

THE POLLS—REVIEW

PUBLIC OPINION RESEARCH AND SUPPORT FOR THE IRAQ WAR

ADAM J. BERINSKY
JAMES N. DRUCKMAN

Abstract Professors Peter Feaver, Christopher Gelpi, and Jason Reifler’s theory of the determinants of public support for war has received a great deal of attention among academics, journalists, and policymakers. They argue that support for war hinges on initial support for military action and the belief in the success of the war. In this review, we take a critical and constructive view of their work, focusing on methodological concerns. We discuss the dependent variable used by the authors—individual casualty tolerance—and argue that it is an insufficient measure of war support. We also make the case that their independent variables of interest—initial support for war and evaluation of war success—may, in fact, be best understood as indicators of latent support for the war more generally. Finally, we discuss the need for more research into the determinants of support for war, focusing on core values and elite rhetoric as potential variables for continued and future study.

The question of whether politicians, particularly the president, attend to public opinion when crafting foreign policy has long intrigued social scientists. For many years, conventional wisdom suggested that the public lacked the sophistication needed to form coherent foreign policy attitudes. From this point of view politicians should—and do—ignore public opinion. More recently, however, scholars have countered this perspective, producing considerable evidence that the public holds consistent opinions on foreign

ADAM J. BERINSKY is an Associate Professor in the Department of Political Science, Massachusetts Institute of Technology, Cambridge, MA 02139. JAMES N. DRUCKMAN is an Associate Professor in the Department of Political Science, Northwestern University, Evanston, IL 60208, USA. We thank Jason Reifler and Christopher Gelpi for making their questionnaires available to us and Chris Federico for providing us with the data from his co-authored article. We thank the anonymous reviewers, Sean Aday, Matthew Baum, Paul Brewer, Sunshine Hillygus, Ron Krebs, Rose McDermott, Peter Miller, Jason Reifler, Mark Schulman, and John Sides for helpful advice, and Toby Bolsen and Matt Gusella for research assistance. Address correspondence to James N. Druckman, e-mail: druckman@northwestern.edu

doi:10.1093/poq/nfl049

© The Author 2007. Published by Oxford University Press on behalf of the American Association for Public Opinion Research. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org

policy, and that politicians respond to these opinions. As Aldrich et al. (2006, p. 496), succinctly summarize, “A mounting body of evidence suggests that the foreign policies of American presidents—and democratic leaders more generally—have been influenced by their understanding of the public’s foreign policy views.”

Understanding the determinants and effects of public opinion on foreign policy becomes increasingly relevant during times of war, such as the current conflict in Iraq. While Presidents, at least since Nixon, have relied on their private public opinion polls in crafting their rhetorical strategies (Jacobs and Shapiro 1995), it is rare that we can actually trace the influence of opinion polls on presidential rhetoric and policy. President Bush’s November 30, 2005, address where he outlined his future strategy for the War in Iraq may have been such a case.

During the speech, Bush heavily emphasized the concept of “victory,” by using the word “victory” 15 times, posting “Plan for Victory” signs on the podium, and entitling an accompanying National Security Council report “National Strategy for Victory in Iraq.” Bush presumably sought to persuade citizens to expect success, thereby increasing support for the war, his foreign policy, and his administration (Shane 2005). A number of news outlets traced the origins of Bush’s victory theme to public opinion survey results. The *New York Times* pointed to the research of National Security Council (NSC) staffer Dr Peter Feaver (who also is a political science Professor at Duke University), along with his colleagues, Professors Christopher Gelpi (of Duke) and Jason Reifler (of Loyola, Chicago) who have argued that the public supports military action when they believe that a war will succeed. This attribution sparked considerable debate in the mainstream media, academic journals, and websites about the origins and nature of public opinion about war, with particular attention to Feaver, Gelpi, and Reifler’s approach.

The purported reliance by Bush on the work of Feaver and his colleagues is beneficial for public opinion researchers. For one, it demonstrates an application of cutting-edge scholarship to ongoing political events. But casting a spotlight on this research also generates debate about Feaver, Gelpi, and Reifler’s particular findings, and more generally about what we, as a research community, know about attitudes toward war.

Feaver, Gelpi, and Reifler have developed an intriguing theory of the determinants of public support for war and have completed a unique data collection effort. In what follows, we provide a critical but hopefully constructive review of Feaver, Gelpi, and Reifler’s work. Our intent is neither to present an alternative theory (we present little new analyses) nor to disparage the authors’ work which we consider rigorous and thought-provoking. Rather, by focusing on Feaver, Gelpi, and Reifler’s influential work, we attempt to determine whether the authors’ test of their theory allows them to adjudicate between competing claims. We also hope to clarify those

issues that need to be further explored to understand citizens' preferences regarding military conflict. Specifically, in this poll review, we raise some methodological concerns from a survey research perspective that give some hesitation to the conclusions of Feaver, Gelpi, and Reifler. For a number of reasons, which we detail subsequently, the analysis that Feaver, Gelpi, and Reifler present makes it difficult to distinguish between their hypothesized causal hypothesis and that of other leading theories of the determinants of war support, such as the casualties hypothesis advanced by Mueller (1973).

Contending Views on War Attitudes

While much ink has been spilled on the study of public opinion concerning war, here we focus on work most pertinent to Feaver, Gelpi, and Reifler. A common starting point is Mueller's (1973) study of public opinion concerning the Korean and Vietnam wars. Though Mueller's book is a comprehensive treatment of several correlates of support for war, it is best known for presenting his argument on the effects of casualties. Mueller argues that support for military action declined as a function of American casualties.¹ This observation led to a conclusion that holds weight with both policymakers and academics: the American public is casualty phobic.

In a series of works (e.g., Feaver and Gelpi 2004, Gelpi, Feaver, and Reifler 2005–06 and Gelpi, Reifler, Feaver n.d.), Feaver, Gelpi, and Reifler argue that casualties are *not* the key element in explaining war support. (We henceforth refer to the latter two articles collectively as FGR).² Instead, FGR make the case that the “public’s tolerance for the human costs of war is primarily shaped by the intersection of two crucial attitudes: beliefs about the rightness or wrongness of the war and beliefs about a war’s likely success . . . [with the] likelihood of success matter[ing] most” (Gelpi et al. 2005–06, p. 8). Perhaps their most relevant (publicly available) evidence comes from surveys on the Iraq war conducted in October 2003 and October 2004, the results of which show that respondents’ tolerance for casualties depends on these two critical attitudes, and an interaction between the two measures.

1. Specifically, Mueller argued that support for war was inversely related to the logarithm of the total number of American casualties. This formulation assumes that sensitivity to casualties declines as the number of war deaths increases.

2. FGR are certainly not the first authors to question Mueller's contention. There is a large and diverse literature on the determinants of support for war—much of it critical of Mueller's original formulation of the casualties hypothesis (Burk 1999; Jentleson 1992; Larson 1996; Gartner, Segura, and Wilkening 1997; Kull and Destler 1999; Gartner and Segura 2000; Klarevas 2002; Holsti 2004; Berinsky 2006). However, because the focus of this review is the FGR work, we set aside discussion of contending theories. We agree with FGR's assertion that support for war depends on various factors in addition to casualties.

In the debate following the aforementioned publicity linking FGR's research and Bush's speech, the distinction between FGR and Mueller rose to the forefront.³ For example, the January/February 2006 issue of *Foreign Affairs* contained a heated exchange between Gelpi and Mueller (also see Mueller 2005, Klarevas, Gelpi, and Reifler 2006), and debates appeared on a number of websites (see, e.g., http://www.mysterypollster.com/main/2005/12/polling_the_str.html from December 5, 2005; http://blogs.washingtonpost.com/earlywarning/2006/02/overstating_the_impact_of_iraq.html from February 23, 2006). While fruitful in many ways, these exchanges also left important issues about FGR's analysis open to question. In the remainder of this article, we largely set aside questions regarding the veracity of Mueller's argument to carefully consider three integrally related dimensions of the FGR analysis: the dependent variable, the independent variables, and, ultimately, their causal claims.

Support for War

The core construct in any study of public opinion and war must be support for that war. This is the fundamental attitude to which policy-makers presumably turn and is the focus of FGR's research. FGR (2005–06, p.8) state, "Our findings imply that the American public makes reasoned and reasonable judgments about an issue as emotionally charged and politically polarizing as fighting a war. Indeed, the public forms its *attitudes regarding support for the war in Iraq* in exactly the way we should hope they would: weighing the costs and benefits" (also see Gelpi et al. 2005; emphasis added).

But, how exactly should support for war be measured? This is a question that Mueller (1973) identified as a perplexing one, due in large part to the well-known existence of question wording effects: seemingly subtle differences in the way survey questions are phrased can lead to large differences in the responses generated and, ultimately, in the shape of aggregate public opinion.

As a solution to this problem, Mueller (1973, p. 43) utilized various similarly worded questions in his study of Korea and Vietnam, with the core construct tapping—as he put it—"a sort of generalized support for the war" through the use of retrospective evaluation "mistake" questions. An example of his Korea item asked, "Do you think the United States made a mistake in going into the war in Korea, or not?" The main Vietnam version queried: "In view of the developments since we entered the fighting in Vietnam,

3. Feaver joined the NSC staff as a special advisor in June, 2005, and the "National Strategy for Victory in Iraq" report posted on the White House Website in November, 2005, showed that the document's author was "feaver-p." White House officials confirmed that Feaver played a significant role in drafting the plan (Shane 2005). While Gelpi and Reifler did not play a role in drafting the document, the document is clearly based on their research.

do you think the United States made a mistake sending troops to fight in Vietnam?” However, in addition to these general questions, Mueller asked several related questions about withdrawal, escalation, and isolation. In each case, these questions led Mueller to similar conclusions regarding the structure of support for war.

Other scholars have recognized the importance of using multiple items to capture the latent construct of war support. For example, Burke (1999) examines support for missions in Lebanon using some questions that measure approval for the decision to send troops, and others that ask if it was a mistake to send troops. Media organizations similarly tend to ask survey questions that tap both global evaluations of support and attitudes on specific policies—for instance whether troop levels should be increased or decreased (see Everts and Isernia 2005 for a comprehensive overview of public support for the Iraq War). Eichenberg (2005, p. 153) summarizes the conventional wisdom well when he writes that “a single question on any issue will be a misleading gauge of the public mood because an infinite variety of question wordings on any issue is conceivable, and each is likely to yield a different set of responses. . . a reliable analysis requires the study of many survey questions that employ a variety of wordings.” Thus, when examining generalized levels of support for war—a concept measured only imperfectly by any single item—it is best to look at multiple indicators of such support.

Such an approach differs from FGR’s strategy of measuring war support with one dependent variable: *individual casualty tolerance*. FGR ask:

regardless of whether you think the President made the right decision in attacking Iraq, as you know the United States is engaged in an ongoing military operation there and has suffered about [INSERT current number] military deaths in combat. Would you support continued US military action in Iraq until a new Iraqi government can take over if it results in no additional US Military deaths?

If the respondent answered “yes,” FGR continued asking the question with increases in the number of resulting deaths (e.g., up to 500 deaths, up to 5,000 deaths, etc.).⁴ An individual’s response then equals the highest number of deaths the individual would tolerate while still supporting the military action.

While FGR’s measure undoubtedly correlates with general support for war, we have some concerns about the reliability and validity of casualty tolerance. One worry is that FGR’s dependent variable inherently assumes

4. FGR also study aggregate presidential approval, aggregate presidential handling of the Iraq war, and an aggregate measure of whether the war was worth it (however they do not present formal results for the latter two variables). When FGR (2005–06, p. 28) turn to the study of vote choice, they argue that their “measure of casualty tolerance does a better job of gauging the policy relevant issue of continued support of an ongoing military operation than do other more commonly used measures of casualty sensitivity.” It is unclear why they do not directly measure support of the ongoing military operation.

that support for war should be measured by the number of American war deaths a respondent is willing to bear. This approach is ironic since FGR in fact argue that casualties are not the primary determinant of support for war. In addition, a burgeoning literature on contingent-valuation suggests that people are not adept at reliably estimating their tolerance for paying costs, such as war deaths (see, e.g., Bartels 2003, p. 52). One particularly pervasive problem is that people base estimates on the prototypical incidence (e.g., the prototypical death) that comes to mind rather than on a consideration of the actual frequencies (e.g., number of deaths) (Lichtenstein, Slovic, Layman, and Combs 1978, Hertwig, Pachur, and Kurzenhäuser 2005). As a result, responses might be quite sensitive to such things as exposure to vivid media images of military deaths (see, e.g., Kahneman, Ritov, and Schkade 1999 for a related example).⁵ In addition, hypothetical valuation assessments, such as FGR's, often significantly differ from individuals' estimates during analogous real-world, ongoing events (Schlöpfer and Hanley 2006).⁶

Our point here is not that FGR's measure is necessarily more problematic than other individual measures. Indeed, the aforementioned "mistake" question which researchers often use, due undoubtedly to its widespread availability on media polls, is not ideal either. For example, a respondent may believe going to war was a mistake but nonetheless he or she might currently support the war due to subsequent developments.⁷ Rather, we subscribe to the aforementioned wisdom regarding the importance of using multiple questions to measure general support for war.⁸ This is a particularly acute concern in the case of FGR because, as we will discuss below, we worry that FGR do in fact measure additional items that capture general war support, but treat these additional items as independent variables aimed at explaining that same theoretical construct.

5. Aday (2005) actually finds that the media limits the extent to which it shows vivid images of casualties; he also points out, however, that "scholars should spend more time exploring" exactly how the content of war coverage affects the audience's opinions."

6. A related possible problem is that people have little understanding of what constitutes relatively small or large numbers of casualties. Thus, though FGR provide casualty anchors, it is likely that people view those anchors quite differently. An example comes from an October 1945 Gallup survey on WWII that asked: "How many American soldiers, sailors, and airmen were killed in the war—just your best guess?" The median response of 300,000–500,000 was in line with the correct answer (~300,000 deaths). However, this accuracy obscures the wide variation in answers to the question. Only 12 percent of respondents were able to give an answer close to the correct answer, in the range of 260,000–449,000, while 25 percent guessed that the war dead stood at over 1 million, and 15 percent guessed that fewer than 200,000 died (Berinsky 2006). These types of misperceptions imply that there may be wildly different interpretations of what constitutes tolerance for a "small" or "large" number of casualties (Kull et al. 2003–04).

7. We thank an anonymous reviewer for making this point.

8. As mentioned, FGR explore other dependent variables including presidential approval, vote choice, and issue salience; however, these measures are not presumably meant as alternative ways to capture the same underlying construct of generalized war support.

Explaining Public Support for War

There is no shortage of theories to explain support for war, including work that focuses on the aforementioned casualty variable (e.g., Mueller 1973), the war's objective (Jentlestone 1992; Eichenberg 2005; Larson 1996), international support (Kull and Ramsay 2001), elite rhetoric (Zaller 1992; Berinsky 2006), and/or individual predispositions (Federico, Golec, and Dial 2005). FGR nicely situate their work relative to these and other theories, explaining that they focus on two key variables: prospective success and retrospective initial support.

PROSPECTIVE SUCCESS

FGR measure success by asking: "regardless of whether you think that the President did the right thing, would you say that the United States is very likely to succeed in Iraq, somewhat likely to succeed, not very likely to succeed, or not at all likely to succeed." Gelpi et al. (2005–06, p. 16) emphasize that this item is meant to gauge "eventual future success," and not "necessarily assessments of how the war is going right now" (Feaver and Gelpi 2004). On their questionnaire, FGR follow this question by probing the meaning of success by asking respondents which of six (or seven) options "best describes what 'success in Iraq' means to you?" (emphasis added). They find that greater than three-fourths of respondents define success as occurring when there is a stable and democratic government in Iraq, when Iraqis provide for their own security, or when Iraqis are able to live peaceful, normal everyday lives. The other options, which significantly fewer respondents chose, included if Iraq is prevented from supporting terrorism, is prevented from producing weapons of mass destruction, is not a threat to its neighbor, and if the economy is rebuilt.⁹

FGR then asked respondents which of eight distinct items they think "is the best way to *judge* whether the United States is likely to succeed in Iraq" (emphasis added). Possible responses included: what services are being provided by Iraqis, how soon Iraqi elections are held, whether Iraqis are cooperating with the United States and not protecting terrorists, how well the Iraqi economy is doing,¹⁰ the number of attacks against U.S. soldiers, the number of U.S. soldiers killed or wounded, the number of terrorists killed or arrested, and the amount of money the United States spends. Respondents tend to focus on one of the first three items just listed (particularly the cooperation with the U.S. item, as time progressed). More importantly, according to FGR, very few respondents measure success either in terms of U.S. casualties (4 percent), or terrorist casualties and arrests (2 percent).

9. The threat to neighbor item appeared only on their final survey.

10. This item was actually listed first on the questionnaire.

FGR conclude, therefore, that success is a distinct construct from casualty tolerance.

However, even if, as FGR argue, perceptions of success are not determined by casualty assessments, those perceptions may not be distinct from global evaluations of support for war.¹¹ Indeed, there is reason to think that evaluations of success could be driven by support for a given war. At the aggregate level, “perception of success” may have a clear meaning: it could vary overtime in reaction to the events on the battlefield. But it is not clear how best to give meaning to the cross-sectional variation in individual perceptions of success. The literature on the effect of perceptions of the economy on vote choice is instructive on this point. First, as Erikson (2004) notes (following Kramer 1983), because cross-sectional variation in perceptions of the economy represents variation in individual perceptions of a fixed quantity, cross-sectional variation in economic evaluations may be, in part, random noise and part determined by an individual’s political predispositions. Similarly, we might expect that cross-sectional variation in evaluations of future military “success”—a quantity with a presumably objective answer—may, in part, also be random noise. But given the partisan nature of patterns of support for the Iraq conflict (Jacobson 2007), this variation is probably less random noise than it is the product of partisan projection effects (Lord, Ross, and Lepper 1979). That is, people may use their political predispositions to assess the likelihood of success. Such projection effects could undermine our ability to effectively estimate causal relationships of interest.

Indeed, recent research has demonstrated that economic perceptions may be determined by vote choice, rather than the converse (Wlezien, Franklin, and Twiggs 1997; Anderson, Mendes, and Tverdova 2004; Erikson 2004).

11. In addition, we have some concerns related to question order: FGR first ask respondents to assess the likelihood that the United States will succeed, at which point it is plausible that respondents are evaluating the likelihood of success based on casualties. They then ask respondents to choose one of six (or seven) ways that captures how they *define* success. However, none of the options include anything approximating defining success as limiting casualties. Rather, the options all concern the future status of Iraq (and it does not appear that respondents could name “other” ways to define success). The query asking respondents how they *judge* success comes last—a point at which respondents were *just* forced to define success in a way unrelated to casualties. It is, therefore, not surprising that respondents did not state that they would judge success using casualties. Given what we know about the process by which people answer survey questions, it is possible that respondents view success, at least in part, in terms of casualties, but then were primed away from citing casualties because they were just forced to define success in ways orthogonal from casualties. Tourangeau, Rips, and Rasinski (2000, p. 218) explain that when a question provides an interpretive framework for the subsequent question, respondents often will assimilate their answers to the second question so as to be in line with the first. It may have been preferable for FGR to separate these items from one another on the survey, or, at least, experiment with alternative orderings to ensure the ordering does not lead to their particular results.

Similarly, just as the observed correlation between vote choice and economic perceptions is a result of voters bringing their economic assessments in line with their political judgments, the causal arrow between perceived success and latent generalized support for war could run from the later to the former, rather than vice-versa, as FGR argue. In practice, both measures could be determined by the same underlying political preferences such as partisan identification. However, since we can never fully account for all the variance in survey items with the measured characteristics of respondents—the R^2 of even the best fitting regressions fall far short of 1.00—simply controlling for the background correlates that we know affect opinion, such as partisanship, will not solve this endogeneity problem.

A cursory empirical analysis suggests that judgments of war success are indeed determined in large part by respondents' general political predilections (rather than careful assessments of ongoing events). In a survey on the Iraq war conducted by Knowledge Networks in August 2004, Berinsky (2006) included the FGR success measure. Mirroring the large partisan differences found on support for the war (Jacobson 2007), 85 percent of Republicans, but only 51 percent of Democrats thought that the United States was very or somewhat likely to succeed in Iraq.¹² These results are comparable to the partisan differences found with somewhat different forms of the "success" question asked by other survey organizations. In October 2004, The Program on International Policy Attitudes asked, "How confident are you that the US intervention in Iraq will succeed. Please answer on a scale of 0 to 10, with 0 being not at all confident and 10 being extremely confident." The mean score among Republicans was 7.0, but only 3.3 among Democrats. Similarly, in December 2005, the *Washington Post* asked, "All told, do you think the United States will win or lose the war in Iraq?" Eighty-nine percent of Republicans, but only 35 percent of Democrats thought that the United States would win. When a slightly different form of the question is asked—"All told, do you think the United States is winning or losing the war in Iraq?"—a similar partisan breakdown emerges: 82 percent of Republicans and 29 percent of Democrats believe the United States is winning.¹³ In sum, it appears that people's beliefs about a war's success depend in large part

12. These differences are highly statistically significant. In addition, the "success" question—like other measures of support for the war—exhibits a polarization pattern (Zaller 1992). The August 2004 Knowledge Network survey also measured the respondents' levels of political engagement. As levels of political engagement increased among Republicans, the estimates of perceived success also increase. Among Democrats, however, increasing political engagement decreased the estimates of success (see Berinsky 2006 for details).

13. These patterns of partisan polarization are not specific to the Iraq war. In June, 1999, with Bill Clinton leading the charge on U.S. intervention in Kosovo conflict, the *Washington Post* asked, "As of now, which side do you think won the Kosovo conflict: Serbia, or the United States and its European allies?" Sixty percent of Democrats but only 41 percent of Republicans thought the United States won.

upon where they sit politically. The point here is not that we should account for the political predilections of respondents when modeling support for war and stop there. If the solution were that simple, we could just control for partisanship and move on to assessing the effect of perceived success on support for war. The problem, we believe, may run deeper. As with economic voting, perceptions of success might be influenced by the respondents stand on the war, even controlling for measured partisanship. If, as we believe, the “success” variable taps the same underlying concept—support for war—as FGR’s dependent variable, statistically controlling for the measured characteristics of respondents will merely parse out some of variance common to the two measures; it will not establish the causal relationship between the two measures. In sum, we suspect that FGR’s “success” variable is an indicator of support for war, not a cause of support for war. Like other measures of war support, it is influenced in part by partisan predispositions.

INITIAL SUPPORT

FGR’s other key variable is initial support. They use this variable to capture the notion that “Retrospectively, voters are judging whether the decision to invade Iraq was the right one” (n.d., 15; also see Gelpi et al. 2005–06, p. 25). Specifically, they ask respondents, “I would like to know whether you think President Bush did the right thing by using military force against Iraq. Would you say that you strongly approve, somewhat approve, somewhat disapprove, or strongly disapprove of his decision?”

It is not clear to us whether FGR intend for this measure to capture the respondent’s current assessment of Bush’s past action of using force, or the respondent’s recall of what he or she thought when the conflict began. (The use of the present tense for the response categories suggests the former usage.) Either way, however, we worry about the extent to which respondents’ evaluation of the initial decision can be differentiated from their contemporaneous evaluation of the war (which is the presumed underlying dependent variable construct). Individuals who currently support the war may be substantially more likely to view the initial decision positively while those who oppose the war could see the launching of the conflict as having been the wrong thing to do. As in the case of the “success” measure, we suspect that the “initial support” variable may be best seen as an indicator of general support for the war.

MULTIPLE MEASURES OF SUPPORT FOR WAR?

The discussion of the problems inherent in FGR’s two independent variables leads us to the conclusion that perhaps these variables, like FGR’s casualty tolerance measure, are best understood as indicators of latent support for the war more generally. Treating these three measures as indicators of war

support rather than pieces of a causal story would fit well with the ‘‘multiple measures’’ approach to gauging support for war used in previous work in the public opinion and foreign policy tradition (see the earlier discussion).

Berinsky’s (2006) 2004 Iraq War survey suggests that the FGR independent variable measures indeed function as alternative indicators of support for war. In that survey, respondents were asked four questions relating to the Iraq War.¹⁴ The first item was similar in tone to FGR’s ‘‘initial support’’ question, and read, ‘‘Do you think the U.S. made the right decision or the wrong decision in using military force against Iraq? Do you feel strongly or not strongly that the U.S. made the [right/wrong] decision?’’ The second question measured general support for the war using the ABC News/Washington Post general evaluation question: ‘‘All in all, considering the costs to the United States versus the benefits to the United States, do you think the current war with Iraq has been worth fighting, or not? Do you feel strongly or not strongly that the war in Iraq [has/has not] been worth fighting?’’ The third item replicated FGR’s success measure exactly: ‘‘Regardless of whether you think the President did the right thing, would you say that the U.S. is: very likely to succeed in Iraq, somewhat likely to succeed in Iraq, not very likely to succeed in Iraq, not at all likely to succeed in Iraq.’’ A final item was not directly related to the FGR analysis but provides another indicator of support for the war. This question measured respondents’ views concerning the proper future course of U.S. action in Iraq asking, ‘‘There is some discussion about how many troops the U.S. should have in Iraq now. Do you think the number of U.S. troops in Iraq should be: increased a lot, increased some, maintained at current level, decreased some, decreased a lot, withdrawn completely.’’¹⁵

For the purpose of analysis, we rescaled all these variables to the 0–1 interval, with ‘‘1’’ indicating the highest level of support for war/level of perceived success and ‘‘0’’ indicating the lowest level on these variables. There is considerable variation in the median level of support for these propositions ranging from 0.33 (somewhat disagree) on the question of whether the war has been worth fighting to 0.66 (the United States is somewhat likely to succeed) on the FGR success question. However this variation obscures the fact that these four items are alternative indicators of a single underlying construct. We ran an exploratory principal components factor analysis on the four items and found strong support for a single-factor solution (the correlations of the four measures are

14. These questions were run as part of an experiment designed to assess the effects of information about casualties on support for the war in Iraq. There were not significant differences on any of these measures across conditions. Thus, for the purposes of presentation, we pool the data across the experimental conditions. We re-ran the factor analysis and scale analysis separately for each condition and achieved results that are identical to those presented here.

15. The results reported here are, if anything, even stronger if we remove this final item.

presented in Appendix A).¹⁶ The Eigenvalue for the first factor is 2.29, dwarfing the 0.07 value of the second factor.¹⁷ In addition a scale formed from the four items is extremely reliable ($\alpha = 0.83$). Furthermore, individual-level regression analysis (available in an online appendix to this article) demonstrates that these variables share common background correlates. While measures that are highly intercorrelated may have distinct properties—given the theoretic ambiguity of the FGR measures, the common strategy of tapping support for war through multiple measures, and the shared correlates of the variables of interest—it is more likely that these measures are multiple indicators of the same construct than are the outcome of a complicated causal model. We, therefore, have highly suggestive evidence that the FGR success question is simply another indicator of general support for war.

A recent study by Federico, Golec, and Dial (2005) echoes these results. The authors seek to explain support for military action against Iraq by combining six indicators. Two of these indicators capture general military action (e.g., take action or use diplomatic means; take action or wait on the United Nations). Three others focus on likely success (e.g., will it stabilize the situation; will it reduce the threat of terrorism; will it further American interests in the Middle East), and a final item measures justification for taking action. These data produce an α of 0.90, and a principal exploratory factor analysis shows an Eigenvalue for the first factor of 3.36 and a value 0.20 on the second factor.

In sum, we suspect that the two independent variables of interest may actually be better indicators of latent support for the war in Iraq than their dependent variable.¹⁸ That is, it appears that the average survey respondent may use the “initial support” and “likelihood of success” items as contemporaneous measures of support for war.

16. We used the principle factor method in STATA 9.0. We come to the same conclusion using the principal-components and iterated principle factor methods.

17. The factor loadings for the four items are all in approximately the same range. Ranging from 0.52 for the “increase troops question” to 0.89 for the question whether the war was “worth it” (the factor score for the FGR success measure is 0.68 and the score for the “right decision” question is 0.88). Thus, the weakest of the four items is the question about the proper level of U.S. troops, but even this seems to be a strong indicator of support for the war.

18. Earlier, we suggested that perceptions of success may be caused by general war support. Here we claim that the measures of perceptions of success and general war support may be indistinguishable. We do not see these points as contradictory. In theory, support for war could influence specific evaluations of success. So it could be that FGR’s theoretic story is problematic. Alternatively, on an empirical level, we never measure the latent construct of interest, general war support. Instead, with the “mistake” question and the “considering the costs and benefits” questions, we simply measure additional indicators of the latent construct. In this case, it could be that our more general measures *and* the success question are both just indicators of generalized support for the war. Our factor analysis supports this position. A similar logic applies to the relationship between the initial support and prospective success variables.

Discussion

The discussion of the nature of the variables involved in the FGR analysis leads directly into questions of causality. The exchange between Gelpi and Mueller in *Foreign Affairs* highlighted the fact that these opposing scholars are arguing over the causal direction of similar variables. FGR claim that casualty tolerance is a function of (initial) support for the war effort and perceived success. Mueller essentially argues that overall support (which, by our accounting, might be indexed by FGR's retrospective support variable) is a function of casualties. FGR attempt to estimate their model using cross-sectional data, but given the common structure to the variables, such analysis will be heavily dependent on the (potentially questionable) modeling assumptions chosen by the researcher. In short, it is very difficult to establish causal relationships in cross-sectional nonexperimental data with such highly related variables.¹⁹ We therefore find ourselves in agreement with former Gallup Vice President David Moore who concludes, "The causal model cannot be proved, as least by the data obtained by the three authors [i.e., FGR]; in this case, causality is more an act of faith than a provable dynamic."²⁰

The implication is that more research is needed to sort out which of various dynamics drive public opinion about war. Two particular, sometimes competing, forces that demand more explicit study are longer term values and elite rhetoric (Herrmann et al. 1999). For instance, Fedirico, Golec, and Dial (2005) show that patriotism, and to a greater extent nationalism (especially among certain types of individuals) shape war attitudes. And, of course, a substantial amount of research documents the importance of media and politician framing and priming of issues in shaping citizens' opinions.²¹ Exactly how these dynamics operate will depend on such things as the mix of

19. FGR acknowledged this endogeneity problem in earlier versions of their work where they sought to overcome it by use of simultaneous equation analysis. While this strategy is potentially useful, in practice such analysis requires the use of appropriate instruments—variables that affect one of the variables of interest (say perceived success) but not the other potentially endogenous variables of interest (initial support and casualty tolerance). FGR used the respondents' views concerning success in hypothetical interventions in Yemen and East Timor and their level of confidence in the President and various departments in the U.S. military as instruments for the perceptions of success. These instruments seem highly suspect to us. For instance, analysis of data from the 2004 National Election Study (available from the authors upon request) suggests that respondents' support for the military and their confidence in Bush's leadership capabilities predict support for the war in Iraq, even controlling for relevant political variables, such as party identification. Thus, it appears that FGR's measures do not meet the requirements of instrumental variable analysis; their instruments, in fact, are likely correlated with their dependent variable.

20. Quoted at http://www.mysterypollster.com/main/2005/12/polling_the_str.html. Accessed April 25, 2006.

21. Baum and Groeling (2006) persuasively demonstrate the impact of elite rhetoric in the case of foreign policy issues (see also Berinsky 2006).

messages from competing elites (of varying credibility), the number of messages, the vividness of the messages, and individual-level variables such as sophistication and values (e.g., Iyengar 1991; Druckman 2001; Sniderman and Theriault 2004). These and other factors undoubtedly shape what people think about the war and its likely success.²²

FGR's research agenda accentuates the extent to which our understanding of attitudes toward war has evolved. It is increasingly clear that citizens can form systematic opinions about war, and policy-makers care what citizens think. The challenge, then, is to isolate the forces that shape public opinion about war. We suspect that these factors do not differ from those found to affect opinion on other political issues and events, including partisan political conflict, elite rhetoric, and individual-level variables such as sophistication. Much more research needs to be done—both in surveys and through experiments—to tease out the dynamics of opinion. Even if, as we argue, support for war is heavily influenced by pre-existing political judgments and the balance of elite rhetoric, the complexity of opinion formation ensures that other factors, such as attachment to ethnic groups and political values, may shape support for war in important ways. Research on public opinion and foreign policy should, therefore, pay closer attention to the voluminous research on opinion concerning domestic politics.

Our main concerns about FGR's research—that their dependent variable may not be reliable, that their two independent variables may be alternative measures of overall war support and the possibility of omitted variable bias—are issues that all researchers need to consider. We believe this will happen as the many researchers, from varying scholarly perspectives, who study war opinions continue to interact and exchange ideas.

22. FGR virtually ignore the impact of political and media elites in their individual level analyses (however, see Gelpi and Reifler 2005). While the authors acknowledge the immense volume of coverage of the war, they offer little discussion of potential fluctuations in the tone of coverage both in balance—positive or negative coverage—and in focus—episodic or thematic coverage. FGR do include a measure of elite influence; however, we worry that this measure does not adequately capture the relevant construct. Their “perceived elite consensus” variable asks “Do you think America’s political leaders—both Republican and Democrat—agree that the U.S. troops should remain in Iraq until the new Iraqi government is stable and secure?” The presumption of such a measure is that mass communication effects are mediated by conscious recognition of an elite consensus (or lack of consensus), which citizens then purposefully connect to their overall opinions. We find this unlikely, as the impact of elite consensus more likely comes from the actual mix of elite cues available to the public. That is, a correlation between war support and elite consensus does not stem from citizens’ recognizing the consensus (or not) but rather, from the uniformity of the cues that everyone receives (e.g., in support of the war). Citizens process elite cues, not merely perceptions of consensus (Zaller 1992, pp. 98–102). Moreover, we expect that this relationship is moderated by political awareness and partisanship (Zaller 1992; Druckman and Holmes 2004). An interaction term among the relevant measures would be necessary to discern these effects. Even if perception of elite consensus matters, it is possible that individuals may not recall their perceptions if they arrive at their attitudes with on-line evaluations (Druckman and Lupia 2000).

Supplementary Data

Supplementary data are available online at <http://pubopq.oxfordjournals.org/>.

References

- Aday, Sean. 2005. "The Real War Will Never Get Televised." In *Media and Conflict in the 21st Century*, ed. Phil Seib. Palgrave.
- Aldrich, John H., Christopher Gelpi, Peter Feaver, Jason Reifler, and Kristin Thompson Sharp. 2006. "Foreign Policy and the Electoral Connection." *Annual Review of Political Science* 9:477–502.
- Anderson, Christopher J., Silvia Mendes, and Yuliya V. Tverdova. 2004. "Endogenous Economic Voting." *Electoral Studies* 23:683–708.
- Baum, Matthew A., Tim Groeling. (2006) "Crossing the Water's Edge" Unpublished Manuscript, UCLA.
- Bartels, Larry M. 2003. "Democracy With Attitudes." In *Electoral Democracy*, ed. Michael Bruce MacKuen, and George Rabinowitz. Ann Arbor: The University of Michigan Press.
- Berinsky, Adam J. (2006) "Assuming the Costs of War." Unpublished Manuscript, Massachusetts Institute of Technology.
- Burk, James. 1999. "Public Support for Peacekeeping in Lebanon and Somalia." *Political Science Quarterly* 114:53–78.
- Druckman, James N. 2001. "The Implications of Framing Effects for Citizen Competence." *Political Behavior* 23:225–56.
- Druckman, James N., and Justin W. Holmes. 2004. "Does Presidential Rhetoric Matter?" *Presidential Studies Quarterly* 34:755–78.
- Druckman, James N., and Arthur Lupia. 2000. "Preference Formation." *Annual Review of Political Science* 3:1–24.
- Eichenberg, Richard C. 2005. "Victory Has Many Friends." *International Security* 30:140–77.
- Erikson, Robert S. (2004) "Macro vs. Micro-Level Perspectives on Economic Voting." Paper presented at the annual meeting of the Political Methodology Society, Stanford University, July 29–31.
- Everts, Philip, and Pierangelo Isernia. 2005. "The Polls—Trends: The War In Iraq." *Public Opinion Quarterly* 69:264–323.
- Feaver, Peter D., and Christopher Gelpi. 2004. *Choosing Your Battles*. Princeton, NJ: Princeton University Press.
- Federico, Christopher M., Agnieszka Golec, and Jessica L. Dial. 2005. "The Relationship Between the Need for Closure and Support for Military Actions Against Iraq." *Personality and Social Psychology Bulletin* 31:621–32.
- Gartner, Scott Sigmund, and Gary M. Segura. 2000. "Race, Casualties, and Opinions in the Vietnam War." *Journal of Politics* 62:115–46.
- Gartner, Scott Sigmund, Gary M. Segura, and Michael Wilkening. 1997. "All Politics Are Local." *Journal of Conflict Resolution* 41:669–94.
- Gelpi, Christopher, Peter D. Feaver, and Jason Reifler. 2005–06. "Casualty Sensitivity and the War in Iraq." *International Security* 30:7–46.
- Gelpi, Christopher, and John Mueller. 2006. "The Cost of War." *Foreign Affairs* 85:139–44.
- Gelpi, Christopher, and Jason Reifler. 2005. "What the Public Really Thinks of Iraq." *Rocky Mountain News* December 12:44A.
- Gelpi, Christopher, Jason Reifler, and Peter D. Feaver (N.d.) "Iraq the Vote." *Political Behavior*. Forthcoming.
- Herrmann, Richard K., Philip E. Tetlock, and Penny S. Visser. 1999. "Mass Public Decisions to Go to War." *American Political Science Review* 93:553–73.
- Hertwig, Ralph, Thorsten Pachur, and Stephanie Kurzenhäuser. 2005. "Judgments of Risk Frequencies: Tests of Possible Cognitive Mechanisms." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31:621–42.
- Holsti, Ole R. 2004. *Public Opinion and American Foreign Policy*. Ann Arbor: University of Michigan Press.

- Iyengar, Shanto. 1991. *Is Anyone Responsible?* Chicago: The University of Chicago Press.
- Jacobs, Lawrence R., and Robert Y. Shapiro. 1995. "The Rise of Presidential Polling." *Public Opinion Quarterly* 59:163–95.
- Jacobson, Gary C. 2007. *A Divider, Not a Uniter*. New York: Pearson Longman.
- Jentleson, Bruce. 1992. "The Pretty Prudent Public." *International Studies Quarterly* 36:49–74.
- Kahneman, Daniel, Ilana Ritov, and David Schkade. 1999. "Economic Preferences or Attitude Expressions?" *Journal of Risk and Uncertainty* 19:203–35.
- Klarevas, Louis J. 2002. "The 'Essential Domino' of Military Operations." *International Studies Perspectives* 3:417–337.
- Klarevas, Louis J., Christopher Gelpi, and Jason Reifler. 2006. "Correspondence: Casualties, Polls, and the Iraq War." *International Security* 31:417–37.
- Kramer, Gerald H. 1983. "The Ecological Fallacy Revisited." *American Political Science Review* 77:92–111.
- Kull, Steven, and Clay Ramsey. 2001. "The Myth of the Reactive Public." In *Public Opinion and the International Use of Force*, ed. Phillip Everts, and Pierangelo Isernia. London: Routledge.
- Kull, Steven, and I.M. Destler. 1999. *Misreading the Public*. Washington, D.C: Brookings.
- Kull, Steven, Clay Ramsay, and Evan Lewis. 2003–04. "Misperceptions, the Media, and the Iraq War." *Political Science Quarterly* 118:569–98.
- Larson, Eric V. 1996. *Casualties and Consensus*. Santa Monica, CA: Rand.
- Lichtenstein, Sarah, Paul Slovic, Baruch Fischhoff, Mark Layman, and Barbara Combs. 1978. "Judged Frequency of Lethal Events." *Journal of Experimental Psychology: Human Learning and Memory* 4:551–78.
- Lord, Charles, Lee Ross, and Mark Lepper. 1979. "Biased Assimilation and Attitude Polarization." *Journal of Personality and Social Psychology* 37:2098–109.
- Mueller, John E. 1973. *War, Presidents, and Public Opinion*. Lanham, MD: University Press of America.
- Mueller, John. 2005. "The Iraq Syndrome." *Foreign Affairs* 84:44–50.
- Schläpfer, Felix, and Nick Hanley. 2006. "Contingent Valuation and Collective Choice." *Kyklos* 59:115–35.
- Shane, Scott. 2005. "Bush's Speech on Iraq Echoes Analyst's Voice." *New York Times*, December 4:1–35.
- Sniderman, Paul M., and Sean M. Theriault. 2004. "The Structure of Political Argument and the Logic of Issue Framing." In *Studies in Public Opinion*, ed. Willem E. Saris, and Paul M. Sniderman. Princeton, NJ: Princeton University Press.
- Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Wlezien, Christopher, Mark Franklin, and Daniel Twiggs. 1997. "Economic Perceptions and Vote Choice." *Political Behavior* 19:7–17.
- Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.