Public Opinion, the War in Iraq, and Presidential Accountability

Erik Voeten
Assistant Professor of Political Science and International Affairs
The George Washington University
voeten@gwu.edu

Paul R. Brewer
Assistant Professor of Journalism and Mass Communication
University of Wisconsin-Milwaukee
prbrewer@uwm.edu

Earlier versions of this manuscript were presented at the 2004 PolMeth Conference at Stanford University, the Annual Meeting of the American Political Science Association, Chicago September 2 2004, and seminars at the University of Maryland and Dartmouth College. We thank the participants in those seminars and Lee Sigelman for many useful suggestions. We also thank Harlan Geer and Ryan Schoen for valuable research assistance.
Public Opinion, the War in Iraq, and Presidential Accountability

Abstract

How do citizens hold their leader accountable during an ongoing war? We distinguish between two models of accountability—the “decision-maker” and “managerial” models—and investigate their implications in the context of the current war in Iraq. We employ a novel measurement model and a database of survey marginals to estimate weekly time-series of aggregate beliefs about various aspects of the war. Consistent with the “decision-maker” model, we find that shifts in aggregate beliefs about whether the war was “worth it” have a greater impact on presidential approval than do equivalent shifts in perceptions of war success or approval of the president’s handling of the war. Conversely, aggregate perceptions of the war’s success are much more responsive to casualties and key events than are beliefs that the war was “worth it.” This suggests that the link from casualties and events to presidential approval is less direct than previously assumed.
Public Opinion, the War in Iraq, and Presidential Accountability

Leaders in democracies differ from their authoritarian counterparts in the conduct of war in that they require higher levels of consent from the populace to initiate and prolong wars. This need helps to explain a broad array of empirical regularities regarding the differential war behavior of democracies. For example, democracies tend to pick fights that they are likely to win quickly even though they are more likely than authoritarian states to accept draws or even defeats as a war goes on (Reiter and Stam 2002). Such patterns suggest a broader and more continuous conception of democratic accountability than one based solely on voters’ use of elections to reward incumbents for triumphant wars and punish them for unsuccessful military adventures (e.g. Bueno de Mesquita et al. 2003, Gaubatz 1999). In this more extensive view, democratic leaders rely on “contemporary consent”: they need high levels of public support to initiate a war and must maintain that support to carry on a war (Reiter and Stam 2002). This assumes a relation between public opinion and policy not unlike that of the “dynamic representation” model for domestic policy (Stimson et al 1995).

How, then, does the public judge a leader during wartime when the outcome of that war is yet unclear? The connection between news from the front and the performance of the incumbent is often ambiguous, providing elite and media discourse with considerable room to shape the formation of political judgments (e.g., Brody 1991, Kull et al 2004, Page and Shapiro 1992, Zaller 1992). Thus, the effects of war events and casualty reports on political judgments may flow through citizens’ beliefs about the war at hand. This logic suggests that we should look beyond the direct links between events, casualties, and evaluations of leaders that scholars have traditionally examined (see
Mueller 1970; 1973 for the seminal analyses; see Berinksy 2004 for an overview). It also raises an additional question: what kinds of beliefs about a war affect evaluations of a commander in chief?

We build on theories of accountability for economic policies to develop two models of public accountability during wartime: a “managerial” and a “decision-maker” model. We then evaluate some of the most important empirical implications of these models using a newly created database on aggregate public opinion towards the war in Iraq. Our findings indicate that events and casualty reports affect perceptions of the war’s success but have relatively little direct impact on beliefs that the “war was worth it.” Conversely, shifts in aggregate beliefs that the “war was worth it” more strongly shape presidential approval ratings than do swings in beliefs about the war’s success or even whether President George W. Bush “is doing a good job” in handling Iraq. We take this as evidence that citizens primarily hold the president accountable for his perceived decision-making qualities.

**Accountability and War**

Most theoretical analyses of how the public holds incumbents accountable for their deeds in office have focused on economic performance (e.g. Barro 1973, Ferejohn 1986, Keech 1995). The main assumption in these models is that, all else being equal, citizens prefer incumbents who are competent economic policymakers to those who are not. Competency is, however, not directly observable. Instead, citizens use their appraisals of discernable economic outcomes to indirectly hold incumbents accountable for their aptitude in managing the economy.
Precisely what perceptions about the economy are politically consequential is a matter of fierce debate in the literature (e.g. MacKuen et al. 1992, Keech 1995). These debates turn around a variety of questions. Do citizens draw political consequences based on the state of their personal financial circumstances or the macro-economy? Do they use prospective or retrospective evaluations? Are voters able to tell favorable economic outcomes created through genuine craftsmanship from those produced through temporary manipulation of economic policy instruments?

Similar issues arise in holding leaders accountable for their competency in security issues. Leaders, and certainly American presidents, have considerable discretionary powers to wage and manage wars. All else being equal, citizens prefer a leader with proven abilities to make good decisions in times of crisis and to manage wars well. The competency of a leader is, however, not directly observable and can only be inferred from outcomes. In the long run, mounting casualties and battlefield defeats ought to undermine beliefs among the public that an incumbent is a competent war leader. In the short run, however, it is not at all clear how citizens will draw inferences from individual outcomes that are only opaquely related to a leader’s competence.

Hence, the accountability process is likely to be indirect: citizens form impressions about the war and then use those in their evaluations of the incumbent. This calls into question a traditional mode of analysis that examines the direct impact of events and casualties on presidential evaluations (e.g., Mueller 1973, Gelpi et al. 1994). Instead, we expect that after controlling for subjective appraisals of the war, casualties and important events will have no direct impact on aggregate support for the president. A similar argument has been advanced and validated empirically in the literature on
economic voting: the impact of changes in objective economic outcomes on changes in presidential evaluations disappears after controlling for changes in subjective evaluations of the state of the economy (MacKuen et al. 1992).

But what kinds of appraisals of a war are politically relevant? Continuing the analogy with economic voting, we may suspect that what matters are citizens’ perceptions of how well the war is going. Based on such assessments, citizens may conclude that the incumbent is (or is not) sufficiently capable of managing wars.¹ A potential downfall of this model is that the link between immediate war success and presidential competence is ambiguous. Though the president, as the commander in chief, is ultimately responsible for managing a war successfully, the actual success of the war effort depends on other factors as well, such as the efforts by other administration officials, the military, foreign coalition partners, and fortune.

The president alone, however, is undeniably responsible for making the decision to launch a war. The obvious form of evidence for whether a leader is competent in this regard is his or her decision-making record. Accordingly, citizens may adjust their evaluations of the incumbent if they change their beliefs about whether a past decision to go to war was right in the first place. Building on this notion, the “decision-maker” model holds that citizens above all else seek a leader whom they perceive as a wise maker of decisions about whether the use of force enhances the security of the country. This model

¹ A variant to this “managerial” account is that what matters are prospective evaluations of success: i.e., the perceived likelihood that a military operation will succeed (Kull and Ramsay 2001, Feaver and Gelpi 2004, Gelpi, Feaver, and Reifler 2004). It is, however, unclear how prospective assessments would differ from contemporaneous evaluations of war success. The literature on prospective voting in economics is motivated by the notion that voters form rational expectations about future economic performance based on the cyclical nature of the economy (e.g., Keech 1995). There is no equivalent theory about future war success.
fits an important empirical regularity: namely, that democracies are disproportionately successful in wars not so much because they have greater resources or pick better strategies but mostly because they generally initiate wars against targets that they can beat relatively quickly (Reiter and Stam 2002). This may be because public opinion constrains leaders to pick their wars wisely. Nevertheless, a leader has considerable private information when initiating a war. As a war goes on, elites may debate whether new information about progress in the war validates the wisdom and sincerity of this initial decision. As the balance of elite discourse shifts, so does aggregate support for the war (Zaller 1992) and, if the “decision-maker model” holds, aggregate support for the leader.

Elite discourse about any given war is, of course, to some degree constrained by the actual things that happen. Yet some types of evaluations of a war are related in a less ambiguous way to events than others and hence less easily manipulated by elites. Reports of mounting casualties or battlefield losses and wins should have relatively straightforward implications for perceptions of how well a war is going, but their implications for war support and presidential evaluations are less clear. For example, Democratic challenger John F. Kerry was quick to call into question President Bush’s abilities as Commander in Chief based on the disappearance of 380 tons of high explosives in Iraq, whereas the White House immediately sought to absolve Bush from any responsibility by claiming that the weaponry may have disappeared before the arrival of American troops.2 No one, however, could have credibly claimed that the loss of such

---

large quantities of weaponry to resistance forces was evidence of success in the war, even if elite messages could (and did) differ on the extent to which it showed incompetence.

Further obscuring the relationship between news from the front and conclusions about decision-making skills are the incentives that leaders face to start wars for opportunistic reasons. A leader may want to divert attention from a poorly performing economy or some other domestic crisis (e.g., Stoll 1984, Russett 1990, Downs and Rocke 1994). A leader may also engage in strategic brinkmanship by inducing a crisis to signal competence in security affairs.\(^3\) This signal may be credible if the risky military adventure is such that only a competent executive would undertake it, thereby constraining a candidate whose competence is questionable. For example, President Bush presented himself in the 2004 election campaign as a leader with the courage and strength to go it alone.\(^4\) The war in Iraq lent credibility to his message. The point here is not to suggest that the war was fought for opportunistic reasons but rather that citizens may be forced to assess the sincerity of a decision to go to war. This provides elites with incentives to question—and, in turn, defend—the motives of the incumbent.

To summarize, we expect that, compared to perceptions of its success, support for the war in Iraq is shaped to a greater extent by relatively unconstrained elite discourse and hence to a lesser extent by events and casualty reports. As a consequence, aggregate opinion about the war’s success is likely to be more volatile than opinion about its merits. In this sense, perceptions of the war’s success are more likely to resemble Almond’s

\(^3\) This rationale is derived from models of signaling competence through fiscal policy (Rogoff and Sibert 1988, Rogoff 1990). It is based on the criticism that traditional political business cycle accounts take insufficient account of the rational expectations voters have regarding pre-election tax cuts, increased spending, or (by extension) wars (see also Hess and Orphanides 1995).

(1960) depiction of public opinion on foreign policy as a capricious “mood.” On the other hand, we anticipate that swings in this mood are relatively inconsequential politically in that we expect shifts in aggregate support for the president to follow shifts in aggregate beliefs that the “war was worth it” more closely than they follow shifts in aggregate perceptions of war success. Moreover, if citizens are truly first and foremost interested in the decision-making skills rather than the management skills of a leader, then we also expect that shifts in support for the war are a stronger predictor for swings in overall presidential job approval than are shifts in a more direct measure of perceived managerial abilities: perceptions of how well Bush is handling Iraq. This is a stringent test given the virtual identical phrasing of items that measure overall Bush job approval and his job performance on Iraq.

**Data and Method**

In testing these claims, we face a choice between individual- and aggregate-level analysis that has long been the subject of another debate in the economic voting literature. Numerous studies have applied the former approach to the study of economic voting (e.g., Kinder and Kiewet 1979). Others, however, have warned of the potential for endogenously induced bias in such analyses (e.g., Kramer 1983; Erikson 2004). The same issue extends to the domain at hand. At the individual level, war support and perceptions of war success are difficult to disentangle (Berinsky 2004). Projection effects tend to frustrate efforts to establish why citizens hold certain opinions. For example, an individual who likes the president for other (partisan) reasons may be more likely both to believe that the war is going well and to believe that the war was worth it, even without much information about the war (on the latter point, see Kull et al 2004).
We circumvent this issue by analyzing changes in aggregate beliefs and controlling for the most obvious alternative explanation that may lead to shifts in assessments of the president: perceptions of the economy. To be sure, our approach opens us to some potential problems characteristic to analyses of aggregate public opinion data and prohibits us from drawing inferences about individual variation in opinion formation (see Berinsky 1999; 2002). A sizable body of recent research has argued, however, that analyses of aggregate public opinion data can shed light on how various aspects of public opinion are related to events, to one another, and to public policy (e.g., Page and Shapiro 1992; Stimson et al. 1995). Moreover, aggregate-level analysis suits our substantive research objectives well: politicians care primarily about the aggregation of individual opinions and the extent to which movement in aggregate opinion carries over to evaluations of their personal performances.

An important departure from previous studies is that we estimate a model that allows us to use multiple indicators to measure the same concept. This approach increases the points in time for which we have observations. Moreover, it makes the analysis less dependent on idiosyncrasies such as the question wording of the most frequently asked item and less likely to mistake outliers due to sampling error for temporary shifts in opinion. Our approach does, however, require the assembly of a database of survey marginals and the development of a measurement model to estimate a common time series from these diverse indicators.

A Database of Survey Marginals
The assembly of response frequency data from assorted surveys has been facilitated by organizations that maintain on-line archives of polling results, often sorted by issue area.
We used two such sources: The Polling Report\(^5\) and the *National Journal’s* Poll Track.\(^6\)

The main criteria for including a question in our database were that it had to provide an unambiguous measurement of one of our concepts of interest and be asked at least five times. For each question, we recorded dates, sample size, survey organization, question wording, answer categories, and marginals for each answer category. Most surveys were conducted over a period of three to four days. We assigned the survey to a week by the mid-date of fieldwork. The data run until the last week of October, also the last week before the 2004 election. Table 1 provides an overview of the data that we collected for each time series. The full data are available from [website deleted].

**TABLE 1 ABOUT HERE**

We have 323 survey items that tap general war support spanning more than two years. The most frequently asked item (32 times) was the CNN/Gallup question, “All in all, do you think the situation in Iraq is/was worth going to war over, or not?” Eight other items were asked at least fifteen times. Some questions were asked both before and after the start of the war with a simple change of verb. The data include two sets of questions that make war support explicitly contingent on some other factor. Twenty-three observations condition war approval on the use of ground troops,\(^7\) and 52 observations explicitly mention American casualties.\(^8\) As explained in the methods section, we correct for the possibility that these items tap differently into our concept of interest.

---

5 [www.pollingreport.com](http://www.pollingreport.com)

6 [http://nationaljournal.com/](http://nationaljournal.com/)

7 Most frequently asked (13 times) was the CNN/Gallup question “Would you favor or oppose invading Iraq with United States ground troops in an attempt to remove Saddam Hussein from power?”

8 Most frequently asked (18 times) was the CBS/*New York Times* item “Do you think the result of the war with Iraq was worth the loss of American life and other costs of attacking Iraq or not?”
The time series for approval of Bush on Iraq begins in mid-January 2003 as too few organizations asked pertinent questions before that time. Most of our observations come from the question, “Do you approve or disapprove of the way George W. Bush is handling the situation in Iraq?” The question was asked with little variation and great regularity by six different organizations. The data set for perceptions of war success comprises 92 observations for 84 weeks. All but ten observations stem from slight variations to the question, “How would you say things are going for the U.S. in Iraq?”

In addition, we collected data on overall evaluations of Bush’s performance as president. Overall job approval is measured using the standard question: “Do you approve or disapprove of the job George W. Bush is doing as president?” This question was asked 323 times by seven survey organizations.

Method

Seemingly innocuous variations in question wording can significantly influence survey responses, and the same question administered by different survey organizations can

---

9 CNN/TIME asked a variant: “Do you think President Bush is doing a good job or a poor job handling the situation in Iraq?” Newsweek originally asked, “Do you approve or disapprove of the way Bush is handling policies to deal with the threat posed by Iraq and its leader Saddam Hussein?”


11 The relative thinness is due partly to a gap in the dataset: in a five-week period following January 11 2004 we have only one observation. A September 19-21, 2003 Gallup poll (N = 1,003) indicates that the overwhelming majority of respondents viewed the war as ongoing even after President Bush had declared an end to major combat operations, with 89% of respondents answering “no” to the question, “Based on what you have heard or read about the events in Iraq over the past few weeks, do you think that for all intents and purposes, the war in Iraq is over, or not?”

12 The most distinctive is a question that CBS asked seven times: “From what you have seen or heard, is the United States in control of events taking place in Iraq, or are the events in Iraq out of U.S. control?”

yield different marginals due to variation in sampling and survey procedures. The methodological task is to disentangle genuine temporal changes from the effects of question wording and random noise. This task is complicated by the fact that the various items are not measured at the same points in time, necessitating a statistical model rather than simple standardizing as a way to separate these sources of variation (see also Stimson 1991).

Let $Y_{jt}$ denote the observed percentage supporting the war (or approving of the president’s handling of it, and so forth) on question $j$ at time $t$. Our objective is to estimate the changing level of support $\theta(t)$ that presumably underlies these observations. Each observation provides a noisy measurement of the public’s true aggregate position at a specific time. The following model describes the observations:

$$Y_{jt} = \alpha_j + \beta_j \theta_t + \epsilon_{jt} \quad (1)$$

In equation 1, $\alpha_j$ represents the “item bias,” a parameter that corrects for the extent to which a question understates or overstates war support in comparison to an anchor item. For example, the Fox News question (“Do you support or oppose the United States having taken military action to disarm Iraq and remove Iraqi President Saddam Hussein?”) on average generated an estimated 11-percentage point higher level of support for the war than did the Gallup question from Table 1, which mentions neither disarmament nor Saddam Hussein.

---

14 The data are centered around zero (based on the overall mean). A focus on support implies the assumption that trends in opposition largely mirror approval, which is standard in the literature. To check the validity of this assumption, we examined whether there were noticeable trends in the “don’t know” categories, which there were not.
\( \beta_j \) is a “weight” parameter that adjusts for the extent to which a question taps our main concept of interest. It is equivalent to a factor loading. Theoretically, we could estimate this parameter for each item. In practice, however, this would be difficult given the limited amount of data for each individual item. We therefore assume that this parameter is a constant except for items for which we have substantive grounds to suspect otherwise. As mentioned before, the dataset for war support includes questions that specifically mention ground troops and casualties. These questions may tap into slightly different aspects of support for the war than do questions that make no explicit mention of either, which we account for by estimating separate weight parameters.

The latent level of support follows a random walk process: \( \theta_t \sim N(\theta_{t-1}, \tau^2) \), with a diffuse prior on \( \theta_t \). This imposes dynamic structure on the estimation. Our specific assumption about persistence is motivated by arguments in the literature that in finite samples, series such as public mood or presidential approval can be expected to mimic the behavior of a unit root process (e.g., Durr 1993, DeBoef and Granato 1997). The random portion of the measurement error is described by \( \varepsilon_{jt} \sim N(0, \sigma^2) \). The models are estimated using MCMC and implemented in WinBugs.

The models fit the raw data very well.\(^{15}\) Moreover, the estimates correlate well with estimates from an alternative procedure: WCALC (Stimson 1991, 1999).\(^{16}\) The main advantage of our approach is that we obtain a series that is simultaneously smooth and allows for sudden shocks, whereas WCALC obscures sudden breaks in the data (and thus

\(^{15}\) In an ANOVA analysis on the raw data controlling for fixed effects of the items, the adjusted \( R^2 \) is over .90 for all series and around .95 for most series.

\(^{16}\) Bivariate correlations are .97 for war support, .96 for Bush Iraq approval, .94 for perceptions of success, .98 for Bush economic approval and .99 for overall job approval.
the short-term impact of events) or leads to seemingly random zigzagging (when the smoothing option is turned off). A web-appendix provides more detail on the method and comparisons to WCALC.

Events, Casualties, and Beliefs about the War
Figure 1 compares war support with Bush job approval on Iraq and perceptions of war success. All three series are scaled to represent the marginals from the main Gallup series. The percentage of Americans who judged the war to be worthwhile hovered at around 55 until late January of 2003. Thus, it appears that the initial efforts by the administration to persuade the public of the war’s necessity had little effect. Support rose sharply starting around the time that Colin Powell gave the speech in the United Nation Security Council that signaled the beginning of the diplomatic endgame. The speech itself did not lead to a one-time shift in support for the war. Instead, such support grew steadily as it became increasingly clear that the U.S. was relatively isolated in its insistence on the use of force. This runs counter to Kull and Destler’s (1999) argument that American public support for military intervention is a function of international public support, as measured by whether a use of force is supported multilaterally.

17 For perceptions of war success the Gallup question is, “In general, how would you say things are going for the U.S. in Iraq?” The marginal represents the percentage of respondents who answered “very well” or “moderately well.”
18 Much of the initial support for an invasion was apparently in place in the immediate aftermath of September 11, 2001. For example, on October 4, 2001, 73% of respondents answered favorably to the question, “Do you favor or oppose the United States taking direct military action in Iraq?” (N = 1005, Gallup and the Los Angeles Times).
20 At the same time, surveys found that support for military action tended to be greater when questions posited international support.
The initial apparent success of the war effort drove support even further upward, to a maximum of around 70%. After that, support dropped to a minimum of 46% in October 2004. The gradual decline was interrupted by what appeared to be pieces of good and bad news, in particular the capture of Saddam Hussein and the revelations of the prison scandal in Abu Ghraib. In addition, the announcement by President Bush in September 2003 that an additional 87 billion dollars was needed to cover military and reconstruction costs was followed by an apparent drop in war support.

Perceptions of war success and Bush job approval on Iraq followed roughly the same pattern, although there were some noteworthy differences.21 By simply looking at the figures, we can see that the relative levels of instability in the series followed our expectations. Though the means of the series were roughly equivalent,22 their variances differed considerably and as expected.23 The estimated percentage of respondents believing that things are going very or moderately well in Iraq varied from a high of 78% in the first weeks after the start of the war to a low of 43% when the revelations about Abu Ghraib were prominently in the news and Nicholas Berg was beheaded (May 8, 2004). Similarly, evaluations of Bush job approval on Iraq varied from 74% to 39%. Clearly, evaluations of war success and Bush’s management capabilities were particularly sensitive to short-term considerations.

21 Bivariate correlations with war support were .85 (war success) and .95 (Bush Iraq approval). The bivariate correlation between the latter two series was .93.
22 After the start of the war they were 55.2 for war approval, 52.6 for war success, and 51.8 for Bush Iraq approval.
23 Standard deviations were 8.9 (perceptions of success), 9.5 (Bush approval on Iraq) and 5.7 (war support).
Casualties and Events

A prediction derived from Mueller (1973) and pursued by others is that support for a war decreases with cumulative American casualties (deaths). There is, however, a problem in evaluating the impact of cumulative casualties on aggregate evaluations of the war in that they are non-stationary. Hence, it is not advisable to regress the levels of these series on each other, as spurious results are likely to occur. Instead, we may analyze the differenced series. This, however, ignores the possible long-run relationships between variables. Error-correction models provide a possible solution to both issues and have become the model of choice in analyses of presidential approval and similar time series (e.g., Beck 1993, Clarke et al. 1998, DeBoef and Kellstedt 2004).

Table 2 reports the results of a single-equation error-correction model. The lagged level of the dependent variable is included because high levels of support in the previous period may be correlated with a negative change in the following period due to equilibrium correction. The capture of Hussein and the Abu Ghraib scandal are included as the most important events that should have positive and negative effects, respectively, on assessments of the war. The immediate effect of casualties is measured by the marginal change in the number of cumulative casualties from the preceding week (see Gartner and Segura 1998). If casualties have a long-term equilibrium correcting effect,

\[ \Delta Y_t = \beta_0 - \beta_1 (Y_{t-1} - \beta_2 X_{t-1}) + \beta_3 \Delta X_t + \epsilon_t \]

The single-equation model is often theoretically preferable (Beck 1993) and more efficient (DeBoef 2001) than the two-step Engle-Granger method (Engle and Granger 1987). Moreover, it is less directly linked to the analysis of cointegrated series as it can be derived from the autoregressive distributed lag model. Results from analyses on the differenced series and Prais-Winsten regressions are available from the authors. The main results are robust across these different specifications.

Models that included the other events noted in the discussion did not produce different results. The casualty data come from http://icasualties.org, which tallies CENTCOM and Department of Defense releases of combat and non-combat deaths. Mueller (1973) advocated taking the
the lag of the cumulative casualties should also have an impact. We limit the analysis to the period after the first three weeks of the war. Expected casualties are higher during the initial fighting than after major combat is over. One might control for this shift in expectations by including a dummy variable for major combat and an interaction of this variable with casualties, but this is not feasible here given the short duration of combat. The results are, however, robust to slight alterations in the starting point of the analysis.

TABLE 2 ABOUT HERE

As expected, casualties had a significant impact on perceptions of war success but not on beliefs that the war was worth it. Even the effect on perceptions of war success is small: 33 casualties led to an estimated one-percentage point drop in aggregate perceptions that the war is going well, not a very large impact given that the mean number of weekly casualties was 14, with a maximum of 52. Lagged cumulative casualties are not significant at the 5% level. The capture of Hussein led to an estimated 7.0 percentage point increase in beliefs that the “war is going well,” but only a 2.6 percentage point increase beliefs that the “war was worth it.”27 By contrast, Abu Ghraib had the strongest effect on perceptions of Bush’s job performance on Iraq, which was reduced by an estimated 6.4 percentage points as a consequence of the scandal.28 This may be an instance where an event revealed something about competence without having a discernable impact on the perceived prospects for winning the war.

Consistent with expectations, casualties and events explained perceptions of the war’s success much better than they explained war support. Casualties and shocks due to

\[ \text{natural log because early deaths supposedly have the greatest marginal impact on opinion. We found no such effect in our data.} \]

27 The capture of Saddam was coded as a two-week effect (week 48-49 in 2003).

28 Abu Ghraib was coded as an event that lasted from week 17-21 in 2004.
high profile events accounted for little of the temporal variation in support for the war. Moreover, as the last column of Table 2 shows, perceptions that the “war is going well” explain only about one fifth of the variation in beliefs that the “war was worth it.” This suggests that swings in support for the war may have been determined in part by shifts in elite discourse that were relatively unconstrained by the happenings of the war.

**Beliefs about the War and Support for the President**

Did changes in beliefs about the war move overall presidential approval? If so, what sort of beliefs mattered most? To address these questions, we need to consider alternative accounts of presidential popularity. The most common explanation has been that evaluations of the president’s job performance are largely driven by subjective assessments of the economy, with objective economic conditions mattering only to the extent that they affect such assessments (e.g., MacKuen et al. 1992). Figure 2 graphs overall presidential job approval, war support, and a Consumer Comfort Index (CCI) constructed by ABC/Money magazine. Bush’s overall job approval jumped by about 7.7 percentage-points at the start of the war. This “rally around the flag” effect was not exceptional compared to the 9-percentage-point average jump recorded for major

---

29 This is a rolling weekly average based on telephone interviews with about 1,000 adults nationwide each month. Field work is by ICR Survey Research of Media, Pa. The published index ranged from +100 (everyone positive on all three measures) to -100 (all negative on all three measures) but was rescaled to the 0-100 interval for comparability. The three questions used to calculate the index are: “Would you describe the state of the nation’s economy these days as excellent, good, not so good, or poor?” “Would you describe the state of your own personal finances these days as excellent, good, not so good, or poor?” “Considering the cost of things today and your own personal finances, would you say now is an excellent time, a good time, a not so good time, or a poor time to buy the things you want and need?” We also estimated a model that substituted Bush economic job approval as an independent variable. This generated similar, but weaker, results for the economic variable.
responses to international crises (Oneal and Bryan 1995). The visual evidence suggests that presidential job approval and war support tracked each other closely. Consumer comfort, on the other hand, followed a different path. In fact, the correlations between consumer comfort and both overall presidential approval and war support were negative and significant over the full two-year period, though not after the war started. It is plausible that the anticipation of the war depressed consumer confidence.

FIGURE 2 ABOUT HERE

Table 3 presents the results of a regression analysis using error-correction models similar to the ones from Table 2. The dependent variable is the weekly change in presidential approval ratings. The lagged level of presidential approval is included to capture equilibrium corrections. The other independent variables are weekly changes in various beliefs about the war and consumer confidence. To ensure comparability the analysis is limited to the period after the war started.

There are several interesting findings. First, weekly shifts in war support had a strong immediate effect on changes in presidential approval. A one-percentage point increase in support for the war led to a .78 percentage point immediate increase in overall presidential approval. By contrast, the corresponding effects for perceptions of war success and Bush approval on Iraq were .34 and .45, respectively. The latter two variables had significant long-run relationships as well. Overall, however, the model

---

30 Some caution is warranted here in that we measured the change from the week before to the week of the actual intervention, whereas previous studies have used a much cruder time frame.
31 The Pearson Rs were -.50 and -.51, respectively over the entire period but were indistinguishable from zero after the start of the war.
33 This is especially true for Bush Iraq approval. The two-step Engle-Granger test shows strong evidence that Bush overall approval and Bush Iraq approval are cointegrated series.
that included changes in overall war support explained the data better than did the other
two models, especially the model that included changes in perceptions of war success.
When this latter variable was included in a model with war support (model 4), we found
no independent effect for perceptions of war success. Together, these findings confirm
our expectation that shifts in aggregate beliefs that the war was worth it were more
consequential politically than shifts in beliefs that it was going well.

The finding that shifts in war support relate more strongly and directly to
presidential approval than even shifts in approval for Bush’s job performance on Iraq is
surprising given that the wording of the questions that assessed overall job approval and
job approval on Iraq differed only slightly. We take this as strong evidence for the
consequences of shifts in aggregate beliefs about the wisdom of initiating the war.

A model that included only the events and casualty variables (Model 5) explained
very little variation in weekly approval ratings. Moreover, when events and casualties
were included in the same model as aggregate beliefs that the “war is worth it,” the event
variables were jointly insignificant (Model 6). These findings confirm our thesis that if
we want to capture the political impact of war, we should look beyond the standard
approach of linking events and casualty figures directly to approval ratings.

Controlling for beliefs about the war, changes in consumer confidence had the
expected positive impact on presidential approval ratings. A one-point increase in the
consumer confidence index led to about a .17 percentage point increase in presidential
approval. This effect was modest given that the range of the index was between 38 and

34 Note, however, that a proper test of the full theoretical model implies a structural equation
approach. We limit the analysis here to a test of several implications of the model.
35 This finding also held for the entire period over which the two series can be compared (94
weeks) with the analyses producing virtually identical coefficients.
48, with a maximum weekly shift of 3.5 points. By comparison, war support had a larger marginal impact on presidential approval and varied over a wider range (46-70). The marginal effect of changes in consumer confidence did not increase as the election approached. Hence, it appears that shifting perceptions about whether the war was worth it were more important in explaining shifts in presidential approval ratings than were shifts in perceptions about the economy.

Conclusions
Our two main findings are (1) that events and casualties accounted for considerable variation in perceptions of the war’s success but did not explain as much variation in support for the war and the president, and (2) that shifts in war support accounted for shifts in presidential approval ratings better than did shifts in perceived war success or even shifts in Bush’s job approval on Iraq. These findings suggest that, the public changed its collective mind on whether the war was going well in response to events and casualties, but also that opinion on this topic was of relatively little consequence for presidential support. On the other hand, when the public changed its collective mind on whether the war was worth it, this had immediate and considerable consequences for the president. Yet here we found weaker relationships between war support, on the one hand, and key events and casualty reports, on the other.

How, then, does the public evaluate and reevaluate presidential decisions to go to war? Like other observers, we suspect that support for war is shaped in part by the signals, or cues, that elites send and that these signals are only partly constrained by

36 For instance, a regression on half the period (the last 41 weeks before the election) yielded regression coefficients of .71 for war support and .18 for consumer confidence.
reality. Research on previous wars—including World War II, the Korean War, the Vietnam War, and the 1991 Gulf War—indicates that the nature of elite discourse about a war is a key determinant of changing patterns in both war support itself (Zaller 1992; Berinsky 2004) and presidential approval during wartime (Brody 1991, 1994). Of particular relevance to our findings is the notion that the extent of elite consensus shapes public support for a war. A plausible interpretation of the trend in support for the war in Iraq is that increasing elite polarization regarding the war produced not only polarization among the public (with Democratic and Republican partisans diverging on its wisdom; see Berinsky 2004) but also declining support for the war in the aggregate.

Another question is whether our conclusions can be extended to other wars. In the case of the 1991 Gulf War, the short (six-week) span of the war and the limited variance in public opinion about the war and the president during this span (see Mueller 1994) render our aggregate time-series approach impractical. For earlier wars of longer durations, the relative paucity of data may do the same. Still, our analyses of the war in Iraq suggest, at the very least, additional caution in judging claims regarding the unmediated effects of casualties and events on support for some of these previous wars (for a complementary argument based on individual-level analysis, see Berinsky 2004).

Even taking into account its limitations, our study has advanced the literature on the relationship between accountability and war considerably. First, we have presented direct evidence regarding the role of aggregate beliefs about a war in shaping aggregate support for a commander in chief. The results for the Iraq war suggest that previous analyses may have underestimated this impact to a considerable degree given that they tended to rely on casualties and events as proxies for aggregate beliefs about the war.
Second, we have distinguished between two models for understanding presidential accountability during wartime and tested these models. In doing so, we have produced evidence that the “decision-maker” model of accountability works better than the “managerial” model. These insights should lead to further theoretical and empirical inquiry into what kinds of subjective beliefs about a war are politically consequential and what shapes these beliefs. Moreover, our competing accountability models have implications for the burgeoning literature on democracies and war: though a wide range of results suggest that democracies behave differently in choosing and fighting wars, there is still a wide gap in understanding the accountability mechanisms that generate these results (see Reiter and Stam 2002). Third, we have developed a new measurement model for estimating time series from survey marginals. This model may prove useful for other applications, especially when data are sparse, given that it allows one to combine data from various survey instruments. Finally, the analyses provide insights regarding the dynamics of public opinion about a war that is interesting in and of itself—one that is ongoing and for which questions regarding presidential accountability are likely to be of enduring interest in the scholarly community and beyond.

References


<table>
<thead>
<tr>
<th>Representative Question</th>
<th>Observations</th>
<th>Items</th>
<th>Weeks</th>
<th>Start Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Do you think the situation in Iraq was worth going to war over, or not?”</td>
<td>323</td>
<td>25</td>
<td>119</td>
<td>8/6/2002</td>
</tr>
<tr>
<td>“In general, how would you say things are going for the U.S. in Iraq: very well, moderately well, moderately badly, or very badly?”</td>
<td>92</td>
<td>8</td>
<td>84</td>
<td>1/16/2003</td>
</tr>
<tr>
<td>&quot;Do you approve or disapprove of the way George W. Bush is handling the situation with Iraq?”</td>
<td>144</td>
<td>8</td>
<td>94</td>
<td>3/25/2003</td>
</tr>
<tr>
<td>&quot;Do you approve or disapprove of the way George W. Bush is handling his job as president?”</td>
<td>339</td>
<td>7</td>
<td>119</td>
<td>1/16/2003</td>
</tr>
</tbody>
</table>
Table 2: Casualties, Events, and Changes in Beliefs about the War (after first 3 weeks of war, N=80, standard error in parentheses).

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(\Delta) War Going Well</th>
<th>(\Delta) Bush Approval Iraq</th>
<th>(\Delta) War Worth It</th>
<th>(\Delta) War Worth It</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-.382</td>
<td>-.387</td>
<td>-.275</td>
<td>-.225</td>
</tr>
<tr>
<td>Lagged Level</td>
<td>(.311)</td>
<td>(.435)</td>
<td>(.164)</td>
<td>(.101)</td>
</tr>
<tr>
<td>Dependent Variable</td>
<td>-.054**</td>
<td>-.055*</td>
<td>-.091*</td>
<td>-.024</td>
</tr>
<tr>
<td>(\Delta) Casualties</td>
<td>-.033*</td>
<td>-.008</td>
<td>-.007</td>
<td>-</td>
</tr>
<tr>
<td>Lagged CumulativeCasualties</td>
<td>.001</td>
<td>.000</td>
<td>-.001</td>
<td>-</td>
</tr>
<tr>
<td>Saddam Capture</td>
<td>3.499**</td>
<td>2.202*</td>
<td>1.276*</td>
<td>-</td>
</tr>
<tr>
<td>Abu Ghraib</td>
<td>-.130</td>
<td>-1.595*</td>
<td>-.762*</td>
<td>-</td>
</tr>
<tr>
<td>(\Delta) War Going Well</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.294** (.060)</td>
</tr>
<tr>
<td>Lagged War Going Well</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.018 (.018)</td>
</tr>
<tr>
<td>R(^2) adj</td>
<td>.395</td>
<td>.163</td>
<td>.162</td>
<td>.222</td>
</tr>
<tr>
<td>S.E. Estimate</td>
<td>1.012</td>
<td>1.354</td>
<td>.754</td>
<td>.774</td>
</tr>
</tbody>
</table>

**p<.01 *p<.05 , all tests are two-tailed
Table 3: Weekly Change in Presidential Approval by Beliefs about the War and the Economy (since start of war N=82, for models with casualties N=80, standard error in parentheses).

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-.475</td>
<td>-.350</td>
<td>-.742</td>
<td>-.883*</td>
<td>-.929</td>
<td>-1.717*</td>
</tr>
<tr>
<td></td>
<td>(.307)</td>
<td>(.839)</td>
<td>(.979)</td>
<td>(.372)</td>
<td>(.599)</td>
<td>(.599)</td>
</tr>
<tr>
<td>Bush Approval t-1</td>
<td>-.103*</td>
<td>-.311**</td>
<td>-.159**</td>
<td>-.158**</td>
<td>-.070*</td>
<td>-.179**</td>
</tr>
<tr>
<td></td>
<td>(.044)</td>
<td>(.079)</td>
<td>(.057)</td>
<td>(.055)</td>
<td>(.028)</td>
<td>(.054)</td>
</tr>
<tr>
<td>∆ War Worth It</td>
<td>.784**</td>
<td>-</td>
<td>-</td>
<td>.672**</td>
<td>-</td>
<td>.750*</td>
</tr>
<tr>
<td></td>
<td>(.095)</td>
<td>-</td>
<td>-</td>
<td>(.108)</td>
<td>-</td>
<td>(.111)</td>
</tr>
<tr>
<td>Lag War Worth It</td>
<td>.072</td>
<td>-</td>
<td>-</td>
<td>.083</td>
<td>-</td>
<td>.222**</td>
</tr>
<tr>
<td></td>
<td>(.047)</td>
<td>-</td>
<td>-</td>
<td>(.047)</td>
<td>-</td>
<td>(.081)</td>
</tr>
<tr>
<td>∆ Bush Approval Iraq</td>
<td>-</td>
<td>453**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(.060)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bush Approval Iraq t-1</td>
<td>-</td>
<td>.179**</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>(.050)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>∆ War Going Well</td>
<td>-</td>
<td>-</td>
<td>.336**</td>
<td>.129</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>(.072)</td>
<td>(.068)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>War Going Well t-1</td>
<td>-</td>
<td>-</td>
<td>.069**</td>
<td>.036</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>(.029)</td>
<td>(.023)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>∆ CCI</td>
<td>.086*</td>
<td>.082*</td>
<td>.092*</td>
<td>.087*</td>
<td>.034</td>
<td>.063*</td>
</tr>
<tr>
<td></td>
<td>(.036)</td>
<td>(.037)</td>
<td>(.044)</td>
<td>(.035)</td>
<td>(.046)</td>
<td>(.036)</td>
</tr>
<tr>
<td>CCI t-1</td>
<td>-.010</td>
<td>-.013</td>
<td>-.044</td>
<td>-.024</td>
<td>-.012</td>
<td>-.020</td>
</tr>
<tr>
<td></td>
<td>(.020)</td>
<td>(.023)</td>
<td>(.027)</td>
<td>(.021)</td>
<td>(.028)</td>
<td>(.022)</td>
</tr>
<tr>
<td>∆ Casualties</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.021</td>
<td>.015</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(.012)</td>
</tr>
<tr>
<td>Cumulative Casualties t-1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.000</td>
<td>.002*</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(.001)</td>
</tr>
<tr>
<td>Saddam Capture</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>.861</td>
<td>-.076</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(.689)</td>
<td>(.551)</td>
<td>-</td>
</tr>
<tr>
<td>Abu Ghraib</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-1.160*</td>
<td>-.226</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(.444)</td>
<td>(.380)</td>
<td>-</td>
</tr>
<tr>
<td>R² adj</td>
<td>.544</td>
<td>.481</td>
<td>.290</td>
<td>.533</td>
<td>.219</td>
<td>.507</td>
</tr>
<tr>
<td>S.E. Estimate</td>
<td>.721</td>
<td>.769</td>
<td>.900</td>
<td>.730</td>
<td>.935</td>
<td>.722</td>
</tr>
</tbody>
</table>

**p<.01 *p<.05 , all tests are two-tailed**
Figure 1: Beliefs about the War

- Is/Was the War Worth It?
- Is Bush Doing Good Job in Iraq?
- Are Things Going Well in Iraq?

Start of War
Powell UN Speech
$87 Billion
Capture Saddam
Abu Ghraib
Figure 2: Beliefs about the War, Consumer Comfort, and Presidential Approval

- Is/Was the War Worth It?
- Consumer Comfort (ABC/Money, 0-100)
- Presidential Job Approval
- Is War Going Well?