Polit Behav
DOI 10.1007/s11109-007-9027-8

ORIGINAL PAPER

Retrospective and Prospective Performance Assessments during the 2004 Election Campaign: Tests of Mediation and News Media Priming

Neil Malhotra · Jon A. Krosnick

© Springer Science+Business Media, LLC 2007

Abstract According to many theoretical accounts of the vote choice, distal determinants (e.g., party identification) influence proximal determinants (e.g., perceptions of candidates), which in turn shape candidate preferences. Yet almost no research on voting has formally tested such mediational hypotheses. Using national survey data collected between February and September of 2004, this paper begins by illustrating how to conduct such investigations. We explored whether public approval of President Bush's handling of a series of specific national problems (e.g., the Iraq war) influenced overall assessments of his job performance and evaluations of his likely future performance versus John Kerry's, which in turn shaped vote choices. The results are consistent with the claim of mediation and shed additional light on the impact of various issues on the 2004 election outcome. We also tested what we term the ''dosage hypothesis,'' derived from news media priming theory, which posits that changes in the amount of media coverage of an issue during the course of a campaign should precipitate changes in the weight citizens place on that issue when evaluating the president's overall job performance, particularly among citizens most exposed to the news. Surprisingly, this analysis did not yield consistent support for the venerable dosage hypothesis, suggesting that the conditions under which priming occurs should be specified much more precisely in future work.

N. Malhotra (✉)
Department of Political Science, Stanford University, Encina Hall West, Room 100, Stanford, CA 94305-6044, USA
e-mail: neilm@stanford.edu

J. A. Krosnick
Departments of Communication, Political Science, and Psychology, Stanford University, 434 McClatchy Hall, 450 Serra Mall, Stanford, CA 94305, USA
e-mail: krosnick@stanford.edu

Keywords Retrospective voting · Presidential approval · 2004 election · News media priming · Mediation

For decades, political scientists have presumed that citizens' presidential vote choices and their overall evaluations of competing candidates are shaped partly by retrospective assessments of the incumbent's performance and by comparative prospective judgments of the likely future performance of the competing candidates. And scholars have long presumed that overall evaluations of the President's job performance are derived partly from assessments of his performance handling various specific national problems. This latter process of derivation has been shown to be shaped in part by the focus of recent news media stories, as outlined by the theory of news media priming.

In this paper, we bring these ideas together into a single, coordinated investigation with a focus on the notion of mediation. The term ''mediation'' has appeared only very rarely in the literature on presidential candidate evaluation and voting. Yet other disciplines, especially psychology, have been proposing and testing mediational hypotheses for many years now, producing valuable theoretical advances as a result. In an effort to illustrate the value of serious attention to mediation, we explored the possibility that domain-specific assessments of presidential job performance may not shape intended vote choices directly but may instead do so indirectly, by influencing overall job approval ratings and comparative prospective evaluations of the likely performance of the incumbent and the challenger. That is, overall and prospective presidential and challenger evaluations may *mediate* the relation between domain-specific assessments and vote choice.

We also tested a widely-believed *moderation* hypothesis: when deriving overall presidential performance evaluations, people place more weight on domains that the news media have covered more extensively (i.e., news media priming). To complement the existing literature on priming, we explored this possibility in a new way, by gauging whether changes in the volume of media attention to issue domains over the course of a presidential campaign predicted corresponding variation in the weights attached to these domains when deriving overall presidential evaluations, particularly among Americans who were highly exposed and attentive to political news.

In the first section, we discuss the notion of mediation and outline its applicability to the study of voting, as well as describe the data we used (from ABC News/*Washington Post* surveys). Then, we outline the news media priming hypotheses that we set out to test and describe the analyses we conducted, the findings we produced, and their implications for understanding the 2004 American presidential election.

The Impact of Retrospective and Prospective Importance Assessments on Vote Choice: Testing Mediation

Overview

The literature on voting is filled with hypotheses that implicate mediational relations: an independent variable (X) affects a mediator (Z), which in turn causes the dependent variable (Y). According to Baron and Kenny (1986), a mediator is a ''variable which represents the generative mechanism through which the focal independent variable is able to influence the dependent variable of interest'' (p. 1173). For instance, in *The American Voter*, Campbell, Converse, Miller, and Stokes (1960) proposed the ''funnel of causality.'' External conditions in the wide part of the funnel were said to affect personal, political attitudes in the narrow region, which were thought ultimately to determine vote choices. Within the narrow region, party identification was a principal independent variable, causing factors more proximate to vote choice, such as candidate evaluations. Party identification was posited to be shaped by more distant factors, such as presidential candidate quality (Jennings & Markus, 1984), policy platforms (Gerber & Jackson, 1993), and government performance (Fiorina, 1981). Campbell et al. offered some more specific mediational hypotheses, each consisting of a chain of multiple causal steps. For example, they conjectured that an individual's employment status may affect his or her view of the incumbent government, which in turn may affect his or her attitudes towards the parties, which may determine vote choices. However, this and other mediational hypotheses were never explicitly tested in *The American Voter*, which instead almost exclusively reported bivariate associations between pairs of variables. The authors showed how X was associated with Z and how Z was associated with Y, but the indirect linkage of X with Y via Z was never evaluated. Yet the observed correlations would have appeared even if X caused Z in one subgroup of the population and Z caused Y in a different subgroup, which would be inconsistent with the notion of mediation.

The literature on retrospective voting also offers many mediational hypotheses. For example, assessments of government performance, particularly with respect to the economy, are thought to influence overall presidential approval, which in turn presumably influences vote choice (Fiorina, 1981; Kinder & Kiewiet, 1979, 1981). A related body of research has claimed that presidential approval is shaped by citizens' beliefs about how the incumbent has helped or hurt social groups with which they identify, and presidential approval in turn influences vote choice (e.g. Conover, 1985; Kinder, Adams, & Gronke, 1989). And some evidence suggests that voters' personal financial circumstances may shape their perceptions of the nation's overall economic health, which in turn shape candidate preferences in elections (e.g., Abramowitz, Lanoue, & Ramesh, 1988). Thus, pocketbook considerations may influence votes, but only indirectly via sociotropic beliefs. However, this hypothesis and others like it have never been explicitly tested in a mediational

framework. Routinely, researchers demonstrate two correlational relations independently, such as the association of economic perceptions with presidential approval and the association of approval with vote choice. But observing two such relations is not sufficient for demonstrating indirect causal influence.

In the literature on news media priming, mediational hypotheses have also often been posited. Laboratory experiments and national survey evidence has suggested that the news media prime citizens by influencing which policy domains citizens use to judge political elites (e.g., Iyengar & Kinder, 1987; Krosnick & Kinder, 1990). This literature has overwhelmingly presumed that media coverage of an issue increases the accessibility of citizens' beliefs and evaluations regarding that issue in their memories, which in turn leads people to weight that issue more heavily when constructing relevant evaluations. However, this causal mechanism (news media coverage leading to accessibility leading to greater evaluative weight) was never explicitly tested until Miller and Krosnick (2000) did so. And they found no evidence that news coverage of an issue causes increased evaluative weight on that issue via increases in accessibility. Thus, widespread theoretical presumptions can be debunked when mediation is finally tested formally.

Although mediational hypotheses have frequently been offered in theories of voting, these hypotheses have rarely been explicitly tested. There are, of course, exceptions, such as Markus and Converse's (1979), Page and Jones' (1979), and later studies of the dynamics of policy preferences, partisanship, and vote choice. And these studies have proposed a series of different, competing hypotheses. For example, Markus and Converse (1979) posited that party identification shaped perceptions of candidates' personalities, that policy issue preferences and perceptions of candidate personalities had independent effects on candidate evaluations, and that candidate evaluations determined vote choices. Although Page and Jones (1979) agreed with this last step (that candidate evaluations determine vote choices), these investigators theorized that party identification could mediate the impact of policy preferences on candidate evaluations and that policy preferences could mediate the impact of party identification on candidate evaluations. Interestingly, their empirical investigation yielded support for the former but not the latter. Remarkably, Kenney and Rice's (1988) more recent investigation concluded that party identification had no direct effect on candidate evaluations and instead influenced them only indirectly, by shaping policy preferences. Thus, their mediational chain was just the opposite of Page and Jones' (1979): party id $\rightarrow$ policy preferences $\rightarrow$ candidate evaluations. But none of these studies conducted formal tests of mediation developed recently by statisticians, so it is difficult to know for sure what to make of these conflicting findings.

To assess more fully the extent to which the voting literature has formally posited and tested mediational hypotheses, we conducted a content analysis of leading publications. The American National Election Study (ANES) has accumulated 1,717 articles published since the 1950s in which vote choice was

the dependent variable. We randomly sampled 40 articles published since 1978 from this collection to assess whether they posited mediational hypotheses and whether they formally tested such hypotheses.

Fully half of these articles treaded into dangerous analytic territory. In all of them, vote choice was regressed on an array of predictors, and among these predictors were variables that have been posited as mediators of the effects of independent variables that were also included in the same regression equation. For instance, in some of these articles, both pocketbook and sociotropic economic judgments were included simultaneously in regression equations. In others, both domain-specific and overall presidential approval measures were employed as predictors simultaneously. When both a mediator and an independent variable are included in the model as such, a researcher may incorrectly conclude that the statistically insignificant independent variable had *no* effect on vote choice, when in reality it may have affected vote choice *indirectly* through the statistically significant mediator. And indeed, this sort of conclusion of no effect was reached regularly in the publications we examined, despite a theoretical rationale suggesting mediation.

One illustration of how mediational thinking has at times made its way into the voting literature is Hetherington's (1996) well-known study of presidential elections. He posited that more media consumption during the 1992 campaign caused Americans to view national economic conditions as being worse, which in turn decreased their likelihood of voting for President George H. W. Bush. Hetherington first regressed perceptions of national economic conditions on media consumption and confirmed the expected negative relation. Then he regressed vote choice on perceptions of national economic conditions and found the expected effect: more positive assessments were associated with a higher likelihood of voting for President Bush. Thus, there has been some explicit testing of mediated effects on voting, though according to Baron and Kenny's (1986) argument, even Hetherington's (1996) analysis did not include all the steps needed to documented mediation completely.

Testing Mediation Explicitly

Baron and Kenny (1986) argued that three analytic steps must be executed in order to fully test a mediational hypothesis: (1) regress the dependent variable on the independent variable, finding a significant relation; (2) regress the mediator on the independent variable, finding a significant relation; and (3) regress the dependent variable on the independent variable and the mediator; the effect of independent variable should be weakened, which can be confirmed statistically with a Sobel (1982) test. Tests of mediation require that investigators have theoretical foundations for assumptions about the directions of the causal relations between variables, because the statistics themselves do not determine causal direction. Recent advances in mediational analysis have focused on relaxing parametric assumptions to better estimate the confidence interval of the mediating effect, especially in small samples (MacKinnon, Lockwood, & Williams, 2004; Shrout & Bolger, 2002). The key,

though, is to assess explicitly whether variance in the mediator that is shared with the independent variable is also shared with the dependent variable. That is the condition necessary for mediation that has so rarely been assessed in past voting studies. Mediational analysis employs the principles of covariance structure modeling and path analysis and entails implementing a prescribed set of analytic steps to explicitly test for conditions that must obtain if hypotheses positing particular indirect causal effects are valid.

Statistical tests of mediation have come front and center in many disciplines in the social and natural sciences, most prominently in psychology, organization theory, and clinical research. Indeed, the canonical Baron and Kenny (1986) article has been cited in 7,452 papers in these and other disciplines. This phenomenal degree of citation has increased exponentially over time; the paper was cited 32 times in 1989, 304 times in 1997, and 1,031 times in 2005. Recent tests of mediation include whether subjective evaluations of alcohol expectancies mediate the association between perceived drinking expectancies and participation in drinking games (Zamboanga, Bean, Pietras, & Pabon, 2005), whether self-assessments of health mediate the relation between body weight and depression (Needhman & Crosnoe, 2005), whether reduction in post-traumatic stress disorder mediates the impact of disability benefits on veterans' odds of poverty (Murdoch, Van Ryn, Hodges, & Cowper, 2005), and many more. In these ways, researchers have suddenly become much more concerned with identifying direct effects and indirect effects via mediators. And with the growth in the popularity of the method has come a surge in the precision and sophistication of theory-building to account for social cognition and behavior.

Hypotheses

In an effort to illustrate the value of this methodology for the study of voting, we explored the possibility that the impact of domain-specific assessments of presidential performance may shape intended vote choices via mediational relations. To do so, we weaved together speculations offered in various separate literatures and also posited some new pathways of influence. Our starting point was the wide-spread presumption that citizens derive overall assessments of the president's job performance from perceptions of his or her performance in handling specific domains. As illustrated at the top of Fig. 1, that effect can be blended in a single causal diagram with the notion that overall performance assessments may shape intended vote choices, thus mediating the impact of domain-specific performance assessments on intended vote choices. As we will discuss more later in this article, the hypotheses we test entail making assumptions about causal direction, and as in all science, our analytic approach can produce evidence challenging those assumptions but not validating them.

As shown at the bottom of Fig. 1, prospective assessments of incumbent and challenger performance may also mediate the relation between domain-specific performance evaluations of the president and intended vote choice.

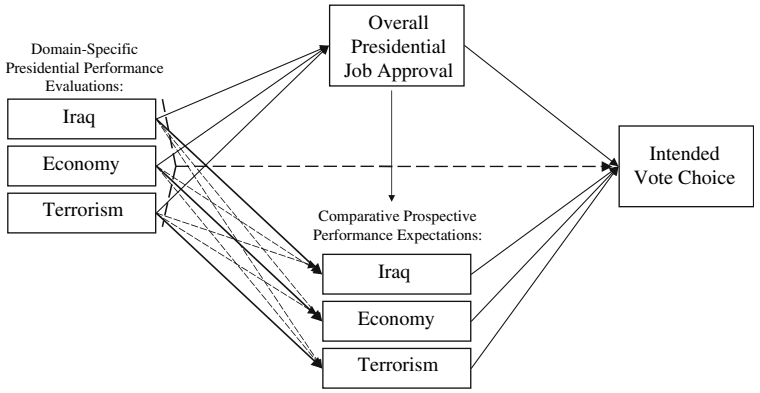

Fig. 1 Model of vote choice

Specifically, a voter who believes that the president has handled a domain more effectively may be more likely to believe that he or she will perform better than the challenger in the future in managing that domain. And beliefs that a particular candidate will handle a specific domain better may have enhanced the likelihood that a citizen will intend to vote for that candidate. When portrayed in this fashion, this analysis raises the interesting question of whether domain-specific performance may shape intended vote choices more via backward-looking assessments of past presidential performance or forward-looking guesses about future performance.

The horizontal path across the middle of Fig. 1 (indicated by the dotted line) represents yet another interesting possibility: domain-specific assessments may have direct effects on intended vote choice, not mediated by retrospective summary judgments or prospective domain-specific judgments. That is, instead of using domain-specific judgments simply to derive overall performance assessments, these domain-specific judgments may be consulted directly at the time of vote choice formation. This may be especially likely, for example, if an issue is especially important to voters and/or an extremely frequent focus of news media attention.

Lastly, people may derive their expectations about likely future performance by the incumbent and challenger partly based upon the incumbent's overall performance to date, shown by the vertical path in the middle of Fig. 1. That is, the better an incumbent is seen as handling *any* problem domain, the more optimistic voters may be about how he or she (or his or her party) will handle any other problem domain in the future.

Data

To evaluate these mediational hypotheses, we analyzed data from RDD (random digit dialing) telephone surveys commissioned by ABC News and the *Washington Post* and conducted by TNS with representative national samples of American adults on five occasions between February and September, 2004

(Ns of about 1,000 respondents per wave).[1] One adult per household was selected to be surveyed using the ''last birthday'' method from among adults at home at the time of the call (Lavrakas, 1993). Because women are more likely to be reached than men, interviewers asked first to speak with the male household member with the most recent birthday on 75% of the calls (chosen randomly) and asked to speak with the female household member with the most recent birthday on the other 25% of the calls. If no one of the requested gender was available, the interviewer asked to speak with the adult household member of the opposite gender who had the most recent birthday. Cell-based weighting was done using the U.S. Census Bureau's Current Population Survey so that the sample proportions matched the population proportions in terms of age, race, gender, and education (Langer & Merkle, 2006).

The entire questionnaire was administered to all respondents in early survey waves, but in later waves, respondents who said they were not registered to vote were not asked all the questions. In order to assure comparability of respondent populations across waves, we focused all of our analyses only on people who said they were registered to vote (79.4% on 2/10, 80.5% on 5/20, 75.7% on 7/22, 72.6% on 8/26, and 75.6% on 9/23).

The following measures were analyzed:

Overall approval of President Bush was measured with the following two questions: ''Do you approve or disapprove of the way George W. Bush is handling his job as president?'' and ''Do you approve/disapprove strongly or just somewhat?'' Responses were coded as follows: ''strongly disapprove'' was coded 0, ''somewhat disapprove'' as coded .33, ''somewhat approve'' was coded .67, and ''strongly approve'' was coded 1.

Domain-specific approval of President Bush was measured with the question, ''Do you approve or disapprove of the way George W. Bush is handling: [the economy, the situation in Iraq, the U.S. campaign against terrorism]?'' ''Approve'' responses were coded 1, and ''disapprove'' responses were coded 0.

Prospective judgments of how President Bush and Senator Kerry would handle the three domains were measured with the question: ''Who do you trust to do a better job handling—[the economy, the situation in Iraq, the U.S. campaign against terrorism]—Bush or Kerry?'' Respondents who chose President Bush were coded 1, and respondents who chose Senator Kerry were coded 0.

Intended vote choice was measured by the question, ''If the 2004 presidential election were being held today, would you vote for: (Response choices: ''Bush (and Cheney)'', ''Kerry (and Edwards)'').'' The question included the names of the running mates in the waves conducted after Cheney and Edwards had been selected. Respondents who did not select a candidate were asked a follow-up question: ''Which one are you leaning toward?'' Responses

[1] ABC conducted more pre-election surveys, but we focused on only these five because these were the only questionnaires including the complete set of measures that we needed to allow identical analyses to be conducted at all waves. This permitted us to make over-time comparisons of coefficients for the same equations.

were coded 1 for President Bush and 0 for Senator Kerry. All other responses were treated as missing data.[2]

Democrat and *Republican.* Respondents were asked: ''Generally speaking, do you usually think of yourself as: a Democrat, a Republican, an Independent, or what?'' A Democrat dummy variable was coded 1 for Democrats and 0 for all others. A Republican dummy variable was coded 1 for Republicans and 0 for all others.

Liberal/conservative ideology. Respondents were asked: ''Would you say your views on most political matters are liberal, moderate, or conservative?'' A liberal dummy variable was coded 1 for liberals and 0 for all others. A conservative dummy variable was coded 1 for conservatives and 0 for all others.

Education was coded using a set of dummy variables representing five different degree levels: high school, associates degree, some college, college, and graduate. Respondents without a high school diploma constituted the omitted, comparison category.

Race was represented by a set of dummy variables for three different racial groups: blacks, non-white Hispanics, and other non-white respondents. Whites constituted the omitted, comparison category.

Gender was coded 1 for males and 0 for females.

Age was a continuous variable coded to lie between 0 (meaning the youngest person across all waves, who was 18 years old) and 1 (meaning the oldest person across all waves, who was 99 years old). In all model tests, we also included a quadratic term to test for nonlinearities.

Results

To illustrate what results would be obtained via the standard approach of including all possible predictors in a single regression predicting intended vote choice, we estimated the parameters of an equation with these predictors: overall presidential approval, domain-specific presidential approval for Iraq, the economy, and terrorism, comparative prospective evaluations of performance by both candidates with respect to those three same domains, and the demographic and political controls (see Table 1). According to this approach, overall presidential approval was the most powerful predictor of intended vote choice (coefficient = 3.71, $p < .001$), and prospective comparative performance evaluations predicted intended vote choice significantly but more weakly (coefficients = 2.81, 2.71, and 1.56, respectively, $p < .001$ in all cases). Assessments of President Bush's handling of the economy were a significant predictor

[2] In all waves except 2/10, Ralph Nader was explicitly mentioned. In all waves except 9/23, Nader voters were asked: ''If Nader does not run or is not on the ballot, for whom would you vote—Bush or Kerry?'' We used responses to this follow-up in constructing our measure of intended vote choice. For the 2/10 wave, ABC News specifically asked a question dealing with a general election contest between Bush and Kerry, making it directly comparable to the questions in later waves. Even though Kerry did not have the requisite number of delegates secured at this point in time, ABC News asked this question because it was evident that Kerry was going to be the nominee based on his victories (often resounding) in 12 of 14 primaries up to that point.

Table 1 Logistic regression predicting vote choice with overall and domain-specific approval ratings

Overall Bush approval	3.71***
Bush current handling: Iraq	–.43
Bush current handling: Economy	.99**
Bush current handling: Terrorism	–.44
Prospective: Iraq	2.81***
Prospective: Economy	2.71***
Prospective: Terrorism	1.56***
Democrat	–.93*
Republican	1.32***
Liberal	.30
Conservative	.48
Age	2.82
Age squared	–2.88
Black	.26
Hispanic	.98
Other	–.28
Male	–.17
High School	–.97
Associates	–.07
Some College	–.32
College	–.70
Graduate	–.82
Constant	–5.96***
Pseudo R^2	.89
N	3,570
Mean VIF	6.30

***$p < .001$; **$p < .01$; *$p < .05$

Note: Registered voters only. All data are weighted

as well (coefficient = .99, $p < .01$), but perceptions of his handling of Iraq and terrorism were not (coefficients = –.43 and –.44, respectively, ns in both cases).

Let us assume for a moment that an analyst has no theory of mediational relations among these variables and had simply thrown them all into this regression in order to explore which drove intended votes. If so, then he or she might conclude that intended votes were a function mainly of a general judgment about the president's past performance and that consequential issue-specific judgments were the prospective ones, not the retrospective ones. Of course, in light of the arguments embodied in Fig. 1, it is easy to recognize that the domain-specific current approval judgments may have had indirect effects on intended vote choice via other predictors in Table 1's regression. Therefore, one should not presume that the current approval judgments had almost no effects at all (as the coefficients in Table 1 might seem to suggest). And indeed, it is easy to recognize that even the strong coefficient for overall Bush approval judgments in Table 1 may underestimate their total impact on intended vote choice, if some of that impact was mediated by prospective performance assessments. But without such theory-based expectations, it is easy to imagine an analyst concluding that candidate preferences during this campaign were driven by the posited causes according to the magnitudes of the coefficients in Table 1.

To test whether overall approval might have mediated the relations between domain-specific approval judgments and intended vote choice, we followed Baron and Kenny's (1986) approach: overall approval (the mediator) was regressed on domain-specific approval judgments (along with control variables); intended vote choice was regressed on domain-specific approval judgments (with controls), and intended vote choice was regressed on both domain-specific judgments *and* overall approval (with controls). If the effect of domain-specific judgments in the first and second equations was statistically significant, and if the effect of the domain-specific judgments on the dependent variable was significantly smaller in the third regression than in the second regression, then support for mediation would be obtained. Sobel tests were computed to formally ascertain whether the effects of domain-specific approval ratings were significantly smaller in the third regression than in the second.

The mediational model illustrated in Fig. 1 was generally confirmed by these data (see Table 2). For each of the three policy areas, domain-specific judgments significantly predicted overall presidential approval (the posited mediator): b = .28, .24, and .16 for Iraq, the economy, and terrorism, respectively (see row 1 of column 1, row 2 of column 4, and row 3 of column 7). Next, we assessed whether domain-specific judgments predicted vote choice more weakly when overall Bush approval (the mediator) was included in the regressions (shown in bold near the tops of columns 3, 6, and 9 of Table 2) than when overall Bush approval was not included (shown in bold near the tops of columns 2, 5, and 8 of Table 2). Overall approval completely mediated the relation between Iraq approval and intended vote choice (reducing the coefficient from 1.26, $p < .001$, to .26, ns, in row 1 of Table 2; Sobel test = 8.59, $p < .001$) and partially mediated the relation of economy approval with intended vote choice (reducing the coefficient from 2.41, $p < .001$, to 1.00, $p < .01$, in row 2 of Table 2; Sobel test = 6.85, $p < .001$). The coefficient for terrorism approval was significantly smaller in the third regression than in the second, but assessments of the president's performance in that domain did not have a significant, independent effect on intended vote choice (coefficient = .45, ns; see row 3 of Table 2) even when not controlling for overall Bush approval, so overall approval could not have been mediator with respect to current terrorism performance assessments.

Next, we explored whether comparative prospective judgments mediated the relations between domain-specific approval ratings and intended vote choice (see Table 3). According to these results, presidential performance handling Iraq, the economy, and terrorism all had effects on their respective prospective performance assessments (Iraq coefficient = 2.70, $p < .001$; economy coefficient = 2.75, $p < .001$; terrorism coefficient = 2.53, $p < .001$; see row 1 of column 1, row 2 of column 4, and row 3 of column 7 in Table 3). Prospective evaluations of the candidates' performance on Iraq completely mediated the relation between judgments of President Bush's handling of the war and intended vote choice (reducing the coefficient from 1.26, $p < .001$, to .47, ns, see row 1 of Table 3; Sobel test = 5.94, $p < .001$). Prospective evaluations with respect to the economy acted as a *partial* mediator (reducing

Table 2 Testing overall approval as a mediator between domain-specific performance evaluations and vote choice

	Dependent variables: Iraq			Dependent variables: Economy			Dependent variables: Terrorism		
	Overall Bush approval	Vote choice	Vote choice	Overall Bush approval	Vote choice	Vote choice	Overall Bush approval	Vote choice	Vote choice
Current handling: Iraq	.28***	**1.26***	**.26**						
Current handling: Economy				.24***	**2.41***	**1.00****			
Current handling: Terrorism							.16***	**.45**	**–.21**
Overall Bush approval			4.26***			3.61***			3.99***
Prospective: Iraq				.35***	3.55***	2.81***	.32***	4.29***	3.36***
Prospective: Economy	.31***	4.07***	3.31***				.31***	4.04***	3.27***
Prospective: Terrorism	.21***	3.29***	2.66***	.18***	2.55***	3.34***			
Democrat	–.01	–.94**	–.83*	–.02	–1.15***	–.94*	–.02	–1.20***	–1.03**
Republican	.03***	1.01**	1.29***	.05***	1.52***	1.19***	.04***	1.09***	1.16**
Liberal	–.03*	–.09	.13	–.01	.06	.23	–.02*	–.11	.13
Conservative	.03***	.81**	.68*	.04***	.53^{+}	.43	.04***	.66*	.46
Age	.01	.48	.97	.18	4.66	1.36	–.03	2.28	.90
Age squared	–.08	–1.26	–1.09	–.25^{+}	–4.50	–.72	–.02	–2.46	–.35
Black	–.02	–.35	.02	–.02	–.14	–.05	–.03^{+}	–.39	–.18
Hispanic	.11**	1.87**	.75	.08^{+}	1.69**	1.16*	.10*	2.24***	1.31*
Other	.02	.06	–.06	–.01	–.33	–.34	.02	.22	–.12
Male	.00	–.37	–.32	.00	–.16	–.16	.00	–.16	–.17
High School	–.01	–.73	–1.03	–.01	–.72	–.70	.00	–.59	–.76
Associates	–.03	–.58	–.56	–.02	–.31	.10	–.01	–.10	.06
Some College	.01	–.10	–.41	.01	–.48	–.08	.01	.06	–.05
College	–.02	–.65	–.81	–.02	–.69	–.36	–.01	–.36	–.29
Graduate	–.03	–.93	–.97	–.03	–.66	–.45	–.02	–.68	–.51
Constant	.11***	–4.22***	–5.33***	.08**	–5.22***	–5.85***	.09**	–4.89***	–5.56***
R^2	.81			.80			.80		
Pseudo R^2		.85	.88		.83	.89		.87	.89
N	3,823	3,779	3,743	3,829	3,774	3,763	3,835	3,787	3,755

***$p < .001$; **$p < .01$; *$p < .05$; $^{+}p < .10$

Note: Registered voters only. All data are weighted. OLS regression used to predict overall Bush approval and logistic regression used to predict vote choice

Table 3 Testing comparative prospective performance evaluations as a mediator between domain-specific performance evaluations and vote choice

	Dependent variables: Iraq			Dependent variables: Economy			Dependent variables: Terrorism		
	Prospective: Iraq	Vote choice	Vote choice	Prospective: Economy	Vote choice	Vote choice	Prospective: Terrorism	Vote choice	Vote choice
Current handling: Iraq	2.70***	**1.26***	**.47**						
Current handling: Economy				2.75***	**2.41***	**1.51***			
Current handling: Terrorism							2.53***	**.45**	**.11**
Prospective: Iraq			3.14***	2.29***	3.55***	3.27***	3.86***	4.29***	3.40***
Prospective: Economy	2.32***	4.07***	3.83***			3.17***	2.01***	4.04***	3.90***
Prospective: Terrorism	3.80***	3.29***	1.58***	1.93***	2.55***	1.72***			1.63***
Democrat	–.12	–.94**	–1.14**	–.64**	–1.15***	–1.09**	–.33	–1.20***	–1.19**
Republican	.42	1.01**	1.14***	1.14***	1.52***	1.09***	.84**	1.09***	1.14***
Liberal	–.20	–.09	–.14	.26	.06	.04	–.37	–.11	–.12
Conservative	.56^{+}	.81**	.63*	.16	.53^{+}	.60^{+}	.03	.66*	.69*
Age	–5.38	.48	2.69	5.42^{+}	4.66	2.09	2.85	2.28	2.52
Age squared	4.23	–1.26	–3.31	–5.07	–4.50	–2.47	–2.38	–2.46	–3.10
Black	–.44	–.35	–.07	–.05	–.14	.09	–.58^{+}	–.39	–.08
Hispanic	–.10	1.87**	2.21**	–.06	1.69**	1.85*	–.15	2.24***	2.23**
Other	.20	.06	.03	–.84^{+}	–.33	–.09	.29	.22	.15
Male	–.38	–.37	–.27	.06	–.16	–.21	.59**	–.16	–.21
High School	.35	–.73	–.78	.34	–.72	–.79	.06	–.59	–.80
Associates	–.67	–.58	–.30	.37	–.31	–.31	.33	–.10	–.28
Some College	.07	–.10	–.06	.21	–.48	–.15	.25	.06	–.13
College	–.09	–.65	–.66	.25	–.69	–.70	.05	–.36	–.55
Graduate	–.15	–.93	–.97	.46	–.66	–.93	–.29	–.68	–.98
Constant	–3.21***	–4.22***	–4.77***	–5.17***	–5.22***	–5.12***	–4.19***	–4.89***	–4.89***
Pseudo R^2	.83	.85	.87	.75	.83	.88	.81	.87	.87
N	3,764	3,779	3,688	3,755	3,774	3,679	3,748	3,787	3,671

***$p < .001$; **$p < .01$; *$p < .05$; $^{+}p < .10$

Note: Registered voters only. All data are weighted. All coefficients from logistic regressions

the coefficient from 2.75, $p < .001$, to 1.51, $p < .001$; see row 2 of Table 3; Sobel test = 9.29, $p < .001$), meaning that evaluations of the President's handling of the economy had an unmediated effect on intended vote choice in addition to the mediated effect. Again, the President's handling of terrorism did not significantly predict intended vote choice even when not controlling for the posited mediator (coefficient = .45, ns, see row 3 of Table 3), so there was no relation for prospective evaluations about terrorism to mediate.

Taken together, this evidence is consistent with the claim that the impact of the President's Iraq performance on intended vote choice was *completely* mediated by *both* overall approval and prospective evaluations regarding Iraq. The vertical arrow in Fig. 1 proposes how this could have occurred: the President's performance in general might have been used to shape predictions of his handling of the three target domains in the future. That is, people may have used perceptions of President Bush's performance in specific domains to derive an overall evaluation of his performance, and they may then have used this general judgment (rather than the domain-specific evaluations) to derive predictions about the President's future performance.

We tested the possibility that comparative prospective judgments may have mediated the relation between overall Presidential approval and intended vote choice by comparing the coefficients for overall presidential approval in row 1 of the second and third columns of Table 4 (which displays the relevant three Baron and Kenny regressions). These coefficients are consistent with the conclusion of partial mediation, because the coefficient for overall Bush

Table 4 Testing comparative prospective performance evaluations as a mediator between overall approval and vote choice (Iraq domain)

	Prospective: Iraq	Vote choice	Vote choice
Overall Bush approval	3.87***	4.44***	3.80***
Prospective: Iraq			2.57***
Prospective: Economy	1.52***	3.31***	3.12***
Prospective: Terrorism	3.79***	2.80***	1.26**
Democrat	–.19	–.85*	–.99*
Republican	.40	1.27***	1.33***
Liberal	–.09	.12	.13
Conservative	.73*	.69*	.44
Age	–4.06	.47	1.59
Age squared	3.04	–.58	–1.48
Black	–.45	.01	.19
Hispanic	–1.43^{+}	.66	1.43*
Other	.01	–.08	–.11
Male	–.16	–.29	–.27
High School	.07	–1.03	–.94
Associates	–.75	–.53	–.02
Some College	–.31	–.42	–.28
College	–.27	–.73	–.62
Graduate	–.22	–.96	–.74
Constant	–3.72***	–5.33***	–5.70***
Pseudo R^2	.83	.88	.89
N	3,773	3,783	3,697

***$p < .001$, **$p < .01$, *$p < .05$, $^{+}p < .10$

Note: Registered voters only. All data are weighted. All coefficients from logistic regressions

approval dropped from 4.44, $p < .001$, when not controlling for prospective Iraq judgments, to 3.80, $p < .001$, when controlling for those judgments (Sobel test = 5.65, $p < .001$). Iraq performance assessments may therefore have shaped intended vote choice partly by influencing expectations about future performance.

Discussion

If an analyst had approached these data employing the conventional approach of including all possible predictors in a regression at once, Table 1's result would have led him or her to conclude that performance assessments of President Bush's handling of Iraq and terrorism had no impact of the outcome of the 2004 U.S. Presidential election. But in fact, the results shown in Tables 2 and 3 suggest that current assessments of Iraq performance had powerful indirect effects on intended votes, mediated both by overall current presidential performance assessments and prospective performance assessments. By contrast, Tables 2 and 3 suggest that current assessments of terrorism performance did not in fact influence intended votes. So all analytic steps were necessary to reach a defensible set of conclusions about the direct and indirect influences that may have been operating during this campaign.

Our findings are encouraging with regard to the notion that some judgments may mediate the impact of others in producing intended vote choices. Specifically, it seems that to tell a full story of how intended vote choices were formed in 2004—domain-specific performance evaluations may have shaped overall performance assessments and comparative prospective judgments, which may in turn have shaped intended vote choices. And overall performance assessments may have shaped comparative prospective judgments, thereby yielding mediated impact on intended vote choices. Thus, to accurately describe the impact of various judgments on election outcomes, ''kitchen sink'' regressions may be misleading—the levers to make vote choices switch may lie farther back in causal chains than direct effects reveal.

Our results are particularly interesting for understanding what took place in 2004. Of course, the data we have analyzed do not permit assessing the predictors of actions made in voting booths that year, because all of these surveys were conducted pre-election. Thus, we have illustrated the sources of vote intentions, not votes themselves. But our analyses are likely to have revealed the processes of candidate evaluation that unfolded ''online'' (see Lodge, McGraw, & Stroh, 1989) during the 2004 campaign. And these results suggest that assessments of President Bush's current handling of the Iraq war and the economy, as well as prospective judgments on those issues, did indeed play significant roles in driving vote choices. In contrast, current handling of terrorism seems not to have played such a role, and prospective judgments of likely performance in handling terrorism had notably weaker effects than did such judgments about Iraq and the economy. According to these results, then, the 2004 election outcome was much more about Iraq and the economy than it was about terrorism.

Evaluating the Dosage–Response Hypothesis of News Media Priming

Overview

Why would this have been so? Why did Iraq and the economy seem to dominate terrorism in shaping candidate preferences in 2004? One possible answer to these questions is offered by the news media priming hypothesis, which contends that the media are a powerful force in determining the considerations people use when judging the incumbent president. According to this hypothesis, when the news media pay a great deal of attention to an issue, people judge the President more based on how well he has handled that issue. Therefore, the ranking of issues in terms of their impact on 2004 intended votes may have been at least partly attributable to differences in the amount of news media attention accorded to those issues during the campaign.

The news media priming hypothesis has been supported by two forms of evidence: (1) laboratory experiments contrasting people who saw or read no news stories or programs about an issue with people who saw or read many such stories (e.g. Domke, 2001; Domke, Shah, & Wackman, 1998; Holbrook & Hill, 2005; Iyengar & Kinder 1987; Iyengar, Kinder, Peters, & Krosnick, 1984; Iyengar, Peters, & Kinder, 1982; Miller, 2004; Miller & Krosnick, 2000; Valentino, Hutchings, & White, 2002); and (2) surveys comparing the correlations among attitudes of Americans when an issue got little or no media attention with those correlations during a later period when the issue received a huge amount of national news coverage (e.g. Hester & Gibson, 2003; Kiousis, 2003; Krosnick & Brannon 1993; Krosnick & Kinder, 1990; Stoker, 1993). Thus, the independent variable (amount of media coverage) was varied between essentially zero and a large amount across samples of people. Both of these methods are valid approaches to assessing whether consumption of a large amount of news media coverage of a domain causes citizens to use that domain in judging the president to a greater extent than people who consumed no coverage.

However, a close examination of the language in these papers and wider discussions of their findings in print reveals an interesting leap being made by analysts: they make the bolder and broader claim that as news media attention to an issue increases, so does the weight that people attach to the domain when evaluating the President. That is, this language posits a *dosage–response hypothesis*. It is not merely that the presence of news media coverage that causes priming, but the *amount* of priming is presumed to increase monotonically with the *amount* of coverage. For example, Iyengar and Kinder (1987) said they wished to ''concentrate on the more elementary relationship between the magnitude of priming and the sheer amount of coverage'' (p. 69). According to Krosnick and Kinder (1990), ''The more attention the news pays to a particular domain—the more frequently it is primed—the more citizens will, according to the theory, incorporate what they know about that domain into their overall judgment of the president'' (p. 500). Stoker (1993) hypothesized that ''Any analysis of the public's reaction to media communication

must recognize the importance of levels of *exposure*...attitude change will be conditioned by the degree of exposure to change-inducing information'' (p. 197, italics author's). And Krosnick and Brannon (1993) posited: ''The greater one's dosage of media content, the more one should be influenced by it'' (p. 965).

But direct tests and evidence in support of the dosage hypothesis have been very rare indeed. Kronsick and Kinder (1990) and Stoker (1993), for example, never explicitly tested the presumed relationship between amount of coverage and degree of priming. Similarly, Miller and Krosnick's (2000) experiment involved respondents either receiving or not receiving exposure; the amount of exposure was not varied in any continuous fashion. In one set of relevant studies we uncovered, Iyengar and Kinder (1987) did manipulate the amount of exposure that their laboratory participants experienced. Yet remarkably, these investigators found no relation between the *number* of stories a person saw and the magnitude of his or her priming effect. And outside of the laboratory setting, Krosnick and Brannon (1993) found stronger priming effects among individuals who were *less* exposed to the news media than among people who were more exposed. Thus, some of the most visible tests of the dosage hypothesis have failed to confirm it, calling into question the language so often used to describe the priming effect in print.

In fact, though, no past studies, either in the laboratory or in the field, have tested the dosage claim most directly by exploring whether naturally-occurring increases and decreases of news coverage of an issue over time lead to parallel increases and decrease in the impact of that issue on presidential evaluations. We therefore set out to conduct such tests. Like all previous studies of priming using national surveys, we estimated cross-sectional regressions predicting overall presidential evaluations with domain-specific evaluations (again regarding Iraq, the economy, and terrorism) separately using the various available ABC News/*Washington Post* surveys. We also conducted content analyses of news media story topics in order to document the natural variation in issue attention during the 2004 campaign. And we explored whether this variation was especially strongly associated with variation in issue domain weights among people who received large doses of news media coverage during this time.

Results

Our first step in this investigation was to gauge the volume of news media attention to the three target issue domains at various points during the 2004 campaign and to assess the extent to which that volume varied over time. Of course, such variation in the posited independent variable is necessary in order to gauge its impact on the dependent variable (the weight that Americans attached to each issue when evaluating the President).

To do so, we tallied stories about each domain in two major national newspapers (*New York Times* and *Washington Post*) and in the news broadcasts of three national television networks (ABC, CBS, and NBC).

Specifically, using the Lexis-Nexis online database of searchable text, we counted the number of stories mentioning each issue during the week prior to the starting date of each survey.[3] The search terms used were: ''Iraq'' for the Iraq war, ''economy,'' ''unemployment,'' and ''inflation'' for the economy, and ''terrorism,'' ''terrorist,'' and ''Al Qaeda'' for terrorism. Staff research assistants (who did not know about the hypotheses to be tested or the regression coefficients we had obtained) counted the number of stories that Lexis-Nexis retrieved for each search. The correlations between newspaper and network news coverage volume over time were .43, .65, and .76 for Iraq, the economy, and terrorism, respectively. Thus, coverage dynamics were not equivalent for the two sets of media, so we combined them together to yield an aggregated portrait of news media focus during these weeks. The resulting story counts appear in Table 5.

During the period of the surveys, the Iraq war drew the most coverage, terrorism attracted the second most, and the fewest stories were written about the economy (see Table 5). Yet as the first three rows of the sixth column of Table 6 illustrates, this ordering of the three issues does not match their ordering in terms of impact of issue-specific evaluations of President Bush's performance on overall presidential performance assessments.[4] In a regression pooling the data from all waves, the Iraq coefficient ($b = .33$, $p < .001$) was significantly larger than the economy coefficient ($b = .25$, $p < .001$; $F(1, 4105) = 7.64$, $p = .006$), which was significantly larger than the terrorism coefficient ($b = .19$, $p < .001$; $F(1, 4105) = 5.83$, $p = .02$).[5] Thus, maximum coverage of Iraq matched maximum impact of Iraq evaluations, which is supportive of the dosage hypothesis. But whereas terrorism received more media attention than the economy, economic evaluations had more impact than did evaluations regarding terrorism.[6]

[3] We tested a series of different lags, all of which yielded similar but weaker results. Standard methodology for event analysis is to assume that the event window that produces the most significant difference is the correct one (see Rogerson, 1989), and we did so here.

[4] Variance inflation factors (VIFs) for the regression models in Table 6 vary between 5 and 6, suggesting the presence of multicollinearity. Multicollinearity does not bias coefficient estimates; it only inflates standard errors. But with our large samples, we had considerable power to detect real differences between coefficients. And the replication of similar coefficient patterns across surveys was reassuring about their validity. Most importantly, when we estimated the regression coefficients removing various terms from the equations, our substantive conclusions were consistently sustained, suggesting that our results do not reflect fragile and unstable parameter or significance level estimates. Finally, multicollinearity would not affect the over-time correlations of coefficient estimates and media content, which are presented in Table 8 and discussed below.

[5] Wald tests were used to test linear hypotheses (see Judge, Griffiths, Hill, Lutkepohl, & Lee, 1985, pp. 20–28). The Iraq coefficient was significantly larger than the economy coefficient only in the 5/20 and 9/23 waves ($p = .03$ and $p = .004$, respectively). The economy coefficient was significantly larger than the terrorism coefficient only in the 2/10 wave ($p = .01$).

[6] Other domains (e.g. education, health care) were also asked about in some individual waves. We included only Iraq, the economy, and terrorism to produce across-wave comparability, because these were the three domains asked about in *every* wave. When we re-estimated the regressions in Table 1 using all possible domains that were asked about, the rank ordering of Iraq, the economy, and terrorism remained unchanged.

Table 5 Media coverage of domains one week prior to the starting date of each survey

News source	Date						
	Feb. 3–9	May 13–19	Jul. 15–21	Aug. 19–25	Sep. 16–22	Correlation[a]	
Iraq							
Newspaper stories	191	361	171	166	242	}	$r = .43$
Television network news stories	103	149	99	134	211		
Total news coverage	294	510	270	300	453		
Economy							
Newspaper stories	88	97	85	76	94	}	$r = .65$
Television network news stories	16	17	12	4	8		
Total news coverage	104	114	97	80	102		
Terrorism							
Newspaper stories	209	235	226	235	313	}	$r = .76$
Television network news stories	56	44	89	53	114		
Total news coverage	265	279	315	288	427		

[a] Correlation across waves between newspaper and television volume

At this relatively crude level, then, we see initial evidence suggesting inconsistency between media coverage volume and issue weight accorded to particular issue domains. But such correspondence may be too much to expect, because the magnitude of issue impact may be determined by many factors in addition to news media volume. It may therefore be more suitable simply to assess whether fluctuation in news media attention to an issue corresponded with fluctuation in the weight attached to that issue when evaluating the President.

To conduct such a test, we must observe variation over time in the number of news stories devoted to an issue. Yet this was clearly not the case during this time period of 2004 for the economy: the number of news stories about the economy held steady between 97 and 114 per week (see the sixth row of numbers in Table 5). Therefore, we cannot use the economy to test for a relation between news volume and issue weight.

Inspection of the numbers in Table 5 for Iraq and terrorism suggest some more promising possibilities. Specifically, whereas total news stories about Iraq were 294, 270, and 300 in February, July, and August, respectively, story volume was 50% greater in September (453) and almost 70% greater in May (510, see row 3 of Table 5). Likewise, whereas terrorism coverage was fairly consistent in February, May, July, and August (265, 279, 315, and 288, respectively), it increased by about 50% to 427 in September. Thus, these two issues seem to offer two opportunities to test whether an increase in news coverage devoted to an issue (from a non-zero starting number to a larger number) yields an increase in weight attached to the issue.

To do so, we gauged wave-to-wave changes in weights by calculating z-statistics testing differences in the domain-specific approval coefficients

Table 6 OLS regressions predicting overall presidential job approval with domain-specific approval ratings

	Feb. 10–11	May 20–23	Jul. 22–25	Aug. 26–29	Sep. 23–26	Pooled	Pooled	High exposure	Low exposure
Bush current handling: Iraq	.33***	.38***	.26***	.31***	.37***	.33***	.32***	.37***	.30***
Bush current handling: Economy	.30***	.21***	.28***	.25***	.22***	.25***	.28***	.33***	.21***
Bush current handling: Terrorism	.17***	.19***	.20***	.21***	.18***	.19***	.18***	.16***	.20***
Time							.02		
Bush current handling: Iraq × Time							.00		
Bush current handling: Economy × Time							–.04		
Bush current handling: terrorism × Time							.02		
Democrat	–.01	–.02	–.02	–.04^{+}	–.07**	–.03**	–.03**	–.04*	–.03^{+}
Republican	.05*	.04*	.11***	.06**	.01	.06***	.06***	.04**	.06***
Liberal	.00	–.02	–.02	–.01	–.03	–.02	–.02	.00	–.02
Conservative	.00	.04^{+}	.03	.07**	.03^{+}	.03***	.03***	.03*	.03*
Age	–.10	.01	–.30	.12	.41^{+}	.06	.06	.09	.04
Age squared	.08	–.05	.30	–.16	–.52^{+}	–.10	–.10	–.11	–.08
Black	–.05	–.06^{+}	–.04	–.04	–.05	–.05**	–.05**	.01	–.08***
Hispanic	.04	.23^{+}	.07	.17*	.00	.10*	.10*	.11	.09^{+}
Other	–.03	.00	.05	–.02	.05	.01	.01	.05^{+}	–.02
Male	–.02	–.01	–.01	–.02	–.01	–.01	–.01	.00	–.02
High School	.12**	.00	–.03	.02	–.13**	.00	.00	–.02	.01
Associates	.07	–.01	–.06	–.02	–.16**	–.03	–.03	–.06	–.01
Some College	.09^{+}	.03	.00	.01	–.10*	.01	.01	–.02	.03
College	.09^{+}	–.03	–.03	.00	–.16**	–.02	–.02	–.05	–.01
Graduate	.05	–.04	–.02	–.02	–.14**	–.03	–.03	–.04	–.03
Constant	.03	.11*	.15*	.08	.23***	.11***	.10**	.08^{+}	.14***
R^2	.74	.76	.78	.81	.81	.77	.77	.85	.72
N	758	780	884	850	852	4,124	4,124	1,832	2,292
Mean VIF	5.49	5.29	5.13	5.22	5.33	5.15	6.67	5.85	4.89

***$p < .001$, **$p < .01$, *$p < .05$, $^{+}p < .10$

Note: Registered voters only. All data are weighted

Table 7 Changes in OLS regression coefficients between survey dates

	Feb. 10–May 20	May 20–Jul. 22	Jul. 22–Aug. 26	Aug. 26–Sep. 23
Full sample				
Bush current handling: Iraq	.05	–.12*	.05	.06
Bush current handling: Economy	–.08^{+}	.06	–.03	–.03
Bush current handling: Terrorism	.02	.01	.01	–.03
Close campaign followers				
Bush current handling: Iraq	.08	–.20	.09	.01
Bush current handling: Economy	–.12	.17	–.15	.02
Bush current handling: Terrorism	.02	.01	.00	.03
Not close campaign followers				
Bush current handling: Iraq	.05	–.11*	.05	.08
Bush current handling: Economy	–.07	.04	.00	–.04
Bush current handling: Terrorism	.02	.02	.01	–.08

*$p < .05$; $^{+}p < .11$ (two-tailed)

Note: Registered voters only. All data are weighted

between adjacent waves.[7] The coefficient for terrorism did not change significantly at all from wave to wave (see the third row of Table 7). This is inconsistent with the expectation that the terrorism coefficient would have become stronger in September, due to the 50% increase in media coverage of the issue at that time. Similarly, the coefficient for Iraq did not increase in September, despite the 50% increase in coverage volume at that time ($\Delta b = .06$, ns, see the last column of the first row of Table 7). And when coverage of Iraq surged by 70% in May, the coefficient for Iraq increased only slightly and not significantly ($\Delta b = .05$, ns, see the second column of the first row of Table 7). Clearly, there is not resounding support here for the dosage hypothesis regarding priming.

Another way to conduct this analysis uses all the months of data at once and correlate coverage volume with regression coefficient magnitude across the survey waves. The dosage hypothesis anticipates a significant positive correlation. As shown in the second row of Table 8, just such a positive correlation appeared for Iraq ($r = .89$, $p < .05$). Even the slight ups and downs in the regression coefficients shown in the first five columns of that row correspond remarkably closely to the ups and downs of media coverage volume in the first row of Table 8. This is, of course, encouraging regarding the dosage hypothesis. But the comparable correlations for the economy and terrorism (shown on the sixth and tenth rows of Table 8) are both negative and not significantly different from zero ($r = –.26$ and –.16, respectively). If we assume that no real change in coverage of the economy occurred over this time, then there is no reason to expect covariation of the fifth and sixth rows of Table 8. But there was real change in the volume of media coverage of terrorism across the months (see row 9 of Table 8), yet no corresponding changes in regression

[7] The z-statistic assessing across-wave changes in the coefficients β_1 and β_2 was simply calculated by: $\frac{\beta_1 - \beta_2}{\sqrt{\text{var}(\beta_1) + \text{var}(\beta_2)}}$.

Table 8 Correlations between media coverage and coefficients from OLS regression of overall presidential job approval on domain-specific approval ratings

	Feb. 3–9	May 13–19	Jul. 15–21	Aug. 19–25	Sep. 16–22	Correlation[a]
Iraq						
Total media coverage	294	510	270	300	453	
Domain weight coefficients: Total sample	.33	.38	.26	.31	.37	.89*
Domain weight coefficients: Close campaign followers	.38	.45	.25	.34	.35	.72^{+}
Domain weight coefficients: Not close campaign followers	.30	.36	.24	.29	.37	.91**
Economy						
Total media coverage	104	114	97	80	102	
Domain weight coefficients: Total sample	.30	.21	.28	.25	.22	–.26
Domain weight coefficients: Close campaign followers	.39	.27	.44	.28	.30	.01
Domain weight coefficients: Not close campaign followers	.25	.18	.22	.22	.19	–.38
Terrorism						
Total media coverage	265	279	315	288	427	
Domain weight coefficients: Total sample	.17	.19	.20	.21	.18	–.16
Domain weight coefficients: Close campaign followers	.14	.16	.17	.17	.20	.92**
Domain weight coefficients: Not close campaign followers	.18	.19	.21	.22	.13	–.77*
N (Total sample)	758	780	884	850	852	
N (Close campaign followers)	288	310	377	409	448	
N (Not close campaign followers)	470	470	507	441	404	

**$p < .001$; *$p < .05$; $^{+}p < .10$

Note: Registered voters only. All data are weighted. Close campaign followers defined as respondents who are following the campaign ''very closely''; all other respondents are considered not close campaign followers

[a] Correlation between total media coverage and domain weight coefficients

coefficients appeared for that issue (in row 10 of Table 8). Again, this challenges the dosage hypothesis.

One possible explanation for these mixed results has to do with exposure levels. Perhaps citizens who were highly exposed to the flow of news stories are influenced by their foci, whereas inattentive citizens were not. Therefore, we might find stronger correspondence of news media coverage with issue impact among the former individuals.

We first explored this possibility by assessing whether the domain rankings in terms of impact on overall evaluations matched the domain rankings in terms of volume of news media coverage most among people most exposed and attentive to political news. To differentiate highly exposed and attentive

people from less exposed and attentive people, we used answers to this question: ''How closely are you following the 2004 presidential race: very closely, somewhat closely, not too closely, or not closely at all?'' Respondents who reported that they followed the race ''very closely'' were considered to be high in exposure, and all other respondents were considered to have experienced low exposure levels.[8]

Disconfirming the dosage hypothesis, high media users manifested the same ordering of domain impact as did less interested citizens (see the top three rows of the final two columns of Table 6). For both groups, Iraq was the most consequential domain, followed by the economy and then by terrorism. Among high-exposure respondents, the terrorism coefficient ($b = .16$, $p < .001$) was significantly weaker than that of the economy ($b = .33$, $p < .001$) and of Iraq ($b = .37$, $p < .001$). But among low-exposure respondents, the Iraq coefficient ($b = .30$, $p < .001$) was significantly larger than those for the economy ($b = .21$, $p < .001$) and terrorism ($b = .20$, $p < .001$), which were not significantly different from one another. Thus, the ranking of issues differs from the ranking of domains in terms of news media coverage just as much in the high and low exposure groups, disconfirming a strict interpretation of the dosage hypothesis.

Next, we tested the dosage hypothesis's predictions regarding over-time shifts in the apparent impact of each domain. This hypothesis anticipates surges in the impact of Iraq in May and September (when Iraq coverage increased) and a surge in the impact of terrorism in September (when terrorism coverage increased) among the highly exposed citizens. Yet as shown in the middle and bottom panels of Table 7, no such surges in impact appeared (Iraq, May: $\Delta b = .08$, ns; Iraq, September: $\Delta b = .01$, ns, see columns 1 and 4 of the fourth row of Table 6; terrorism, September: $\Delta b = .03$, ns, see the last column of the sixth row of Table 7).

Finally, we tested the dosage hypothesis by seeing whether media coverage volume correlated more strongly with the coefficients across months among people high in media exposure than among people low in media exposure. Just such a pattern appeared for terrorism: the correlation for high exposure respondents was .92 ($p < .001$; see the last column of row 11 of Table 8), whereas the correlation for low exposure respondents was –.77 ($p < .05$, see the last column of row 12 of Table 8). And no correlations appeared in either group for the economy, which again might be expected based upon no real change in media coverage volume across the months ($r = .01$, ns, and –.38, ns, for the high and low exposure groups, see rows 7 and 8 of the last column of Table 8). But the expected difference between the high and low exposure groups for Iraq did not appear ($r = .72$, $p < .10$, and .91, $p < .001$, respectively). This may have occurred because even among low exposure respondents, the Iraq war was highly salient in the news during its heavy exposure times. Regardless of the explanation, some of our findings here are consistent with the predictions of the dosage hypothesis, whereas others are not.

[8] We dichotomized in this fashion to yield two groups of about equal size.

Discussion

Why did Americans place the most weight on Iraq performance when making overall presidential approval judgments, less weight on the economy, and still less weight on terrorism across the 2004 campaign? When we tested the dosage hypothesis to explain this pattern, we found some evidence supportive of the theory, but other evidence less consistent with this hypothesis. News media coverage volume did not perfectly correspond with the rank ordering of predictive power of each of the three domains, either cross-sectionally or longitudinally. However, in both analyses, some findings on variation in the coverage of Iraq and lack of variation in coverage of the economy were consistent with the dosage hypothesis. We did not find consistent evidence of better correspondence between media coverage and issue weight among people who were more attentive to the flow of political news, perhaps because of the salience of these issues leading up to a presidential election. In sum, some of our findings are consistent with the dosage hypothesis, but others are not.

Considering the great deal of past evidence that has offered support for the news media priming hypothesis, what explains the inconsistencies we observed? We have used exactly the same estimation approach used in all past studies of priming (cross-sectional regressions predicting overall presidential performance using issue-specific performance assessments), so that cannot explain the discrepancy. Consequently, it makes sense to consider other possible explanations.

One such explanation has to do with the magnitude of variation in coverage volume that we observed. Although we certainly saw a few substantial changes from wave to wave in the amount of attention paid by the media to Iraq and terrorism, even these shifts could be considered to be relatively small. We never examined a time period in which any of these issues received no attention at all. In their lab experiments, Iyengar et al. (1984) consistently found that priming effects were no different depending on whether participants saw a moderate volume of issue coverage or a large amount. Any coverage at all was sufficient to induce priming. And past non-lab documentations of priming have involved much more dramatic shifts in coverage volume over time than we observed for our three target issues during 2004 (e.g., Krosnick & Kinder, 1990). So perhaps the range of coverage volume of these issues was too narrow to produce changes in weights. If correct, that explanation has important implications for the normative significance of priming, because it may happen much more rarely than many authors have presumed—only when huge shifts in coverage volume occur.

A second possible explanation involves the unprecedented level of interest in politics that typified the American public during the 2004 presidential election campaign. Many data sources showed that Americans' cognitive engagement in politics surged to a relatively high level unusually early in the 2004 campaign (Abramowitz & Stone, 2006; Patterson, 2005; Pew Research Center for People and the Press, 2004; Wattenberg, 2005). Perhaps this high

level of thinking about the candidates early on led people to crystallize the weights they attached to issue domains when evaluating the President, and these crystallized weights may have remained fairly stable across the campaign as a result.

Another possible explanation for our failure to find evidence supporting the dosage hypothesis is suggested by the recent work of Miller and Krosnick (2000). They found that news media priming effects appeared only among people with two characteristics: high political expertise and high trust in the news media to provide accurate, unbiased coverage. In the surveys we analyzed, no measures of political expertise or trust in the media were included, so we could not explore this possibility. Perhaps with such measures in hand, we would have uncovered more compelling evidence of priming during the 2004 campaign.

Although we could not assess moderation of priming using these two variables, we did explore whether priming effects were stronger among people who received a larger dose of political news. The surveys we analyzed did not include a direct measure of respondent exposure and attentiveness to news coverage, so we tested the dosage hypothesis using a measure of attention paid to the presidential campaign. According to 2004 American National Election Study survey data, the correlation between closeness following the campaign and a composite measuring exposure and attention to news about politics generally was .57. This is a reasonably strong correlation, suggesting that our ''campaign following'' measure may have been effective at identifying those people who were most exposed to news stories. But it is also possible that our tests of moderation may have been more successful if we had a more direct measure of exposure and attention to all news.

The failure of news media priming to have occurred with regard to terrorism is surprising in light of widespread popular speculation that the Bush administration activated Americans' fear of future attacks for political benefit. One possible explanation is suggested by Miller and Krosnick's (2000) theoretical account of priming, pointing to assessments of national importance as the mediator. Perhaps terrorism had been on the national radar for such a sustained period of time that Americans had fairly well crystallized view of the national importance of the problem, so relatively small changes in the volume of news media coverage of it did nothing to alter those beliefs. Thus, perhaps the perceived national importance of terrorism surged just after the September 11th attacks, but national importance judgments may have remained relatively constant in Americans' minds over the ensuing years. We look forward to seeing new tests of news media priming during 2004 using other datasets and other measures of exposure to see whether an explanation for our findings regarding terrorism might be validated.

Another possibility is that the media coverage of terrorism during the 2004 election was positively correlated with coverage of the Iraq war (in our data, $r = .35$). Given the Bush Administration's efforts to link the two issues together (Gershkoff & Kushner, 2005), it is possible that the media's discussion of one issue may have primed the other in Americans' minds.

However, we replicated our analyses in Table 8 examining the *sum* of media coverage for Iraq and terrorism, and the correlations were similar to those reported in the table for both domains and for both heavy and light media users. Nevertheless, a comprehensive analysis of overlapping domains requires a further gradation of news media content, a fruitful avenue for future research.

Until such evidence is generated, it seems best to conclude that there are some indications that news media priming may have occurred with regard to Iraq, the economy, and terrorism during the 2004 campaign in response to the shifts in media coverage volume that we documented. But if the theory of news media priming cannot tell us the whole story, then we are still left with lingering questions about why citizens may have put more weight on Iraq than the economy, and more weight on the economy than terrorism when deriving overall presidential performance assessments. Our own inclination in this regard is to recognize that during a presidential election campaign, news media messages about issues exist in an environment in which citizens are also prompted to think about issues by paid advertisements and by conversations with other people. Furthermore, individuals may be quite thoughtful about how much weight to place on issues during times when a lot seems to be at stake in the world of politics, as is the case when the nation is at war. As a result, people may make weighting decisions based upon their own principles about national importance and/or presidential responsibility rather than based simply upon the volume of attention to a domain in the media or any other stream of public communication.

Conclusions

This investigation has shed light on many aspects of how Americans evaluated George W. Bush and John Kerry during the 2004 presidential election campaign. First, we saw that perceptions of President Bush's handling of the war in Iraq, the economy, and terrorism all had impact on vote choices, partly mediated by overall assessments of his job performance and by guesses about how well he and Senator Kerry would handle these issues in the future. We saw as well that these domain-specific prospective assessments were derived in part from perceptions of the President's handling of these domains in the past and partly by perceptions of his overall performance. Perhaps most interestingly, Iraq seemed to carry the most weight in driving overall performance evaluations and vote choices, whereas the other two domains had less impact.

We also found no evidence to support the claim that this ordering of issues was attributable to variation in news media coverage according to the news media priming hypothesis. The priming hypothesis does not claim to account for all variation in issue impact on presidential approval, so there is no reason to expect to see a perfect match here. But the fact that we typically found no match at all between media coverage and issue weight in the many analyses we conducted suggests that in 2004, the weights that American voters placed

on various policy domains when evaluating President Bush were determined by factors other than news media volume.

It is important to note that the evidence reported here is not meant to portray the impact of the three performance domains on votes on election day in 2004, for a number of reasons. First, we did not estimate any regression equations predicting vote choices made on election day. Rather, we focused exclusively on the processes by which citizens constructed intended vote choices prior to the election. And instead of focusing on only people who voted (or were likely to do so), our analyses examined the judgments of all Americans registered to vote. So it would be inappropriate to infer that our results indicate something about the weights attached to these domains on election day.

Furthermore, because the surveys we analyzed were all independent cross-sections with no reinterviewing of the same respondents and involved relatively short questionnaires, we could not take analytic steps to discern the directions of causality underlying the relations we examined. We are inclined to assume that at least some causality flows in the directions illustrated in Fig. 1. And our tests of mediational hypotheses indirectly assessed the plausibility of these hypotheses. But it seems quite possible that causality runs to at least some degree in the reverse direction in some cases. For example, once a voter decides that the President is doing a good job in general, that may color the voter's assessments of the President's performance in particular domains through a process of rationalization (see, e.g., Festinger, 1957), akin to the famous ''halo effect'' (Thorndike, 1920). We look forward to future research using data that permit explicit testing of hypotheses regarding directions of causality. But in the meantime, we can take steps to interpret the results reported here as if the causal flows in Fig. 1 are plausible, knowing that some of the observed relations between variables may reflect reverse causal influences as well.

Another limitation of the present data for our purposes involves the measurement of some key variables. Although overall presidential performance evaluations were assessed on a four-point rating scale, domain-specific performance assessments were assessed as dichotomies, as were prospective comparative performance assessments. We suspect that all of these judgments are in fact made in people's minds on continua. By measuring only dichotomies, we have differentiated people above and below the midpoint but have lumped together members of these two groups, ignoring some of the variance in the underlying judgments. Consequently, our regression coefficients may gauge the impact of only a portion of the real variance in these judgments. It is impossible to anticipate how our results would have been different had we had continuous measures of all constructs. We look forward to future studies using continuous measures.

Coda

We hope that the findings reported here set the stage for future work on the role that domain-specific candidate appraisals play in determining election

outcomes. We have shown that careful attention to mediational pathways can reveal indirect effects that would be hidden when investigators do not carefully consider such possibilities before putting many predictors into a single regression equation. We hope that our documentation of the Baron and Kenny analytic approach and our illustration of the sorts of results that it can produce will inspire more voting researchers to take on the problem of mediation seriously in future work.

In light of all this, it is interesting to return to the paper by Hetherington (1996) that we mentioned at the outset was unusually attentive to mediation. Although he did a nice job positing and testing mediation, one aspect of his analyses looks different in retrospect. When exploring the impact of media exposure on perceptions of national economic condition and of this latter variable on voting, Hetherington controlled for overall attitudes toward President Bush (measured with a feeling thermometer). Yet according to the causal models posited by both Page and Jones (1979) and Markus and Converse (1979), overall attitudes toward the candidates are the most proximate determinant of vote choices. That is, citizens are presumed to vote for the candidate they evaluate most favorable.

Thus, in retrospect, it seems that Hetherington controlled for the principal ingredient of his ultimate dependent variable when searching for its prior causes. As a result, the only variance in perceptions of national economic conditions that media consumption could explain in Hetherington's regression was variance not shared with candidate evaluations or vote choices (if candidate evaluations do indeed fully mediate the impact of predictors on votes). And the only variance in vote choices that perceptions of national economic conditions could explain was variance not shared with candidate evaluations. So perhaps even Hetherington's regressions might have yielded different results if a likely mediator had not been included among the predictors and if indirect effects via candidate evaluations had been explicitly modeled. We should all do more of this in the future, and if we do, we are likely to push the study of voting up a notch or two in sophistication and conceptual payoffs.

Acknowledgment We thank Gary Langer, Dan Merkle, and Jon Cohen of ABC News for sharing these data with us and for their helpful comments on the manuscript. Jon Krosnick is University Fellow at Resources for the Future.

References

Abramowitz, A. I., Lanoue, D. J., & Ramesh, S. (1988). Economic conditions, causal attributions, and political evaluations in the 1984 presidential election. *Journal of Politics, 50*, 848–863.

Abramowitz, A. I., & Stone, W. J. (2006). The Bush effect: Polarization, turnout, and activism in the 2004 presidential election. *Presidential Studies Quarterly, 36*, 141–154.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. New York: Wiley.
Conover, P. J. (1985). The impact of group economic interests on political evaluations. *American Politics Quarterly, 13,* 139–166.
Domke, D. (2001). Racial cues and political ideology: An examination of associative priming. *Communication Research, 28*, 772–801.
Domke, D, Shah, D. V., & Wackman, D. B. (1998). Media priming effects: Accessibility, association, and activation. *International Journal of Public Opinion Research, 10*, 51–74.
Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
Fiorina, M. P. (1981). *Retrospective voting in American national elections*. New Haven, CT: Yale University Press.
Gerber, E. R., & Jackson, J. E. (1993). Endogenous preferences and the study of institutions. *American Political Science Review, 87*, 639–656.
Gershkoff, A., & Kushner, S. (2005). Shaping public opinion: The 9/11-Iraq connection in the Bush administration's rhetoric. *Perspectives on Politics, 3*, 525–537.
Hester, J. B., & Gibson, R. (2003). The economy and second-level agenda setting: A time series analysis of economic news and public opinion about the economy. *Journalism & Mass Communication Quarterly, 80*, 73–90.
Hetherington, M. J. (1996). The media's role in forming voters' national economic evaluations in 1992. *American Journal of Political Science, 40*, 372–395.
Holbrook, R. A., & Hill, T. G. (2005). Agenda-setting and priming in prime time television: Crime dramas as political cues. *Political Communication, 22*, 277–295.
Iyengar, S., & Kinder, D. R. (1987). *News that matters*. Chicago: University of Chicago Press.
Iyengar, S., Kinder, D. R., Peters, M. D., & Krosnick, J. A. (1984). The evening news and presidential evaluations. *Journal of Personality and Social Psychology, 46*, 778–787.
Iyengar, S., Peters, M. D., & Kinder, D. R. (1982). Experimental demonstrations of the 'not-So-minimal' consequences of television news programs. *American Political Science Review, 76*, 848–858.
Jennings, M. K., & Markus, G. B. (1984). Partisan orientations over the long haul: Results from the three-wave political socialization panel study. *American Political Science Review, 78*, 1000–1018.
Judge, G. G., Griffiths, W. E., Hill, R. C., Lutkepohl, H., & Lee, T. C. (1985). *The theory and practice of econometrics* (2nd ed.). New York: John Wiley & Sons.
Kenney, P. J., & Rice, T. W. (1988). Presidential prenomination preferences and candidate evaluations. *American Political Science Review, 82*, 1309–1319.
Kinder, D. R., Adams, G. S., & Gronke, P. W. (1989). Economics and politics in the 1984 American presidential election. *American Journal of Political Science, 33*, 491–515.
Kinder, D. R., & Kiewiet, D. R. (1979). Economic discontent and political behavior: The role of personal grievances and collective economic judgments in congressional voting. *American Journal of Political Science, 23*, 495–527.
Kinder, D. R., & Kiewiet, D. R. (1981). Sociotropic politics: The American case. *British Journal of Political Science, 11*, 129–161.
Kiousis, S. (2003). Job approval and favorability: The impact of media attention to the Monica Lewinsky scandal on public opinion of President Bill Clinton. *Mass Communication & Society, 6*, 435–451.
Krosnick, J. A., & Brannon, L. A. (1993). The impact of the gulf war on the ingredients of presidential evaluations: Multidimensional effects of political involvement. *American Political Science Review, 87*, 963–975.
Krosnick, J. A., & Kinder, D. R. (1990). Altering the foundations of popular support for the president through priming: Reagan and the Iran-contra affair. *American Political Science Review, 84*, 495–512.
Langer, G. E., & Merkle, D. M. (2006). ABC news' polling methodology and standards. http://abcnews.go.com/US/PollVault/story?id=145373&page=1 (June 9, 2006).
Lavrakas, P. J. (1993). *Telephone survey methods: Sampling, selection, and supervision*. Thousand Oaks, CA: Sage.
Lodge, M., McGraw, K. M., & Stroh, P. (1989). An impression-driven model of candidate evaluation. *American Political Science Review, 83*, 399–419.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*, 99–128.

Markus, G. B., & Converse, P. E. (1979). A dynamic simultaneous equation model of electoral choice. *American Political Science Review, 73*, 1055–1070.

Miller, J. M. (2004). Media agenda-setting: Does content matter? *Presented at the annual meeting of the Midwest Political Science Association*, Chicago, IL.

Miller, J. M., & Krosnick, J. A. (2000). News media impact on the ingredients of presidential evaluations: Politically knowledgeable citizens are guided by a trusted source. *American Journal of Political Science, 44*, 295–309.

Murdoch, M., Van Ryn, M., Hodges, J., & Cowper, D. (2005). Mitigating effect of department of veterans affairs disability benefits for post-traumatic stress disorder on low income. *Military Medicine, 170*, 137–140.

Needham, B. L., & Crosnoe, R. (2005). Overweight status and depressive symptoms during adolescence. *Journal of Adolescent Health, 36*, 48–55.

Page, B. I., & Jones, C. C. (1979). Reciprocal effects of policy preferences, party loyalties, and the vote. *American Political Science Review, 73*, 1071–1089.

Patterson, T. E. (2005). Young voters and the 2004 election. Report, Joan Shorenstein Center on the Press, Politics, and Public Policy.

Pew Research Center for People and the Press. (2004). Voters more engaged, but campaign gets lukewarm ratings. Survey report.

Rogerson, W. P. (1989). Profit regulation of defense contractors and prizes for innovation. *Journal of Political Economy, 97*, 1284–1305.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*, 422–445.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology, 13*, 290–303.

Stoker, L. (1993). Judging presidential character: The demise of Gary Hart. *Political Behavior, 15*, 193–223.

Thorndike, E. L. (1920). A constant error on psychological rating. *Journal of Applied Psychology, 4*, 25–29.

Valentino, N. A., Hutchings, V. L., & White, I. K. (2002). Cues that matter: How political ads prime racial attitudes during campaigns. *American Political Science Review, 96*(1), 75–90.

Wattenberg, M. P. (2005). Turnout in the 2004 presidential election. *Presidential Studies Quarterly, 35*, 138–146.

Zamboanga, B. L., Bean, J. L., Pietras, A. C., & Pabon, L. C. (2005). Subjective evaluations of alcohol expectancies and their relevance to drinking game involvement in female college students. *Journal of Adolescent Health, 37*, 77–80.