

THE POLLS—REVIEW

REPLY TO BERINSKY AND DRUCKMAN: SUCCESS STILL MATTERS

CHRISTOPHER GELPI
JASON REIFLER

We appreciate Berinsky and Druckman’s thoughtful critique of our recent work regarding the public’s willingness to bear the costs of war (Berinsky, A., J. Druckman. 2007. “Public Opinion Research and Support for the Iraq War.” *Public Opinion Quarterly* 71:126–41). We are fortunate to have such pre-eminent scholars constructively engaging our work. In their review, the authors raise four major concerns about our analysis: (1) our measurement of “war support,” (2) our measurement of perceptions of success, (3) our claim that perceived success is a cause of war support (measured as casualty tolerance), and (4) our lack of attention to elite rhetoric as a cause of war support. We address each of these concerns below. While we remain confident in our original conclusions, we believe that their comments identify important questions that remain unanswered in this area of research.

Casualty Tolerance as “War Support”

Berinsky and Druckman argue that our principal dependent variable is a poor measure of support for war. Instead, they argue for a more generalized measure of support. We utilize the casualty tolerance dependent variable because it directly relates to the question most central to academics and policy-makers—how does the public respond to casualties? The centrality of casualty tolerance to the question of war support is nearly impossible to overstate: the public’s response to casualties is of great concern to civilian and military policy makers; the media make American casualties the dominant theme of its coverage, and even our foreign adversaries, such as Saddam Hussein and Osama Bin Laden,

CHRISTOPHER GELPI is with Political Science, Duke University, PO Box 90204, Durham, NC 27708, USA. JASON REIFLER is with Political Science, Georgia State University, 1019 General Classroom Building, Atlanta, GA 30303, USA. Peter Feaver was unable to attach his name to this response because of his concurrent service with the National Security Council. He agrees with the conclusions presented here. Address correspondence to Jason Reifler; e-mail: poljar@langate.gsu.edu.

base their strategies on their belief that the American public is weak-willed and will fold if soldiers die in conflict.

The crux of this debate over casualties focuses on a policy choice facing the American public: at what level of costs does the public decide that it prefers to withdraw rather than continue to suffer the costs of casualties? In order to address this central theoretical and policy question, we constructed a dependent variable that focuses on the impact of casualties on respondents' policy preferences. Our question specifically asks respondents if they are willing to support the continuation of an ongoing conflict if the United States suffers a particular number of casualties. Our measure allows us to study the effects of war costs on support in the context of a cross-sectional survey (despite the fact that casualties are functionally constant) because our dependent variable can be understood as a within-subject experiment. Each respondent is asked about supporting the war after being exposed to successive treatments about the number of casualties the United States could suffer.¹

The "mistake" question, on the other hand, does not lend itself well to studying policy preferences or costs in cross-sectional studies. Mueller (1973) adopts the mistake question because it is the only one that "tapped a sort of generalized support for the war and that was asked repeatedly in both wars" and "seems to be a sound measure of a sort of general support for the war" (Mueller 1973, p. 43). Thus this question became the standard question not because it ties to a specific theoretical or empirical concern, but because pollsters asked about it during both the Korean and Vietnam Wars. Among the very notable weaknesses of the mistake question is that it "says little about policy preferences at any given moment" (Mueller 1973, p. 43).

We are fundamentally interested in the conditions under which the (American) public is willing to support policies that result in having to bear the costs of war. Moreover, we argue that the willingness to bear costs is, *inter alia*, conditional on perceptions that the war effort will be successful. It is unclear how we could test this proposition without somehow incorporating costs, especially in cross-sectional surveys. Nonetheless, we recognize that answers to questions about cost may depend on how the choice is framed, the information available to respondents, frequencies versus "prototypical incidents," and so on. We sincerely hope that our work to date is not the final word on the subject—costs and war clearly is a topic that will be studied for years to come. As of right now, Berinsky and Druckman essentially argue that costs are too hard to study, and therefore should be avoided (even though they advocate use of "mistake" and

1. This dependent variable is easily exported to other experimental settings, such as hypothetical missions, that a "mistake" question would not easily translate to. In our book, we develop this dependent variable for a number of different scenarios and show that the concept has a strong validity. Unlike the mistake question, our question can be asked *before* a war starts. We believe that a complete understanding of support for war—and how support responds to the costs of war—requires measures that can be asked before a war starts and again once underway.

“worth it” questions which are implicitly about costs). Even if our measure is imperfect, it is still less imperfect than *not* incorporating costs into war support.

Nonetheless, Berinsky and Druckman rightly point out that there may be slippage between how much cost respondents *say they will bear* and how much cost *they will actually bear*. At the same time, the amount of slippage between survey estimates of a willingness to bear costs and the “actual behavior” of bearing costs is a question that may never be fully resolved. First, even the most tolerant Internal Review Board (IRB) is unlikely to allow us to start a war to see how well people estimated the costs they were willing to bear. Second, this slippage suggests that answering survey questions is a fundamentally different behavior than whatever behavior *actually* constitutes bearing the costs of war. Citizens are rarely asked to directly pay the costs of war. If citizens express their preferences about war costs through answering surveys and voting, Berinsky and Druckman’s concerns about slippage that stem from the consumer behavior analogy—where spending money differs qualitatively from answering survey questions—may be overstated.

We fully recognize that measuring the impact of costs is difficult. We disagree that the solution to the problem of incorporating costs is to ignore them. Paying the cost of war is the fundamental question in the war support literature (see, for example, Mueller 1973, 2005; Larson 1996; Gartner and Segura 1998; Burk 1999; Klarevas 2000). General war support questions can mislead when making inferences about what policies the public would like government to pursue. There is rich and important variation in questions about the Iraq War. At the time of this writing, for example, disapproval of President Bush’s handling of the Iraq War is over 70 percent, while somewhere between 50 percent and 60 percent think the original decision to attack Iraq was a mistake, but less than 40 percent of the public wants to set a specific date for the complete withdrawal of US forces from Iraq. If one wants to understand the dynamics of how respondents construct their attitudes toward the Iraq War, therefore, one would not want to conflate disapproval of Bush with support for withdrawing US forces. Opinion about the war is likely to be a bundle of related but not identical attitudes and our model explicitly recognizes that fact.

Perceptions of Success

Berinsky and Druckman also raise concerns about the measurement of our key independent variable: perceptions of success. The authors are correct in pointing out that the public must define “success” as something other than “few casualties” if our argument is to be anything more than a tautology. We defend this claim by asking respondents how they define success and how they judge whether the United States is succeeding. We show that only 4–5 percent of respondents state that they judge whether the United States is succeeding based on the number of American deaths. Berinsky and Druckman contend that

this result is contaminated by the fact that our previous question—which asks respondents to select a definition of success from a list—does not offer anything related to “limiting casualties” as an option. Having just forced respondents to define success in other terms, they argue that respondents were primed not to judge success in terms of casualties. However, in our October 2004 survey, we dropped the “success definition” question but kept the “success metric” question. Results for the metric question were unchanged—less than 5 percent of respondents stated that they measured progress toward success in terms of US casualties. Indeed, this holds true despite the fact that our casualty tolerance question arguably primed respondents to think about casualties.

Another concern is that perceptions of success may be a function of other attitudes and demographic factors—especially party identification. We agree that perceptions of success are likely to be a consequence of a variety of factors and that the construction of these perceptions is an interesting question. We further agree that this is perhaps the most fruitful area for future research. Limited journal space has prevented a nuanced discussion of this subject, but our forthcoming book devotes an entire chapter to endogenizing both perceptions of success and perceptions of whether Bush “did the right thing” in attacking Iraq. Party identification has an effect on these attitudes and we are explicit in stating that we do not view the analyses in our article as suggesting that partisanship has no impact on tolerance for casualties (Gelpi, Feaver, and Reifler 2005–06, p. 40).

The analyses of casualty tolerance in our article, however, control for partisanship and a variety of other demographic factors. Thus any effects that perceptions of success and whether Bush “did the right thing” have on casualty tolerance in our analyses exist over and above the direct impact of party identification. Party identification may have additional indirect effects on casualty tolerance, but any causal relationship between partisanship and perceptions of success and “right thing” does not bias our estimates of the direct effect of those attitudes on casualty tolerance.² In short, party identification has a modest direct impact on casualty tolerance, but the analyses in our book suggest that its greatest impact is through its indirect effect through its influence on perceptions that Bush “did the right thing.” Neither of those facts has any influence

2. Future research on the direction of causal influence among attitudes toward war should also investigate the conditions under which partisan predispositions may be endogenous to policy attitudes. For example, the last several years have brought dramatic changes in party ID marginals. According to Pew, party identification was roughly 43 percent Democrat–43 percent Republican in 2002. Through surveys of March 2007, party ID is now 50 percent Democrat–35 percent Republican. Moreover, these changes are not just limited to Pew, but found across polling firms (see footnote 7 for Charles Franklin’s compilation of presidential approval polls). Berinsky (2007) argues that partisan predispositions are important for understanding war support and that what actually happens in war has little to no independent effect. These substantial shifts in party identification would suggest that the relationship is more complex. While it is hard to say for sure what has driven the substantial change in party ID over the last several years, we strongly suspect that the Iraq War has played the central role in this change.

on the conclusions that we draw in our article about the impact of perceptions of success on casualty tolerance.

Success as a Cause of Casualty Tolerance

Berinsky and Druckman question the causal claims we make regarding the impact of perceptions of success on casualty tolerance in three ways. First, they argue that our measures of casualty tolerance, perceived success, and whether force was the “right thing” to do are all indicators of a general “support for the war” construct. To bolster their claim, Berinsky and Druckman perform a factor analysis of four questions about Iraq and claim that these variables are best understood as a single latent variable representing general Iraq War support.

It is not surprising that the factor analysis of four questions about the war finds the war to be an important structuring dimension. We notice several aspects of their analysis. The “right thing” and “mistake” questions have nearly identical factor loadings (~ 0.88), but the factor loadings for “success” (0.68) and “increase troops” (0.52) are substantially lower. There does appear to be some difference between the questions—if all were exactly equal substitutes for one another, they should all have similar factor loadings. In fact, Berinsky and Druckman freely admit that their policy question about withdrawing troops—which we expect to be most closely related to our casualty tolerance variable—does not scale well with the other variables. Factor and alpha scores do better when the question is removed! Berinsky and Druckman find that “general support” measures can be surprisingly divorced from policy preferences. The inability of the Berinsky policy choice question to load well confirms our interpretation that our casualty tolerance measure represents an underlying construct beyond “rightness” or “mistake.” Moreover, we believe it validates our strategy to see how different aspects of “general support” (like retrospective views of “rightness” and “success”) explain policy preferences such as continued support of a mission in the face of escalating human costs.

Moreover, Berinsky and Druckman point to a low eigenvalue for a second factor as evidence that there is only one dimension of Iraq attitudes. This analysis of the dimensionality is hampered in two important ways. First, it is hard to find support for multiple dimensions in exploratory factor analysis using just three or four indicators. A common rule of thumb in SEM is that latent dimensions should have *at least* three indicators per dimension. Under that requirement, Berinsky and Druckman only have enough indicators to find one dimension. Therefore, finding only one dimension should hardly be surprising. Second, Berinsky and Druckman’s factor analysis assumes that the factors are orthogonal. We never claim that the three dimensions should be perfectly orthogonal—there is clearly some common structure to the different facets of opinion concerning Iraq. Our measures of success, “right thing,” and casualty

tolerance are correlated with one another. We would be surprised if they were not so. Thus we would not expect to see an absolutely clear multifactor solution, in part because we are saying there are subtle yet substantively important differences in what these variables measure. Berinsky and Druckman find there are no unsubtle differences, and therefore conclude that all questions about Iraq mean exactly the same thing.

It is well accepted in political science that there are subtle, yet important distinctions between prospective and retrospective evaluations—particularly when it comes to perceptions of the economy. We simply extend this perspective by suggesting that there may also be subtle but substantively important differences in how prospective and retrospective evaluations matter for war and in explaining policy preferences about war.

But the strongest evidence against the Berinsky and Druckman claim of a single dimension is their own argument that perceptions of success may be endogenous to war support. If success and casualty tolerance are simply different indicators of the same underlying construct, then causal direction is irrelevant. If causal direction matters, then they are distinct enough concepts to worry about causal direction. We view these attitudes as sufficiently distinct as to warrant the investigation of causal relationships among them. We share Berinsky and Druckman's concerns about establishing causal direction, and in our book we make every effort to do so with the available data—including aggregate time-series data, two-stage least-squares analysis of cross-sectional data, and question wording experiments.³ Nonetheless, we agree with Berinsky and Druckman that establishing causal direction among attitudes is difficult and additional experimental research will be extremely important in teasing out these effects.

Elite Rhetoric

Berinsky and Druckman also question the validity of our causal inferences because they believe that we have not paid sufficient attention to the influence of political and media elites. Moreover, Berinsky (2007) characterizes our argument as stating that only “unmediated battlefield events” matter in shaping citizen opinions.⁴ We disagree with this characterization of our argument both

3. As they acknowledge, we use simultaneous equations to address endogeneity. The authors are skeptical of these efforts because they fear that our instruments are correlated with our endogenous variables. The analyses in our forthcoming book make every effort to investigate the plausibility of our exclusion restrictions. Specifically, we relaxed each restriction individually and found that none of the excluded variables were statistically significant when allowed to appear in the equation from which we had excluded them. The accuracy of simultaneous equations analyses are always subject to the validity of the specification assumptions, and these assumptions can never be fully tested.

4. See Johnson and Tierney (2006) for compelling evidence that perceptions of success need not overlap with actual battlefield success.

because we never make such an assertion, and because we agree with them that the impact of elite priming, framing, and agenda setting on the formation of public perceptions of success and the justification for war are extremely important issues that require further investigation. At the same time, we would also contend that arguments which emphasize the pre-eminence of elite cues (such as Berinsky 2007) are not wholly satisfying either. For example, an elite cues approach does nicely explain the high levels of polarization in Iraq attitudes. But an elite-driven perspective does not help explain why the rank and file members of the Democratic Party opposed the Iraq War and sought withdrawal from the conflict long before their leaders were articulating this position.

More generally, as we noted above, explaining how individuals form perceptions of success and whether using force is the “right thing” is an extremely complex and important question that we only begin to address in our book. Berinsky and Druckman’s critiques—while certainly less pleasing than sycophantic applause—will undoubtedly help us in the future. If there is any one point where we most agree with Berinsky and Druckman, it is the need for more research in this area, especially experimental research. We enthusiastically reiterate Berinsky and Druckman’s call for more research in this area.

Whatever these studies may discover about the impact of elite rhetoric versus “battlefield events,” we would also note that identifying elite messages as a source of public attitudes raises the question of how elites form their own beliefs regarding military conflict. The elite cues approach of explaining mass opinion says absolutely nothing about how elites may form judgments about the war. Even if Berinsky and Druckman are correct that the public is a hollow vessel that unthinkingly echoes the views of preferred elites, it still fundamentally does not tell us much about how future policy choices will be shaped by a war’s progress (or lack thereof). We would suggest that elites may also focus on success in forming their opinions. If so, then success may affect public opinion through its influence on elite rhetoric.

Bush Rhetoric, Perceived Success, and War Support

Finally, while conclusive evidence on the interplay of elite rhetoric and “battlefield events” in shaping public attitudes must await future research, we may be able to gain some tentative insight in the intersection of these variables by observing the public’s reaction to major communication efforts by the Bush Administration which Berinsky and Druckman note have been associated with our research.

In December 2005, President Bush gave a series of speeches leading up to the parliamentary elections in Iraq. We recognize that key elements of those speeches and the accompanying National Strategy for Victory in Iraq

contained rhetorical approaches consistent with our research findings that give explanatory power to perceptions of success.⁵ While we do not have definitive data available to us at the individual level on the impact of this elite rhetoric, we do have some aggregate data that can at least tentatively speak on this issue.⁶

The President began rolling out this communication effort on November 30, 2005 with a much publicized speech in Annapolis, MD, and concluded it with a primetime televised address from the Oval Office on December 18, 2005 in which he reflected upon the success of the Iraqi elections. The rhetorical effort was dramatized by the sharp contrast with the most salient Democratic alternative—Congressman Jack Murtha’s proposal for complete withdrawal within six months. Interestingly, Murtha’s opinion is also consistent with our model—he once supported the war, but as he saw success as impossible, he was unwilling for the United States to incur additional costs.

Prior to Bush’s speeches, public approval of the Bush Administration was at a low point. ABC news found that approval of Bush’s handling of the Iraq War was at 36 percent on November 2, and his overall presidential approval rating was at 39 percent. Just a few weeks later, on December 18, in the wake of the “victory” speeches and the successful Iraqi election, approval of Bush’s handling of Iraq jumped to 46 percent and his overall approval moved up to 47 percent. This 10-point bounce was not the result of a few isolated polls, but represented an upward shift across the board that lasted for several weeks. Charles Franklin’s compilation of presidential approval polls, for example, shows a clear bounce of about 10 points lasting through the month of December and into January 2006.⁷

But by March of 2006, President Bush found his approval numbers sagging again and he gave another series of speeches on March 13, 20, and 29 emphasizing the progress he saw in Iraq. The President also held a press conference on March 21 that focused largely on Iraq. This time, however, the President lacked a “real-world” successful event to which he could anchor his speeches. As a result, the presidential rhetoric had no perceptible impact on public attitudes toward the war whatsoever. In early March, the President’s approval ratings—both on Iraq and overall—were at approximately 35 percent and they remained unchanged throughout early April.

These contrasting events suggest that elite rhetoric matters in some circumstances and not in others. We cannot say with confidence what these

5. We would note, however, that those who might seek to increase public pressure to withdraw from Iraq could just as easily make extrapolations from our research to craft a communication strategy that seeks to demonstrate prospects for success in Iraq are dim.

6. For the claim that perceptions of success do not matter, we would actually have to know whether elites were incorporating judgments of success or failure in their rhetoric. Our argument would still have substantial merit if appeals to success could persuade the public better than other appeals. Our argument would be even stronger (and be normatively more appealing) if the persuasive power of appeals to success is contingent on actual signs of success or progress.

7. Data available at <http://www.pollster.com/presbushapproval.php>.

circumstances might be, nor can we say with certainty what the mechanism is by which elite messages matter. The above examples suggest that elite messages matter in combination with events in the “real world” and that they may have little influence absent such anchors.

Conclusion

In sum, we believe the evidence continues to support our claim that perceptions of the likely success of a military mission play an important—and even primary—role in determining the public’s willingness to pay costs in support of that mission. Data at both the aggregate and individual level appear to support this claim. Berinsky and Druckman never offer evidence that directly challenges our finding. Their critique rests solely on measurement and estimation issues that *might* threaten our inferences. We believe that we have answered those critiques. However, Berinsky and Druckman have drawn attention to several important areas that we believe are in need of additional research: measurement of cost sensitivity, how individuals construct their perceptions of the likely success of a mission, and the role of elite rhetoric in shaping opinion about war. We especially echo their plea for careful experimental work on this subject that can strengthen our ability to understand the formation of public attitudes toward war.

References

- Berinsky, Adam J. 2007. “Assuming the Costs of War: Events, Elites, and American Public Support for Military Conflict.” *Journal of Politics* 69:975–97.
- Berinsky, Adam J., James N. Druckman. 2007. “Public Opinion Research and Support for the Iraq War.” *Public Opinion Quarterly* 71:126–41.
- Burk, James. 1999. Public Support for Peacekeeping in Lebanon and Somalia: Assessing the Casualties Hypothesis. *Political Science Quarterly* 114:53–78.
- Gartner, Scott S., Gary M. Segura. 1998. “War Casualties and Public Opinion.” *Journal of Conflict Resolution* 42:278–320.
- Gelpi, Christopher., Peter D. Feaver, Jason. Reifler. 2005–06. “Success Matters: Casualty Sensitivity and the War in Iraq.” *International Security* 30:7–46.
- Gelpi, Christopher, Jason. Reifler, Peter D. Feaver. 2007. “Iraq the Vote: Retrospective and Prospective Foreign Policy Judgments on Candidate Choice and Casualty Tolerance.” *Political Behavior* 29:151–74.
- Johnson, Dominic D. P., Dominic. Tierney 2006. *Failing to Win: Perceptions of Victory and Defeat in International Politics*. Cambridge: Harvard University Press, p. 360.
- Klarevas, Louis J. 2000. “Trends: The United States Peace Operation in Somalia,” *Public Opinion Quarterly* 64:523–40.
- Larson, Eric V. 1996. *Casualties and Consensus: The Historical Role of Casualties in Domestic Support for U.S. Military Operations*. Santa Monica: Rand, p. 126.
- Mueller, John E. 1973. *War, Presidents, and Public Opinion*. New York: Wiley, p. 300.
- Mueller, John E. 2005. “The Iraq Syndrome.” *Foreign Affairs* 84:44–55.