence. Moreover, the two-step flow has received little to no empirical support to date (Bennett and Manheim 2006; Chaffee and Hochheimer 1985; Gitlin 1978). Nonetheless, we replicated the analyses in Tables 2 and 3 (and Tables C1 and C2) controlling for the time-varying impact of interpersonal discussion (that is, with interactions between Wave and interpersonal discussion), and replicated the analyses in Table 4 including interpersonal discussion as a control variable. The results were unchanged.

here, we eliminate the need for quantification of hours, minutes, or days for each program. But we inevitably lose information about fleeting, incidental instances of exposure, though few believe that survey respondents can accurately self-report viewing of this kind in any case (e.g., Prior 2009b). In addition, asking people to select programs from a list is feasible only when there is a screen or written document in front of the respondent, making it impractical for telephone surveys. Given the declining ease and popularity of representative telephone surveys (see, e.g., Chang and Krosnick 2003; Keeter, Christian, and Dimock 2010), and the growing use of high-quality probability samples collected over the Internet, we suggest that this issue will not substantially hamper the ongoing feasibility of this approach. Moreover, face-to-face surveys such as the ANES now regularly use laptops on which respondents respond to lists of this kind.

Another drawback to this approach in a fragmented media environment stems from the small audiences attending to individual political programs. Fragmented audiences require longer lists of programs in order to capture the bulk of political television viewing. And the more lengthy the program list, the larger the amount of survey time required. Four separate screens of programs were used to create our measures, and the median time per screen was roughly 23 seconds. Overall, the entire series took under 2 minutes (a median of 100 seconds) to complete, a substantial amount of time for a large omnibus survey devoted to many research purposes.

To evaluate the potential loss of information that would come with a more efficient version of the program list measure, we evaluated a three-screen version by eliminating one screen's worth of programs (the 12 programs with the lowest viewership).<sup>19</sup> A 3-screen version of the program list technique took just over a minute to answer, which is not substantially different in survey time from what the ANES currently allots for measuring television exposure.<sup>20</sup> Moreover, the correlations between the 3-screen and 4-screen measures were consistently high, ranging from .98 to .99. The true-score reliabilities also remained high, with .82 and .83 for *Total Number of Political TV Programs* and *Total TV Weighted by Campaign Content*, respectively. These results suggest that a 3-screen

<sup>19</sup>These programs had audiences that ranged from zero to 7 percent of the sample.

<sup>20</sup>Although questions have varied over time, the ANES asks a series of separate questions addressing any television exposure to the campaign, how much, how many days per week, how much attention to local news on television, to national news, and to news about the campaign. Each of these items takes between 7 and 12 seconds each.

(37 program) version of the program list method would work equally well for most purposes.

A final potential concern is whether the ability of these measures to specify convincing causal effects from political television exposure depends on the unusually large sample size (approx. 10,000 respondents) of the NAES panel survey. Zaller (2002) argues that most surveys do not have the statistical power to identify media effects. To evaluate this possibility, we conducted a series of simulations randomly sampling subsets of cases from this unusually large sample. As shown in Table D1 in Appendix D, although larger samples certainly make more subtle effects easier to identify, samples much smaller than  $n = 10,000$  (e.g.,  $n = 3,000$  or  $2,000$  or  $1,000$ ) should be able to identify effects using measures of this kind.

## Conclusion

Survey measures of self-reported exposure to political television are a staple of studies of campaigns, but measurement techniques that produce demonstrably valid and reliable indicators of exposure to political television have frustrated scholars for many decades. This problem has been widely blamed for a lack of progress in this area of study (e.g., Bartels 1993; Zaller 2002), and the problem has been exacerbated by the increasing fragmentation of the media environment (Williams and Delli Carpini 2011). By using the names of specific television programs to measure exposure, this approach eliminates the problem of determining whether the researcher and the respondent share the same definition of what constitutes “news” or “politics.” It also incorporates programs watched in real time versus delayed viewing, and programs watched via a television as well as those watched online via computer. Moreover, by using three waves of panel data repeating the measures of exposure and political knowledge, we were able to demonstrate the true-score reliability and predictive validity of the exposure indicators in a far more powerful way than in previous cross-sectional attempts at this same task.<sup>21</sup>

Based on our evaluations of true-score reliability and predictive validity, the best indicators of exposure to political television appear to be a simple count of the total number of political television programs viewed, or the same item weighted by a simple scheme accounting for variations in the extent to which a program covers the

<sup>21</sup>In some cases, it was not clear that the dependent variables were best analyzed using linear fixed effects models as opposed to ordered logit fixed effects models. We repeated these analyses using ordered logit and found precisely the same pattern of results.

campaign. In every analysis, without exception, these two indicators consistently performed the best. This approach produced more detailed indicators of people’s exposure (not just how much in total, but which programs), while simultaneously generating high true-score reliabilities, at least higher than many of the traditional outcome variables that media exposure variables are expected to predict.

Our analyses further suggest that political television exposure is a relatively stable behavior, at least during the time period we studied. Though there was some change over time, (an overall increase as the election drew close), people tended to watch the same number of political programs across the three waves; for the most part, they even stuck with the exact same programs. The importance of this observation is easily overlooked. It suggests that watching television is a highly habitual behavior. People do not approach each viewing opportunity anew (e.g., Adams 2000; Wood, Quinn, and Kashy 2002).

Measuring televised exposure to politics outside of laboratories may be less intransigent than it seems. The program list approach to measuring political television exposure produces indicators with relatively high true-score reliabilities, very promising predictive validity, and impressive discriminant validity. Finally, because of the relative stability and habitual nature of television exposure, it may not be necessary to repeat measures over time in order to evaluate exposure’s ability to predict other outcomes. The study of media effects outside of experimental contexts could still benefit from repeated measurement of dependent variables in order to avoid confounding individual differences with effects, but political television exposure may not change enough to warrant that same investment.

## Acknowledgements

We are grateful to the Annenberg Public Policy Center of the University of Pennsylvania for enabling the collection of data used in this study. We would also like to thank the editors of the *American Journal of Political Science* for their helpful comments and suggestions.

## References

- Achen, Christopher H. 2000. “Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables.” Paper presented to the Annual Meeting of the Political Methodology section of the American Political Science Association, July 20–22, 2000.

- Adams, William, J. 2000. "How People Watch Television as Investigated Using Focus Group Techniques." *Journal of Broadcasting & Electronic Media* 44:78–93.
- Allison, Paul D. 1990. "Change Scores as Dependent Variables in Regression Analysis." *Sociological Methodology* 20:93–114.
- . 2009. *Fixed Effects Regression Models*. Sage Publications.
- Althaus, Scott L., and David H. Tewksbury, 2007. "Toward A New Generation of Media Use Measures for the ANES." Report to the Board of Overseers, American National Election Studies.
- . 2009. "Measuring News Exposure in Contemporary Media Systems." Paper presented at the Annual Meeting of the International Communication Association, May 21–25, Chicago.
- Bartels, Larry M. 1993. "Messages Received: The Political Impact of Media Exposure." *American Political Science Review* 87:267–85.
- Bennett, W. Lance, and Jarol B. Manheim. 2006. "The One-Step Flow of Communication." *ANNALS of the American Academy of Political and Social Science* 608(1):213–32.
- Chaffee, Steven H., and John L. Hochheimer. 1985. "The Beginnings of Political Communication Research in the United States: Origins of the "Limited Effects" Model." In Everett M. Rogers and Francis Balle (Eds.), *The Media Revolution in America and Western Europe* (pp. 267–96). Norwood, NJ: Ablex.
- Chaffee, Steven H., and Joan Schleuder. 1986. "Measurement and Effects of Attention to Media News." *Human Communication Research* 13:76–107.
- Chang, LinChiat, and Jon A. Krosnick. 2003. "Measuring the Frequency of Regular Behaviors: Comparing the 'Typical Week' to the 'Past Week.'" *Sociological Methodology* 33:55–80.
- . 2009. "National Surveys Via RDD Telephone Interviewing Versus the Internet: Comparing Sample Representativeness and Response Quality." *Public Opinion Quarterly* 73(4):641–78.
- Gitlin, Todd. 1978. "Media Sociology: The Dominant Paradigm." *Theory and Society* 6(2):205–53.
- Heise, David R. 1969. "Separating Reliability and Stability in Test-Retest Correlation." *American Sociological Review* 34(1):93–101.
- Huesmann, L. Rowell, Kirsti Lagerspetz, and Leonard D. Eron. 1984. "Intervening Variables in the TV Violence-Aggression Relation: Evidence from Two Countries." *Developmental Psychology* 20(5):746–75.
- Huesmann, L. Rowell, Jessica Moise-Titus, Cheryl-Lynn Podolski, and Leonard D. Eron. 2003. "Longitudinal Relations Between Children's Exposure to TV Violence and their Aggressive and Violent Behavior in Young Adulthood: 1977–92." *Developmental Psychology* 39(2):201–21.
- Katz, Elihu, and Paul F. Lazarsfeld. 1955. *Personal Influence: The Part Played by People in the Flow of Mass Communications*. Free Press.
- Keeter, Scott, Leah Christian, and Michael Dimock. 2010. "The Growing Gap between Landline and Dual Frame Election Polls." Pew Research Center for the People & the Press. Available at: <http://pewresearch.org/pubs/1806/growing-gap-between-landline-and-dual-frame-election-polls>
- SUSANNA DILLIPLANE, SETH K. GOLDMAN AND DIANA C. MUTZ
- King, Gary. 1986. "How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science* 30(3):666–87.
- Lodge, Milton, and Patrick Stroh. 1993. "Inside the Mental Voting Booth: An Impression-Driven Model of Candidate Evaluation." In Shanto Iyengar and William J. McGuire (Eds.), *Explorations in Political Psychology*. Durham: Duke University Press.
- Lord, Frederic, and Melvin Novick. 1968. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.
- Milavsky, J. Ronald. 1992. "How Good is the A. C. Nielsen People-Meter System? A Review of the Report by the Committee on Nationwide Television Audience Measurement." *Public Opinion Quarterly* 56:102–115.
- Napoli, Philip M. 2003. *Audience Economics: Media Institutions and the Audience Marketplace*. New York, NY: Columbia University Press.
- Price, Vincent. 1993. "The Impact of Varying Reference Periods in Survey Questions About Media Use." *Journalism Quarterly* 70:615–27.
- Price, Vincent, and John Zaller. 1990. "Evaluation of Media Exposure Items in the 1989 NES Pilot Study." Report to the Board of Overseers, American National Election Studies.
- . 1993. "Who Gets the News? Alternative Measures of News Reception and their Implications for Research." *Public Opinion Quarterly* 57(2):133–64.
- Prior, Markus. 2009a. "The Immensely Inflated News Audience: Assessing Bias in Self-Reported News Exposure." *Public Opinion Quarterly* 73(1):130–43.
- . 2009b. "Improving Media Effects Research Through Better Measurement of News Exposure." *Journal of Politics* 71(3):893–908.
- Schwarz, Norbert, and Daphna Oyserman. 2001. "Asking Questions about Behavior: Cognition, Communication, and Questionnaire Construction." *American Journal of Evaluation*, 22, 127–160.
- Wiley, David E., and James A. Wiley. 1970. "The Estimation of Measurement Error in Panel Data." *American Sociological Review* 35:112–17.
- Williams, Bruce A., and Michael X. Delli Carpini. 2011. *After Broadcast News: Media Regimes, Democracy, and the New Information Environment*. Cambridge University Press.
- Wood, Wendy, Jeffrey M. Quinn, and Deborah Kashy. 2002. "Habits in Everyday Life: Thought, Emotion and Action." *Journal of Personality and Social Psychology* 83: 1281–97.
- Zaller, John. 2002. "The Statistical Power of Election Studies to Detect Media Exposure Effects in Political Campaigns." *Electoral Studies* 21:297–329.

## Supporting Information

Additional supporting information may be found in the online version of this article:

**Appendix A:** Details of the Political Television Measures

**Table A1:** Consistency in Program Viewership across Waves 2, 4, and 5

**Table A2:** Raw Over-Time Correlations of Political Television Exposure Measures

**Table A3:** Stability of Political Television Exposure Measures

**Appendix B:** Question Wording and Coding of Variables

**Appendix C:** Supplemental Analyses of Predictive Validity

**Table C1:** The Impact of Within-Person Change in Political TV Exposure on Within-Person Change in Campaign Participation (Fixed Effects Panel Analysis)

**Table C2:** The Impact of Stable Levels of Political TV Exposure on Within-Person Change in Campaign Participation (Fixed Effects Panel Analysis)

**Table C3:** The Impact of Within-Person Change in Political TV Exposure on Within-Person Change in Knowledge of Candidate Issue Positions, Controlling for Exposure to Other Media (Fixed Effects Panel Analysis)

**Table C4:** The Impact of Stable Political TV Exposure on Within-Person Change in Knowledge of Candidate

Issue Positions, Controlling for Exposure to Other Media (Fixed Effects Panel Analysis)

**Table C5:** The Impact of Within-Person Change in Political TV Exposure on Within-Person Change in Campaign Participation, Controlling for Exposure to Other Media (Fixed Effects Panel Analysis)

**Table C6:** The Impact of Stable Political TV Exposure on Within-Person Change in Campaign Participation, Controlling for Exposure to Other Media (Fixed Effects Panel Analysis)

**Appendix D:** Supplemental Simulation Analyses

**Table D1:** Percentage of Simulations in which Political TV Exposure Predicts Outcomes, by Sample Size

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.