EXPERIMENTS ON PARTISANSHIP AND PUBLIC OPINION:

PARTY CUES, FALSE BELIEFS, AND BAYESIAN UPDATING

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF POLITICAL SCIENCE

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

John G. Bullock

June 2007

ii

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

(Paul M. Sniderman) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

(Simon Jackman)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____
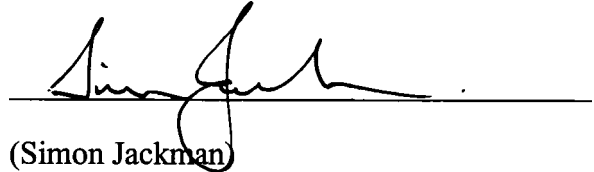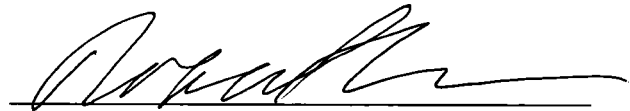
(Robert C. Luskin)

Approved for the University Committee on Graduate Studies.

iii

iv

Preface

This dissertation contains three parts—three papers. The first is about the effects of party cues on policy attitudes and candidate preferences. The second is about the resilience of false political beliefs. The third is about Bayesian updating of public opinion. Substantively, what unites them is my interest in partisanship and public opinion. Normatively, they all spring from my interest in the quality of citizens' thinking about politics. Methodologically, they are bound by my conviction that we gain purchase on interesting empirical questions by doing things differently: first, by bringing more experiments to fields still dominated by cross-sectional survey research; second, by using experiments unlike the ones that have gone before.

Part 1: It is widely believed that party cues affect political attitudes. But their effects have rarely been demonstrated, and most demonstrations rely on questionable inferences about cue-taking behavior. I use data from three experiments on representative national samples to show that party cues affect even the extremely well-informed and that their effects are, as Downs predicted, decreasing in the amount of policy-relevant information that people have. But the effects are often smaller than we imagine and much smaller than the ones caused by changes in policy-relevant information. Partisans tend to perceive themselves as much less influenced by cues than members of the other party—a finding with troubling implications for those who subscribe to deliberative theories of democracy.

v

Part 2: The widely noted tendency of people to resist challenges to their political beliefs can usually be explained by the poverty of those challenges: they are easily avoided, often ambiguous, and almost always easily dismissed as irrelevant, biased, or uninformed. It is natural to hope that stronger challenges will be more successful. In a trio of experiments that draw on real-world cases of misinformation, I instill false political beliefs and then challenge them in ways that are unambiguous and nearly impossible to avoid or dismiss for the conventional reasons. The success of these challenges proves highly contingent on party identification.

Part 3: Political scientists are increasingly interested in using Bayes' Theorem to evaluate citizens' thinking about politics. But there is widespread uncertainty about why the Theorem should be considered a normative standard for rational information processing and whether models based on it can accommodate ordinary features of political cognition including partisan bias, attitude polarization, and enduring disagreement. I clarify these points with reference to the best-known Bayesian updating model and several little-known but more realistic alternatives. I show that the Theorem is more accommodating than many suppose—but that, precisely because it is so accommodating, it is far from an ideal standard for rational information processing.

Acknowledgments

In his preface to *The Nature and Origins of Mass Opinion*, John Zaller thanks an adviser for teaching him that graduate school is "more about taste and judgment than about facts and theories." I blanched at this when I read it as an undergraduate. It seemed like a recipe, if not an apology, for bad work.

I was slow to see that this was callow of me. But I was wise enough—lucky enough—to choose a dissertation adviser who knew that Zaller was on to something. So far as my tastes have been refined, I owe the most to Paul Sniderman, who spent many hours gently turning me toward some projects and away from others. I found it tough to accept that not every interesting project was worth pursuing, but thanks to Paul, I did learn to accept it. In this matter and in others, his advice has stood me in good stead throughout my time at Stanford.

I took more classes with Simon Jackman than with anyone else. He has always been patient: to teach so much about statistics to people who know so little, he has to be. He is generous, too. Just last week, I asked him if we might meet for ten or fifteen minutes to discuss a paper that I am writing. We met for an hour. That is characteristic: however much he is burdened by other duties, Simon makes time to explain matters that are old hat for him but novel and sometimes difficult for me.

Simon is also—and this is not yet sufficiently recognized—one of our clearest writers. I marked that in my first year at Stanford, when I read his trio of articles about Markov Chain Monte Carlo techniques in political science. Ever since, I have tried to

bring to my own writing the clarity of exposition that he brings to his. I don't think that I've quite succeeded, but you must be the judge of that.

A lot seems to depend on serendipity. During my first year at Stanford, Bob Luskin was a Fellow at the Center for Advanced Study in the Behavioral Sciences. We had never met, but I invited him to attend a political communication class about deliberation. He accepted that invitation and a later invitation to join my dissertation committee. We've worked together since that first meeting, and he has done much to shape my ideas—not just about political sophistication, the subject of our collaborations, but about what constitutes good research and what does not. Bob, too, would agree that Zaller was on to something when he wrote about taste and judgment. He has plenty of both.

Serendipity also takes the credit for my long relationship with Jeffrey Friedman—and thus for my graduate career, for I would not have gone to graduate school if I hadn't met Jeff. He was the teaching assistant for a class I took in my sophomore year of college. But there were other TAs, and I met Jeff only because I overslept on the first day of class, missing the section that I'd meant to attend and going to Jeff's instead. We kept in touch after that, and he often told me that I should go to graduate school. I rarely had the heart to tell him that I hadn't the slightest desire. Then, one day, I did have the desire. And that was because Jeff had made me aware that political behavior was a rich and rewarding subject. The current dissertation may seem to bear little relation to his own work, which has its roots in political philosophy, but the tokens of his influence are here for anyone who sees fit to read these papers alongside *Critical Review*, which he edits now just as he did then.

Early in my studies, I had the good fortune—this was another happy accident—to meet people who took experiments seriously. Alan Gerber and Don Green were the first, and it pleases me immensely to think that I'll soon rejoin them at

Yale. They were followed in my first year at Stanford by Shanto Iyengar, Simon, and Paul. David Brady does not conduct his own experiments, but he gave me a great deal of help when I was preparing mine.

All of these people gave up many hours of their own time to talk with me about my work. It heartens me now to recall that they were not alone. On the contrary: so many people in the discipline who scarcely knew me have been generous with their time. Jamie Druckman and Skip Lupia met with me at all manner of conferences. I took a memorable class on bounded rationality with Jon Bendor in my second year at Stanford and have had many exchanges with him since then; I've learned something during each of them. Jon Krosnick was always a willing sounding board for my ideas, and his ability to refer me to relevant work in social psychology was, and is, nonpareil. Larry Bartels met with me at a Budapest café to help me think through my inchoate ideas about partisan bias and Bayesian updating. After a while—it did not seem long—I realized quite suddenly that I was late for another meeting. I bolted almost without saying a word and absolutely without paying for my drinks. Larry picked up the tab, and he wouldn't let me reimburse him the next day.

I scarcely knew any of these people when I moved to Stanford. But I had Marisa Galvez, my only friend from college who joined me here as a new graduate student. As she fed me—she always fed me, and so well—we talked about how strange it felt to live in California after growing up in the Northeastern cities. It still feels strange, and I will miss it.

I knew Matt Levendusky, too, a little; he endured most of what I did, and much that I didn't, as a fellow American Politics student who entered the Stanford program when I did. Catherine Duggan came to my aid when I was evicted from my apartment in the middle of the night. Though they will rush to deny it, I found kindred spirits in Kimuli Kasara, Nathan Collins, Stephen Jessee, and Neil Malhotra. Laura Miller

ix

helped me work through too many math problems, and I profited almost as much from that help as from her quiet unflagging good cheer.

Like other members of the Stanford department, I've been helped more times than I can recall by the indefatigable office staff. It is not enough to say that they made the trains run on time. They slowed the trains when I needed more time, or sped them up to match my pace, or talked sense to me when I wanted to lay a bomb on the tracks. My debts to Jeanette Lee-Oderman, Angelita Mireles, Rowel Padilla, and Eliana Vasquez are therefore considerable. I dealt most with Eliana and Jeanette, who were always diligent and always kind.

It is bracing to write this and to see how much help I have received from so many quarters. And I have not even thanked my family, who have borne with me longer than any others. Let me single out my brother Will and my father. Will is now a graduate student at Princeton; if I can save him from a few of my mistakes, I'll be doing fine. My father, John C. Bullock, has been devoted to me for thirty years.

I hope that all of the people whom I've mentioned will pardon me as I dedicate this dissertation to Nora Ng. She helped me with every part of it. 我很幸運. Everyone should be so lucky.

John G. Bullock

June 2007

Stanford

x

Contents

xi

List of Tables

xiv

List of Figures

xv

Part I

Party Cues and Policy Information:

Real and Perceived Effects

The standard story about political source cues is easy to tell. They have large and pervasive effects on political attitudes. People use them to infer the consequences of enacting policies and electing candidates (Downs 1957; Popkin 1994). And people respond much more to them than to relevant policy information (Rahn 1993). But *ceteris paribus*, the more policy information that people have, the less they will rely on source cues (Petty and Cacioppo 1986; Mondak 1993b).

How much of the story is true? We cannot say, for our hypotheses about political source cues far outstrip our knowledge of them. Surprisingly, this is especially so of *partisan* source cues. There are dozens of types of source cues (McGuire 1969), but given the central place that party identification assumes in our literature, partisan cues should have a special claim on our attention. It seems that citizens are affected by them (e.g., Rahn 1993; Ansolabehere et al. 2006; Conover and Feldman 1989; Druckman 2001) and that the more knowledgeable are less affected (Kam 2005, though see Lau and Redlawsk 2001). But there is much more that we don't know. How

I

large are the effects? Do they diminish as knowledge of the issue at hand increases? Are people more influenced by party cues than by policy details?

Evidence is slight on all of these matters, and most of it comes from surveys that are unsuited to answering questions about party cues. This article is an attempt to plug the gap. In experiments inspired by real-world health care legislation, I manipulate both party cues and policy information prior to asking subjects' attitudes. The results suggest that party cues affect even the extremely well-informed but that their effects are modest relative to those of policy-specific information. The more that people are exposed to information about a policy, the less their attitudes are influenced by cues. Democrats and Republicans believe that they are much less influenced by cues than members of the other party—a finding with unsettling implications for those who subscribe to deliberative theories of democracy.

Theory

A *cue* is a datum that people may use to infer other information and, by extension, to make decisions.[1] *Party cues* come in two forms. They may reveal a person's party affiliation: "John is a Democrat." Or they may link a party to a stand on an issue: "The Democrats voted for tax cuts." Of course, elite political communication to mass audiences involves more than party cues: in particular, it entails framing issue positions in persuasive ways (e.g., Chong and Druckman 2006). But it is worth examining party cues independent of the frames and affect-laden rhetoric in which they are often couched. That is the point of this article.

---

[1] This definition is superficially dissimilar to the one favored by many social psychologists: "cues refer to stimuli in the persuasion context that can affect attitudes without necessitating processing of the message arguments" (Petty and Cacioppo 1986, 18). In practice—that is, in the discussion of particular experiments—all distinctions between these definitions melt away.

Concerted theorizing about the role of cues in decision-making dates to Herbert Simon's extensive but superficially apolitical treatments (Simon and Newell 1958; Simon 1955, 1959; Newell, Shaw, and Simon 1958). Downs (1957) brought the discussion to politics and paid special attention to party cues. And his work differed from Simon's in a second respect. Simon stressed the role of cues as "cognitive shortcuts"—effort-saving devices that can help us make reasonable decisions without thinking hard. Downs also attended to the use of cues as cognitive shortcuts, but he placed greater emphasis on the way that cues can substitute for other kinds of information.

Simon's discussions prefigure psychological work that elaborates his ideas. This work on "dual-process" models of attitude change distinguishes between what Petty and Cacioppo (1986) call "central" and "peripheral" routes to attitude change. Peripheral-route information processing entails no active thinking about the "central merits" of an issue. Instead, attitude change is fostered by immediately-understood information that is only indirectly related to the issue at hand. Classic examples of such information include the attractiveness and reputed expertise of the source: messages that come from attractive or reputedly expert sources are more likely to be persuasive (Chaiken 1979; Bochner and Insko 1966). Importantly, the extra persuasiveness of such messages has nothing to do with analysis of their content.

Political scientists typically argue that reliance on source cues is another form of peripheral-route processing (Rahn 1993; Mondak 1993a; Kam 2005; but see Kerkhof 1999). They maintain that it permits quick decisionmaking, less thinking about the policy issue or candidate choice at hand, and less attention to other available information that may be harder to interpret. They also maintain that the antithesis of peripheral-route processing is attention to detailed information about political issues. "When voters can expertly judge every detail of every stand taken and relate

it directly to their own views of a great society," Downs tells us, "they are interested only in issues" and do not rely on party cues (Downs 1957, 98). But that is the only condition under which they engage in pure policy-based voting. When voters are not expert about every political matter—i.e., always—Downs suggests that they rely on information distant from the "details of every stand," including party cues. And it is easy to infer a tradeoff from Downs' ideas: the less issue-relevant information a person possesses, the more he will rely on cues.

Normative concern about party cues often rests on the possibility that they lead us to policy attitudes that we would not hold if we were well-informed. A second concern is that cues are associated with a special kind of partisanship: people may see themselves as unaffected by cues but believe that members of the other party are extremely affected by them. This is not the contrast that empirical work on actor-observer differences leads us to expect (Jones and Nisbett 1972), but it is suggested by the finding that people generally deem themselves more objective and thoughtful than those with whom they disagree (Ross and Ward 1996; see also Epley and Dunning 2000). The possibility is especially troubling to those who subscribe to deliberative theories of democracy, because deliberative theories make special demands about citizens' perceptions of each other. Rawls elaborates: citizens' duty is to "explain to one another . . . how the principles and policies they advocate and vote for can be supported by the political values of public reason," which are broadly shared ideals about justice (Rawls 1993, 97). Each citizen must believe that the others are equally willing (or at least moderately willing) to make explanations of this sort (Rawls 1993, esp. 243; Cohen 1989; Knight and Johnson 1997). But if a citizen believes that he is independent-minded while members of the other party are unduly influenced by their party leaders, it may be hard for him to believe that they are his deliberative equals or that their positions are rooted in broadly shared ideals of any sort.

Research Designs for the Study of Party Cues

Most studies of party cues are based on nonexperimental surveys. Typically, respondents are asked where they stand on issues and where political parties stand on the same issues (e.g., Conover and Feldman 1989). Those who answer the questions about parties' stances are deemed to have received cues conveying those stances. And if their answers to those questions are correlated with their issue attitudes, cues are said to affect their attitudes. One difficulty with this approach is that many people express views on issues that they have never heard about (Schuman and Presser 1981, ch. 5); merely answering a question about a party's stand, then, is no indication that one has received a party cue. A second problem is reciprocal causality: there is every reason to expect that people's own issue stances influence their perceptions of parties' stances, in which case those perceptions are a murky amalgam of party cue information and projection effects (Page and Brody 1972; Jessee and Rivers 2007).[2] Even apart from this, the receipt of cues may be confounded with other variables that are responsible for the observed effects. For example, very informed respondents are more likely to receive cues and to take their parties' positions, but it may be their knowledge of policy details, rather than their receipt of cues, that causes them to take those positions. Of course, one can attempt to control for political knowledge and to model the relations between it, receipt of cues, and policy attitudes. But political knowledge is almost always measured crudely with just a handful of items, as are other potential confounds; and uncertainty about the correct form of the model is the rule rather than the exception (Leamer 1978, 1983; Freedman 1991). That brings us to a

---

[2] A few authors acknowledge the problem and try to overcome it with two-stage least squares regressions. But the instruments that they use—typically, party identification, gender, race, and a few other demographic variables—invariably perform poorly, rendering the regressions suspect ((Bound, Jaeger, and Baker 1995); for an example, see Conover and Feldman 1989, 928n12). Party identification poses an additional problem because it is unlikely to satisfy the exclusion restriction.

final problem: most surveys do not include the items that a good observational study of party cues requires. Questions about policy attitudes are often ill-suited to the study of cue effects, questions about cue knowledge are rare, and items measuring potential moderators of cue effects are even more lacking. (This is only a slightly lesser problem in the National Election Studies than in commercial surveys.)

All of these problems can be countered by experiments in which some subjects are randomly assigned to receive party cues. But most experiments have shortcomings of their own. Some are shortcomings of realism: see Appendix A for details. Others bear directly on causal inference. In particular, many experiments are poorly suited to testing hypotheses of interest, which are often not just about the size of cue effects but about their relation to other political information. Do people respond more to party cues than to policy information? Do the effects of party cues increase as the availability of policy information declines? Do people rely more on cues when the alternatives under consideration are similar or when they are different? Answering these questions requires more than the manipulation of cues: it requires the separate but simultaneous manipulation of cues and other aspects of the information environment. This simultaneous manipulation of different informational factors has been done occasionally in work on other kinds of partisan messages. But it has never been done in experiments on party cues, save for the one reported by van Houweling and Sniderman (2004).[3]

The experiments reported here feature this sort of manipulation. In the first two, subjects receive information about a policy debate in the form of a newspaper article. Party cues are manipulated across versions of the article. Characteristics of the policy information provided in the articles are separately manipulated, too. I close by

---

[3] See Rahn (1993) for the closest approximation. Cohen's very good 2003 article is also close, but he manipulates party cues and arguments about policies together rather than separately, making it difficult to isolate the effects of party cues or to study their relation to other variables.

reanalyzing the van Houweling and Sniderman data, showing that findings from the first two experiments generalize across issue areas. They are also robust to changes of the location of policy alternatives in policy space.

Hypotheses

Four hypotheses are at stake. Following Rahn (1993), Cohen (2003) and others, I expect that party cues affect even the extremely well-informed and that they are more influential than information about the provisions and likely consequences of policies. But following Downs (1957), people who have more policy-relevant information should be less influenced by party cues. And following Cohen (2003), I expect that members of each party think themselves uninfluenced by cues but think that members of the opposite party are very influenced by cues.

Experiment I: Simultaneous Manipulation of Party Cues and Policy Direction

Democrats and Republicans were presented with a newspaper article about health care for the poor in Wisconsin.[4] The article contrasted the status quo with a series of changes that had just been passed by the state House of Representatives in a party-line vote.

The experiment had a 2 × 3 factorial design. Subjects were randomly assigned to read about liberal or conservative changes to the status quo. They were also assigned to one of three cue conditions: "intuitive," in which Democratic legislators supported

---

[4] The article was adapted from one about Medicaid cuts that were passed by Missouri's legislature in 2005 (Lieb 2005). See Appendix B.

the liberal changes or opposed the conservative ones; "counterintuitive," in which Democratic legislators supported the conservative changes or opposed the liberal ones; or a third condition in which no cues were provided. In each of the first two cue conditions, Republican legislators opposed their Democratic counterparts.

Participants, Design, and Procedure

Seven hundred and sixty subjects, all identifying with either the Democratic or Republican party, were recruited from two national participant pools—one maintained by Survey Sampling International, the other by a large private university—to participate in a study about "news media in different states." 50% were Democrats, 51% were male, and 48% had graduated from a four-year college. Their median age was 39. All were presented with a newspaper article and asked to read it carefully, "as most of the questions that follow will be about your reactions to it."

Policy information. The status quo was held constant across all versions of the article. It allowed Medicaid recipients one eye care exam every two years, required co-payments of 50 cents to three dollars for every visit to a doctor, and offered some coverage for wheelchairs, artificial limbs, and children not covered by Medicaid. Single parents of two were eligible for coverage if they earned less than $1,334 per month. Disabled adults aged 18 to 64 qualified for coverage if they earned less than $1,940 per month. In the past twelve years, Medicaid costs had tripled, and at the time of the article, Medicaid accounted for 29% of Wisconsin's budget.

The status quo was contrasted with changes that would either restrict or expand health care for the poor. Liberal changes would increase coverage for 100,000 of the state's one million Medicaid recipients, chiefly by loosening eligibility standards. Conservative changes would reduce coverage for the same number of people by tightening eligibility standards. (See Table B1.)

	liberal changes	conservative changes
no party cues	some support changes; others oppose them	some support changes; others oppose them
intuitive party cues	Democrats support changes; Republicans oppose them	Republicans support changes; Democrats oppose them
counterintuitive party cues	Republicans support changes; Democrats oppose them	Democrats support changes; Republicans oppose them

Table 1.1: Design of Experiment 1. Experiment 1 had a 2 × 3 design. Each subject read about proposed changes, either liberal or conservative, to a state's Medicaid policy. In the "intuitive party cues" condition, Democratic legislators supported the liberal changes or opposed the conservative ones while Republican legislators did the opposite. In the "counterintuitive party cues" condition, Republican legislators supported the liberal changes or opposed the conservative ones while Democratic legislators did the opposite. In the "no party cues" condition, subjects read about support for and opposition to the proposed changes, but the positions were not linked to political parties.

Party cues. In every article, the proposed changes were said to have passed the House by an 87-71 vote, with 90% of one party's legislators voting for the changes and 90% of the other party's legislators opposing them. In the intuitive cue condition, legislators conformed to party reputations: most Democratic legislators supported the changes if they were liberal or opposed them if they were conservative, while most Republicans did the opposite. In the counterintuitive cue condition, legislators defied their parties' reputations. For example, counterintuitive cues in an article about conservative changes indicated that most Democratic legislators supported the changes and that most Republican legislators opposed them. (See Table 1.1.)

Policy arguments. Unlike policy information and party cues, the content of arguments about the proposed changes was not manipulated. Opponents of the liberal changes—whether Republicans (in the intuitive condition) or Democratic (in the counterintuitive condition)—argued that they would make other welfare services unsustainable and lead to reduced school funding, a budget deficit, and higher taxes.

Proponents emphasized the need to protect the disabled, the elderly, and parents who lacked coverage. The governor, a proponent, emphasized that the bill's anti-fraud provisions ensured that all new spending would be directed to the state's neediest residents.

When the changes were conservative, policy arguments were reversed. Opponents argued that the changes would threaten the disabled, the elderly, and parents who would lose coverage. Proponents emphasized that the cuts would allow a balanced budget while increasing school funding and not raising taxes or cutting other welfare services. The governor, a proponent in this condition as well, again touted the bill's anti-fraud provisions.

Post-manipulation measures. After reading the article, subjects were asked to report their mean attitude toward the policy changes on a scale ranging from 1 ("disapprove strongly") to 7 ("approve strongly"). They were then asked to rate the effect that cues had or would have had on them and members of the opposite party. Subjects who received cues were asked "How much is your opinion influenced by the positions of Republican and Democratic legislators?" and "how much would [Democrats | Republicans] be influenced by the positions of Republican and Democratic legislators?" Those who did not receive cues were asked how they would have been affected by intuitive and counterintuitive cues—

> Suppose you learn something new about the policy changes: they are supported by 90% of Democratic legislators and opposed by 90% of Republican legislators. How much would you be influenced by this new information?

> Suppose instead that the policy changes are supported by 90% of *Republican* legislators and opposed by 90% of *Democratic* legislators. How much would you be influenced by *this* information?

—and how members of the other party would have been affected by intuitive cues, e.g.,

Suppose that Democrats who read the news article learn that the

policy changes are supported by 90% of Democratic legislators and

opposed by 90% of Republican legislators. On average, how much

would they be influenced by this new information?

All of these ratings were made on a scale ranging from 1 ("not at all influenced")

to 5 ("extremely influenced"). At the end of the experiment, party identification

was measured on a seven-point scale from 1 ("strong Democrat") to 7 ("strong

Republican").

Randomization checks. The success of the randomization to cue conditions

was gauged by testing it against the subjects' self-reported party identification.

Using a chi-square test, the null hypothesis of independence cannot be rejected

$(\chi_4^2 = 1.5, p = .82)$. For the randomization to policy change conditions (liberal

or conservative), $\chi_2^2 = .47, p = .79$. To ensure that the two randomizations were

independent of each other, they were tested against each other: $\chi_2^2 = 1.7, p = .44$.

Results

Consider the first hypothesis: party cues affect even the very well-informed. By

the standards of political science, all subjects in Experiment 1 were exposed to an

extraordinary amount of information about a policy debate. Figure 1.1, which reports

the main results, shows that party cues did affect attitudes in many cases—and, by

extension, that they affected even the very well-informed.

Unexpectedly, Figure 1.1 also suggests that counterintuitive cues are more

effective than intuitive ones. Democrats are most supportive of liberal policy changes

when Democratic legislators support them (mean attitude = 4.99) and least supportive

when Democratic legislators oppose them ($M = 4.47$); the difference is significant
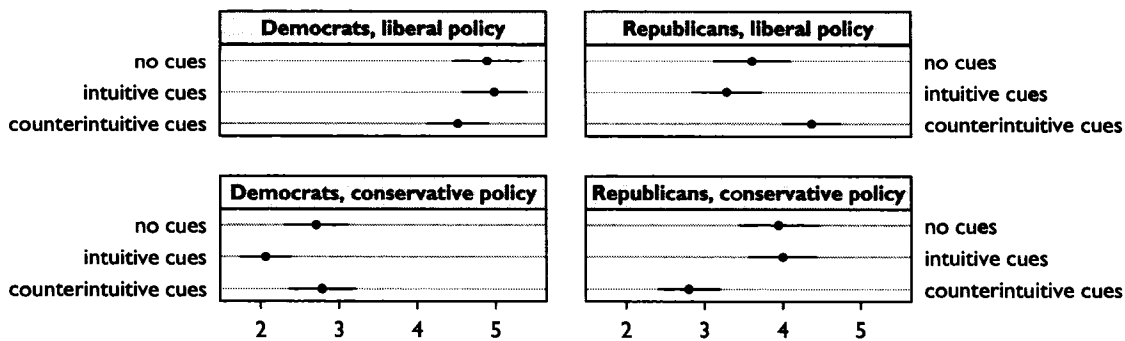
Figure 1.1: Effects of Cues and Policy Information in Experiment 1. All panels plot mean attitude toward the proposed policy changes. Responses range from 1 ("disapprove strongly") to 7 ("approve strongly"). Black bars represent 95% confidence intervals.

Two general results are apparent. Counterintuitive cues change attitudes relative to a no-cue condition; intuitive cues do not. And changes in policy information are almost always more influential than changes in cue information.

at $p = .04$.[5] In this case—establishing a trend—the counterintuitive condition, in which Democratic legislators oppose the liberal changes, depresses Democratic subjects' attitudes toward the changes relative to the no-cue condition (for which $M = 4.90, p = .08$). But the intuitive-cue condition scarcely differs at all from the no-cue condition ($p = .38$)—again, establishing a trend. Note, though, that the difference between the magnitudes of the two cue effects is statistically insignificant ($p = .42$, two-tailed).[6]

The same patterns emerge—in the opposite directions, of course—for Republicans reading about liberal changes. They are most likely to oppose the changes when Republican legislators oppose them ($M = 3.31$), most likely to support them when Republican legislators support them ($M = 4.37$). Again, this difference is significant ($p < .001$), as is the difference between the counterintuitive-cue condition

---

[5] Because there are clear directional hypotheses about the directions of cue effects, significance tests reported in this article are one-tailed unless otherwise noted.

[6] This is a difference of differences: (|attitude in counterintuitive cue condition| − |attitude in no-cue condition|) − (|attitude in intuitive cue condition| − |attitude in no-cue condition|).

and the no-cue condition (for which $M = 3.61$, $p < .01$). And again, the difference between the intuitive-cue and no-cue conditions does not reach conventional standards of significance ($p = .16$). The difference between the magnitudes of the cue effects is again statistically insignificant: $p = .29$, two-tailed.

When subjects read about potential conservative policy changes, the patterns hold save for one exception. For both Democratic and Republican subjects, policy attitudes differ between the intuitive and the counterintuitive cue conditions ($p < .01$ in both cases). For Democrats, the regular pattern is reversed: when they read about a conservative policy alternative, it is the intuitive cues, rather than the counterintuitive ones, that cause a significant change from the no-cue condition. But when we look at Republican responses, we see a return to form. Their attitudes in the counterintuitive cue condition, but not those in the intuitive cue condition, differ significantly from their attitudes in the no-cue condition. It appears, then, that counterintuitive party cues generally affect even the well-informed. But how much?

By conventional standards, not much. The largest observed mean attitude difference due to cues is the one between Republicans reading about conservative policy changes in the intuitive condition ($M = 3.96$) and the counterintuitive condition ($M = 2.76$). This is a shift of 20% on the 1-7 attitude scale—sizable but not extraordinary. And the average differences are much smaller. In absolute value, the average mean shift caused by intuitive cues relative to no cues was .32, or 5.3% of the scale. For counterintuitive cues relative to no cues, it was .59, or 9.8% of the scale. These effects are smaller than many others that we observe in mass behavior—smaller even than effects that are elsewhere treated as "small" (e.g., in the literature on issue framing: Druckman and Nelson 2003; Sniderman and Theriault 2004). And in this experiment, they are swamped by the effects of changes in policy direction (which

no cues (intuitive counterfactual)				
no cues (counterintutive counterfactual)				
intuitive cues				
counterintuitive cues				

2　　2.5　　3　　3.5　　4

Mean Perceived Influence of Cues
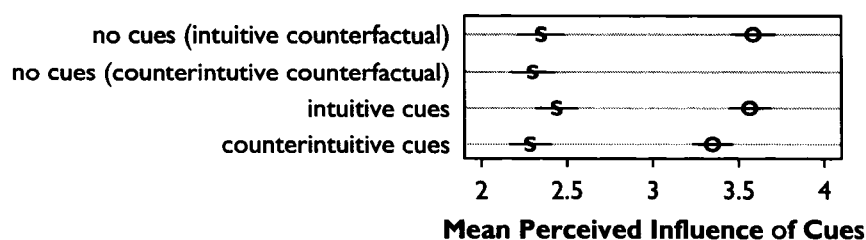
Figure 1.2: People Think Members of the Other Party Are Much More Affected By Cues. Each row plots the mean perceived influence of cues for oneself (S) and members of the other major political party (O). Black bars represent 95% confidence intervals.

All data are from Experiment 1. Subjects who received cues were asked how much they were influenced by the cues and how much members of the other party would be influenced. Subjects who did not receive cues were asked how much they and other-party members would be influenced by intuitive cues. They were also asked how much they would be affected by counterintuitive cues. In every case, responses ranged from 1 ("not at all influenced") to 5 ("extremely influenced").

caused a mean change of 1.58, covering 26.3% of the scale) and party ID (1.05, 17.5%).

Turn now from the real effects of party cues to their perceived effects. Figure 1.2 shows that most people hold an accurate impression of the modest extent to which their judgment is affected by cues. But they radically overestimate the extent to which members of the other party are affected. Overall, only 14% of Democratic and Republican subjects estimated that they were or would be "very" or "extremely" influenced by cues. But 54% believed that members of the other major party were or would be so influenced. Of course, the difference may be due to a belief that others, regardless of party, are more affected by cues than oneself; or it may be due to beliefs about partisan differences. In either case, the finding is the same, and so is its importance: members of both parties think themselves relatively immune to elite influence while seeing members of the other party as highly susceptible. This does not bode well for any form of political activity in which members of different parties are called on to interact with each other. Note too that the difference is especially striking

because the questions about cue influence on oneself and on other-party members appeared consecutively in the questionnaire, creating a clear contrast. If they had not, the differences between the responses might have been greater still (Schuman and Presser 1981, 28-30; Hovland and Sherif 1957).

The hypothesis about perceived cue effects is therefore supported by the data, but the hypotheses about actual cue effects generally are not. Party cues did affect even the extremely well-informed subjects of Experiment 1, but the effects were small and far outweighed by changes in policy information. Note, though, that these are findings from just one experiment. And they do not show all that one might like. The instruction to read the article carefully may have artificially heightened attention to policy information, thereby diminishing cue effects. And by political standards, the article contained an extraordinary amount of information. Cues may have larger effects when subjects are exposed to less information. Experiment 2 speaks to these concerns.

Experiment 2: Simultaneous Manipulation of Party Cues and Amount of Policy Information

Subjects, all Democrats, were presented with a newspaper article about health care for the poor in Wisconsin. It contrasted the status quo with a series of liberal changes that had just been passed by the state House of Representatives in a party-line vote. Subjects were randomly assigned to read high-, medium-, or low-information versions of the article. They were also assigned to receive either counterintuitive party cues or no party cues.

Participants, Design, and Procedure

Four hundred and one subjects were recruited from two national participant pools—one maintained by Survey Sampling International, the other by a large private university—to participate in a study about "news media in different states." All had previously self-identified as Democrats in unrelated studies. 53% were male, 52% had graduated from a four-year college, and the median age was 43. All were presented with a newspaper article, but they were not asked to read it carefully, as they had been in Experiment 1.

Amount of policy information. The high-information versions of the article contained nearly all of the policy information that appeared in Experiment 1 articles about liberal changes. The medium-information versions omitted discussion of copayments, disability cutoffs, and coverage for children, wheelchairs, prostheses, and eye care. The low-information version further omitted information about dental care, eligibility standards, the costs of the changes, and the number of state residents who stood to benefit from them (100,000). It only noted that the legislation would expand coverage for "tens of thousands" of low-income residents and that it contained provisions to guard against fraud and waste.

Party cues. Subjects were randomly assigned to receive either no cues or counterintuitive cues, which took the same form that they did in Experiment 1.

Policy arguments. Policy arguments took the forms that they did in Experiment 1.

Post-manipulation measures. Experiment 2 used the measures that were used in Experiment 1.

Randomization checks. The success of the randomization to cue conditions was gauged by testing it against the subjects' level of education, which was measured on an eight-category scale. Using a chi-square test, the null hypothesis

Mean Attitude Rating
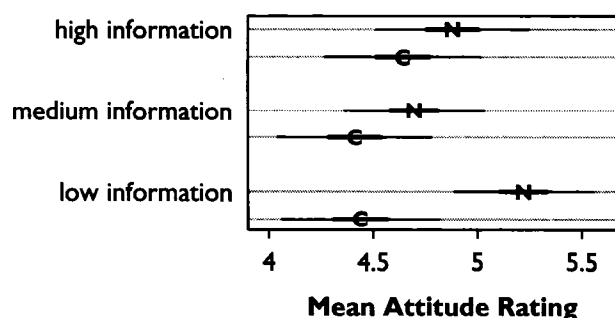
Figure 1.3: Counterintuitive Cues Grow Powerful As Policy-Relevant Information Declines. Each row plots mean attitude toward the proposed policy changes in the no-cue condition (**N**) or the counterintuitive cue condition (**C**). Black bars represent 50% and 95% confidence intervals. Individual attitude responses ranged from 1 ("disapprove strongly") to 7 ("approve strongly").

As the amount of information provided to subjects declines from high to low, the effect of the counterintuitive cue—the distance between **N** and **C** within a tier—more than quadruples.

of independence cannot be rejected ($\chi^2_7 = 3.8, p = .80$). For random assignment to the level-of-information condition, $\chi^2_{14} = 16, p = .30$. To ensure that the two randomizations were independent of each other, they were tested against each other:

$\chi^2_2 = .22, p = .90$.

Results

Consider the likely pattern of responses. If cue effects are stronger when less information is available, they should be strongest for those in the low-information condition, weakest for those in the highest. This is the pattern that appears in Figure 1.3. The mean attitude rating of subjects in the low-information condition is 4.44 if they received counterintuitive cues, 5.22 if they received no cues at all ($p = .001$). In the medium-information condition, the gap is barely one-third as big: $M = 4.41$ in the counterintuitive condition, 4.70 in the uncued condition ($p = .13$). It closes further still in the high-information condition: $M = 4.64$ for those who received cues, 4.88 for those who did not ($p = .18$).

medium-information condition	−.25	.30
low-information condition	.40	.31
counterintuitive cues	−.26	.31
medium-information condition × counterintuitive cues	−.06	.43
low-information condition × counterintuitive cues	−.65*	.44
$\tau_{\text{disapprove strongly | disapprove somewhat}}$	−4.60	.49
$\tau_{\text{disapprove somewhat | disapprove slightly}}$	−2.67	.27
$\tau_{\text{disapprove slightly | neither approve nor disapprove}}$	−1.48	.23
$\tau_{\text{neither approve nor disapprove | approve slightly}}$	−.51	.22
$\tau_{\text{approve slightly | approve somewhat}}$	.48	.22
$\tau_{\text{approve somewhat | approve strongly}}$	1.86	.25
Log likelihood	−696.94	
Likelihood ratio test	13.48	.02
Number of observations	399	

Table 1.2: Exposure to Policy-Relevant Information Moderates Cue Effects in Experiment 2. Cell entries in the top rows are ordered logistic regression parameter estimates and standard errors. Entries in the "likelihood ratio test" row are χ^2 statistics and corresponding p-levels from tests against a model with no predictors. The dependent variable is attitude toward the proposed policy changes, which ranges from 1 ("strongly disapprove") to 7 ("strongly approve"). All predictors are dichotomous. The baseline conditions are "high-information condition" and "no cues." Interesting estimates significant at 95% using a one-tailed test for $H_A > 0$ are denoted by *. Two of the 401 subjects in the experiment did not report their attitude toward the policy changes; they were omitted from this analysis.

Table 1.2 verifies that cue influence is moderated by the amount of policy-relevant information provided to subjects. It reports an ordered logistic regression model in which attitude toward the policy changes is a function of cue condition, information condition, and the interaction between cues and information. If the influence of cues increases as exposure to policy information declines, the estimated coefficient for *low information* × *counterintuitive cues* should differ significantly from zero. And it does: $\beta = -.65, p = .07$.

But note again that the cue effects are small. Even in the low-information condition, they change the mean policy attitude by .78, or 13% of the range of the seven-point scale. In the medium-information condition, the effect drops to .29, or five percent of the range of the attitude scale. And in the high-information condition, it drops to .24, or four percent of the scale's range.

Figure 1.4 provides the evidence on perceptions of cue influence. As in Experiment 1, there is a stark difference between the perceived influence of cues on oneself and on others. Pooling over the amount-of-information conditions, 13% of subjects who received the counterintuitive cues thought that they were "very" or "extremely" influenced; 27% of those who did not receive cues thought that they would be so influenced, as did 30% when asked how they would be affected by intuitive cues. But fully 49% of those who received cues thought that Republicans would be "very" or "extremely" influenced, and 57% of those who did not receive cues believed that Republicans would be so influenced by intuitive cues.

Experiment 2 thus buttresses the findings from Experiment 1. Party cues have weak effects on attitudes. The strength of the effect depends, as Downs theorized, on the amount of policy information to which one is exposed. And partisans believe themselves far less affected by cues than members of the other party. Still, these are two experiments, designed and directed by the same author. And they do not show all that one might like. In particular, party cues may matter much more to judgments made with *very* little information—even less than the amount provided by the low-information condition in this experiment. They may also matter more when the policy alternatives under consideration are more alike. Experiment 3 speaks to these possibilities.

Mean Perceived Influence of Cues
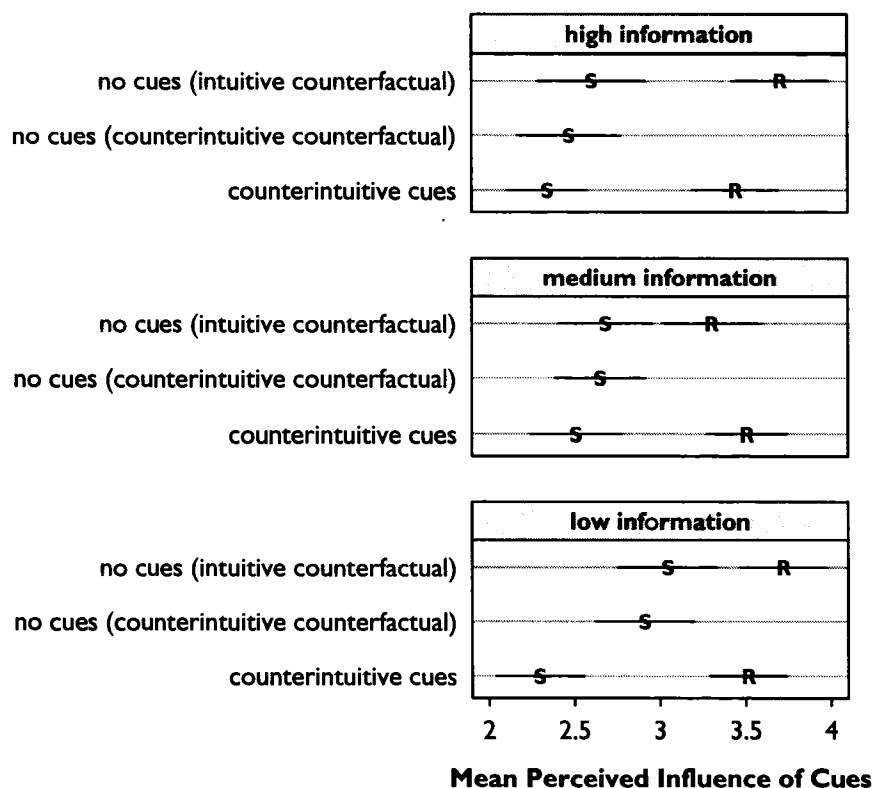
Figure 1.4: Cues Make Democrats Believe that Republicans Don't Think for Themselves. Each row plots the mean perceived influence of cues for oneself (**S**) and members of the Republican Party (**R**). Black bars represent 95% confidence intervals. All data are from Experiment 2. Subjects who did not receive cues were asked how much they and Republicans would be influenced by counterintuitive cues; this produced the "counterintuitive counterfactual" data. They were also asked how much they would be influenced by intuitive cues; this produced the "intuitive counterfactual" data. In every case, responses ranged from 1 ("not at all influenced") to 5 ("extremely influenced").

Experiment 3: Simultaneous Manipulation of Party Cues and Candidate Policy Locations

Experiments 1 and 2 may have produced small party cue effects for any of several reasons. Perhaps even those subjects in the "low-information" condition of Experiment 2 had too much information to permit cues to matter much. Perhaps there is something special about spending on social services that limits cue effects: for example, people

may have stronger prior beliefs about it than about other issues. And perhaps the previous experiments produced relatively small cue effects because subjects were presented with very different alternatives. Had the alternatives been more alike, cue effects might have been larger. Experiment 3, first analyzed in van Houweling and Sniderman (2004), permits examination of all of these possibilities.

Participants, Design, and Procedure

Nine thousand three hundred and thirteen subjects were recruited from a nationally representative pool maintained by Knowledge Networks. Seven thousand five hundred and thirty-three (81%) completed the survey. 45% of the completers identified with the Democratic Party, 37% identified with the Republican Party, 50% were male, 59% were age 45 or older, and 30% had graduated from college.

Prior issue positions. Subjects were asked to indicate their preferences over spending on social services by placing themselves on a 1-7 scale, where 1 stood for "fewer government services even in areas such as health and education to reduce spending" and 7 stood for "important for the government to provide many more services even if it means an increase in spending." They were asked to do the same for preferences over "protecting the environment versus jobs," where 1 stood for "It is important to protect the environment even if it costs some jobs or otherwise reduces our standard of living" and 7 stood for "Protecting the environment is not as important as maintaining jobs and our standard of living"; and for "government aid to blacks," where 1 stood for "Government in Washington should make every effort to improve social and economic position of blacks" and 7 stood for "Government should not make any special effort to help blacks because they should help themselves."

Party cues, policy information, and candidate preferences. At the end of the experiment, the same issue scales were presented to each subject. For each issue,

subjects were told the positions of two candidates, which were determined through random assignment of the candidates to integer positions on the scale subject to the constraint that both candidates could not occupy the same position. Subjects were also told that they were considering a different set of candidates for each issue.

Two thirds of subjects were assigned to always receive party labels for both candidates ("Democrat" and "Republican"); the other third was assigned to never receive party labels, in which case the Democrat was always labeled "Candidate A" and the Republican "Candidate B." The candidates' positions and their labels were presented simultaneously, after which all subjects were asked "which candidate better represents your position on the issue?" The response options were "Candidate A" or "Democrat", "Candidate B" or "Republican," "neither candidate," and "can't say or don't know." In the analyses that appear below, subjects saying "can't say or don't know"—between three and eight percent of the sample, depending on the question—are treated as though they said "neither candidate."

Results

Subjects in this experiment received almost no information about the candidates' policy positions. Figure 1.5 shows that party cue effects are larger under these conditions than they were in previous experiments—though not by much. Among Democrats, the average effect of cues was to boost support for the Democratic candidate by 22% when spending on social services were considered, 13% when the environment was at issue, and 7% when aid to blacks was at issue. Among Republicans, cues depressed support for the Democratic candidate by 17%, 15%, and 7%. Note that spending on social services, the issue that yields the largest cue effects, is the issue most similar to the one considered in the previous experiments.

Proportions Preferring Democratic Candidates
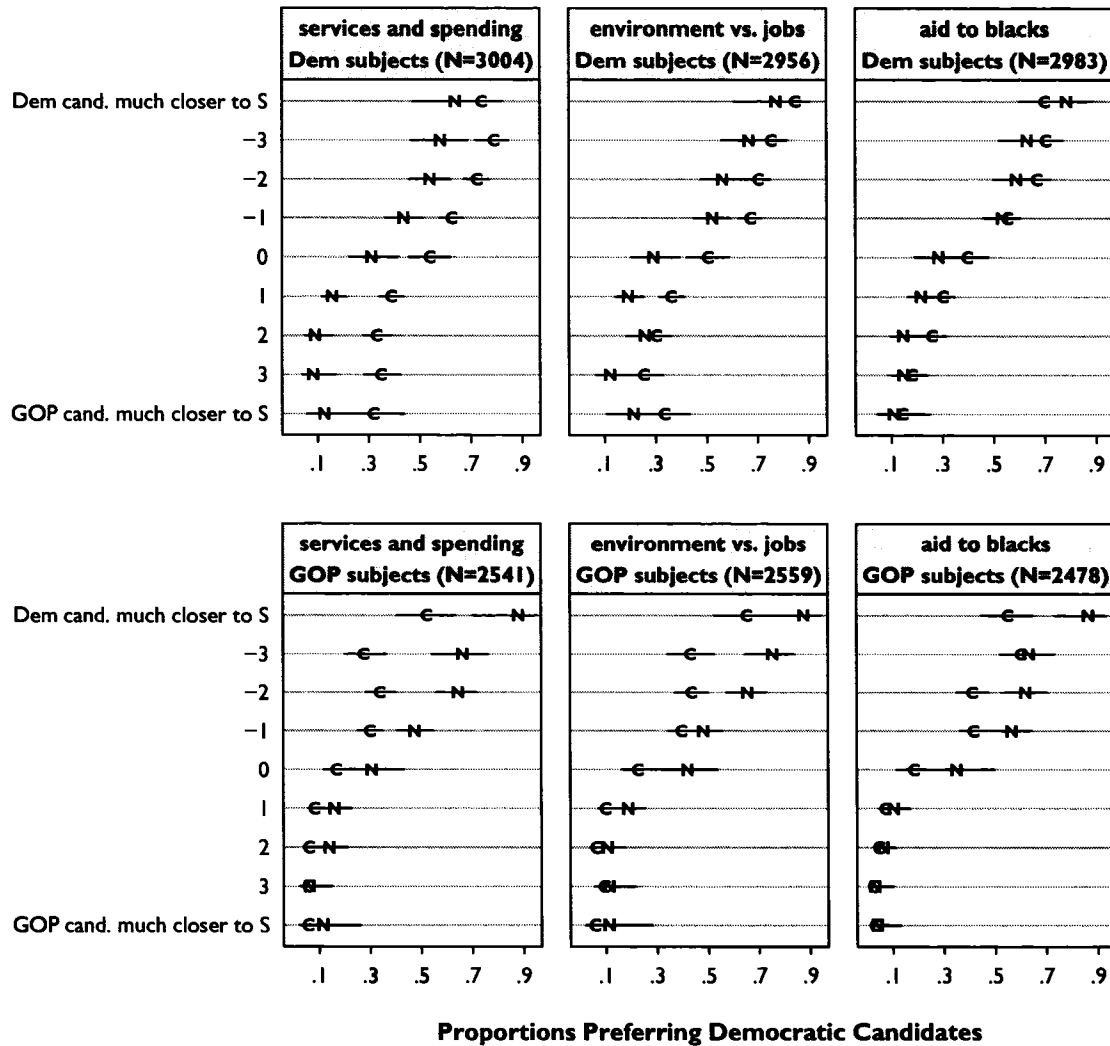
Figure 1.5: Party Cue Effects on Candidate Preference by Issue and Proximity.
Each row of each panel plots proportions of subjects (S) preferring the Democratic candidate's position. **N** indicates the proportion in the no-cue condition, in which candidates were identified only as "Candidate A" and "Candidate B." **C** indicates the proportion in the cued condition. Black lines are 95% confidence intervals.

Three empirical regularities emerge. Party cue effects are usually larger—slightly—in these extremely-low-information experiments than in the previous experiments. Subjects are less influenced by party cues than by the meager information that they have about the candidates' issue positions. And cue effects vary little with the distances of candidates from subjects' ideal points.

This suggests that the effects observed in the previous experiments would have been even smaller had subjects in those experiments considered different issues.[7]

Figure 1.5 also reveals that subjects are again much more responsive to policy information than to party cues. In this experiment, the only policy information that subjects receive is about the locations of the candidates on the seven-point issue spectrum. Changing those locations makes a massive difference to subjects' preferences over the candidates, whether or not they receive cues. To see this, consider any panel in the figure. The bottom row of the panel shows that the Democratic candidate receives very little support in either the cued or the uncued condition when he is much further than the Republican candidate from subjects' ideal points. The rows above show that he gradually gains support as he moves closer to subjects' ideal points and as the Republican candidate moves away from them. And the top row of the panel shows that he is preferred by a large majority when he is much closer than the Republican to subjects' ideal points. The almost complete shift in preferences from the bottom row to the top is evidence of strong responsiveness to policy information.

Even smaller changes in policy information cause preference reversals much greater than the ones effected by party cues. This is easily seen when we define the *simple proximity difference* for each subject on a given issue,

$$P_i = |S_i - D| - |S_i - R|,$$

where i indexes subjects, S_i is subject i's ideal point on the issue in question, and D and R are the positions staked out by the Democratic and Republican candidates (in either the cued or uncued conditions). This quantity is the difference of the

---

[7] The same patterns appear when we consider the effect of party cues on support for the Republican candidate, as seen in Figure C1. This is unsurprising, as levels of support for either candidate are statistically dependent on each other: for any issue, (proportion supporting the Democratic candidate) + (proportion supporting the Republican candidate) + (proportion with no preference) = 1.

Euclidean distances between the subject's ideal point and each candidate's position. Ranging from −6 to 6, it is negative when the Democratic candidate is closer to a subject's ideal point, positive when the Republican candidate is closer, and zero when the candidates are equidistant from the subject.[8] Responsiveness to the proximity difference is responsiveness to policy information; indeed, the proximity difference neatly summarizes the only policy information that subjects have. Among subjects for whom the proximity difference was −3—i.e., those for whom the Democrat was somewhat closer than the Republican—56% preferred the Democrat when considering spending on social services, 62% when considering the environment, and 63% when considering aid to blacks. Among subjects for whom the proximity difference was +3, the proportions preferring the Democrat were 15%, 15%, and 10%, respectively. Thus, merely traversing the middle half of the proximity difference scale changed preferences by 41%, 47%, and 53%—differences two to two and a half times as large as the largest that are caused by party cues in these experiments.

These findings are affirmed by the ordered logistic regression models reported in Table 1.3. The models treat candidate preference as a function of receipt of cues, the proximity difference, and their interaction; they are subsetted by issue and party ID.[9] In every case, the parameter estimates show that subjects are far more responsive to the proximity difference than to party cues. On the logit scale, the greatest possible change in the proximity difference has an effect more than three times greater than that of cues

---

[8] Proximity differences greater than 4 or lower than −4 are extremely rare among uncued subjects. On the social services scale, for example, P_i = −6 for four Democrats and one Republican, −5 for 19 Democrats and eight Republicans, 5 for 12 Democrats and five Republicans, and 6 for three Democrats and four Republicans. In short, there are not enough data to make meaningful inferences about subjects' preferences over candidates at these levels of the proximity difference. Consequently, Figure 1.5 only displays data for subjects for whom $|P_i| \leq 4$.

[9] Following Achen (2003, esp. 446-47; also Achen 2005), the models in Table 1.3 are subsetted to ease comparisons of cue effects to proximity effects. But the finding in each of them—that proximity effects far outweigh cue effects—holds up perfectly in the models of Table C1, which pools subjects from both parties and uses a seven-point measure of party ID.

	social services Dem subjects		environment Dem subjects		aid to blacks Dem subjects	
cues	$-.51^*$	.25	$-.58^*$	.24	$-.08$	.24
proximity difference	5.37^*	.40	4.85^*	.38	5.48^*	.38
cues × proximity difference	$-.91^*$	.48	$-.12$	.46	$-.49$	.45
τprefer A or Democrat | prefer neither	1.80	.21	1.86	.20	2.10	.20
τprefer neither | prefer B or Republican	3.14	.21	2.90	.20	3.45	.21
Log likelihood	-2879.73		-2683.74		-2780.72	
Likelihood ratio test	635.27		923.83		934.54	
Number of observations	3004		2868		2899	

	social services GOP subjects		environment GOP subjects		aid to blacks GOP subjects	
cues	1.65^*	.28	1.44^*	.28	1.30^*	.26
proximity difference	6.39^*	.44	6.96^*	.47	7.57^*	.43
cues × proximity difference	-1.38^*	.54	-1.54^*	.56	-1.55^*	.50
τprefer A or Democrat | prefer neither	2.51	.23	2.91	.24	2.97	.23
τprefer neither | prefer B or Republican	3.89	.24	4.01	.24	4.41	.24
Log likelihood	-2349.79		-2233.82		-2137.34	
Likelihood ratio test	614.49		819.43		1112.99	
Number of observations	2541		2478		2493	

Table 1.3: Cue Effects in Experiment 3. Cell entries in the top rows of each tier are parameter estimates and standard errors from ordered logistic regressions. Entries in the "likelihood ratio test" rows are χ^2 statistics from tests against a model with no predictors; all are significant at $p < .001$. The dependent variable is the answer to "which candidate better represents your position on this issue?" Larger parameter estimates indicate a greater tendency to pick the Republican candidate, who is identified as "Candidate B" in the uncued condition. In the "cues" and "proximity difference" rows, interesting estimates significant at 95% using a one-tailed test for $H_A > 0$ are denoted by *; at 90%, by $^+$. In the "cues × proximity difference" row, the tests are two-tailed.

"Cues" indicates whether the subject was told the candidates' party affiliations: subjects who did are coded as 1; those who didn't are coded as 0. "Proximity difference" is the variable defined in the text but rescaled to range from 0 to 1. Values below .5 indicate that the Democratic candidate was closer to the subject's ideal point. Values above .5 indicate that the Republican was closer. A value of .5 indicates that the candidates were equidistant from the subject's ideal point. Cues have significant effects in almost every case, but their effects are always swamped by the effects of the proximity difference.

Change in Support for Democratic Candidate
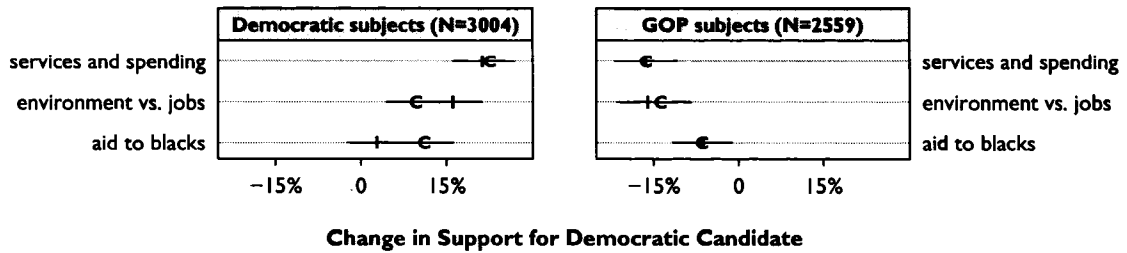
Figure 1.6: Cue Effects Do Not Depend on Which Candidate Is More Liberal. Each row of each panel plots the difference between the percentage of subjects preferring the Democratic candidate in an "intuitive" or "counterintuitive" cued condition and the percentage preferring him in the corresponding uncued conditions. I indicates the net change in preference when the Democrat takes a more liberal position than the Republican. C indicates the net change in preference when the Democrat takes a more conservative position. Black bars represent 95% confidence intervals. I and C are close in each row, indicating that cue effects do not depend much on whether the Democrat is to the left or right of the Republican. Where I and C overlap almost perfectly—as for Republican subjects considering services vs. spending or aid to blacks—it makes almost no difference at all whether the Democrat is to the left or right of the Republican.

among cued Democrats when social services are at issue.[10] And in every other case, the proximity effect outstrips the party cue effect by a greater distance. In the most extreme case—among Republicans considering the environment—the proximity effect is 70 times greater than the cue effect.[11]

The data also permit us to revisit the question of whether cues are more powerful when candidate positions are counterintuitive, and they let us distinguish between two meanings of "counterintuitive." The term can mean that the Democratic candidate stakes out a more conservative position than the Republican, or it can mean that the candidate from one's own party is more distant from one's ideal point than

---

[10] This is the effect of a full swing in the proximity difference among cued Democrats, divided by the effect of cues among Democrats who are much closer to the Republican candidate: $(5.37 - .91)/(-.51 - .91) = -3.14$.

[11] This is the effect of a full swing in the proximity difference among uncued Republicans, divided by the effect of cues among Republicans who are much closer to the Republican candidate: $6.96/(1.44 - 1.54) = -69.60$.

the candidate from the other party. Whether candidate positions are counterintuitive in the first sense makes little difference to cue effects. It matters most to Democratic subjects considering candidates who stake out positions on aid to blacks: when the Democratic candidate is more liberal than his opponent on this issue, he gets 3% more support from cued Democratic subjects than from their uncued counterparts; but when he is more conservative than his opponent, the cue effect rises to 11%. The difference is significant at $p = .02$, two-tailed. The only other statistically significant difference is in the opposite direction: when the Democratic candidate is more liberal than his opponent on the environment, he gets 16% more support from cued Democratic subjects than from their uncued counterparts; but when he is more conservative than his opponent, the cue effect drops to 10%. The difference is significant at $p = .08$, two-tailed. As Figure 1.6 shows, the difference for Democrats considering social services is under 3%, as are all the differences for Republicans. None of these approach statistical significance.

The second sense of "counterintuitive" matters much more. When the Democratic candidate is closer than the Republican to Democratic subjects' ideal points, cues boost his support among them by 18% on social services, 13% on the environment, and 4% on aid to blacks. But when the Democratic candidate is further from Democratic subjects' ideal points, the cue effects are 24%, 12%, and 7%. Among Republicans, the greater effect of counterintuitive cues is starker. When the Republican candidate is closer to Republican subjects' ideal points, the cue effects are 7%, 2%, and 2%. But when he is further from them, the cue effects are 26%, 19%, and 16%. Figure 1.7 displays these differences, all of which are significant at $p = .002$ or below (two-tailed) save for the difference among Democrats considering the environment ($p = .80$, two-tailed) and aid to blacks ($p = .07$, two-tailed).

It is reasonable to expect party cues to matter more when the proximity difference between the candidates is small, and to matter most when the proximity

Change in Support for Democratic Candidate
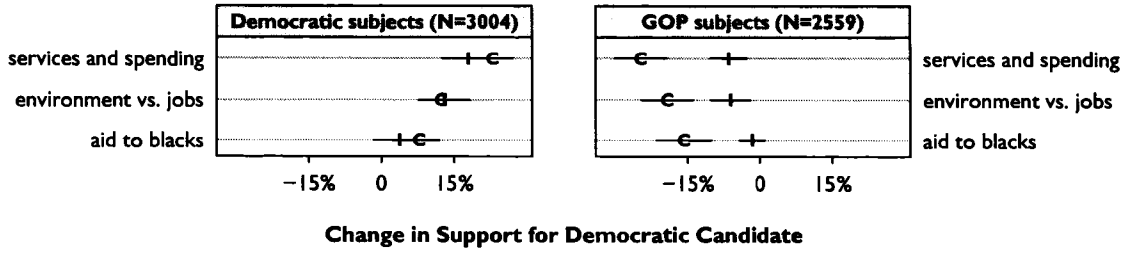
Figure 1.7: Cue Effects Depend on Whether the Candidate From One's Own Party Is Closer to One's Ideal Point. Each row of each panel plots the difference between the percentage of subjects preferring the Democratic candidate in a cued condition and the percentage preferring him in the corresponding uncued condition. **I** indicates the net change in preference when the candidate of subjects' own parties is closer to their ideal points. **C** indicates the net change when the candidate from the other party is closer to subjects' ideal points. Black bars represent 95% confidence intervals. All of the within-row differences-in-difference are significant at $p < .10$ (two-tailed) save the difference between **I** and **C** among Democrats considering the environment, where $p = .80$ (two-tailed). For example, among Republicans considering aid to blacks who have ideal points closer to the Republican candidate, party cues decrease support for the Democratic candidate by 2%. But among Republicans who are closer to the Democrat, party cues decrease support for the Democrat by 16%.

difference is both small and counterintuitive (in the second sense). To see this, consider a spatial model of voter choice as a function of cues and candidate proximity. For members of any one political party,

$$U_{Ai} = \gamma_1 \left[f\left(S_i - A\right) \right] + \gamma_{A2} cue_{Ai} + \epsilon_{Ai}$$

$$U_{Bi} = \gamma_1 \left[f\left(S_i - B\right) \right] + \gamma_{B2} cue_{Bi} + \epsilon_{Bi}$$

$$p(i \text{ prefers candidate A} \mid S_i, A, B, cue_{Ai}, cue_{B_i}) = p\left(U_{Ai} > U_{Bi}\right)$$

$$= p\left(\gamma_1 P_{*i} + \gamma_{A2} cue_{Ai} - \gamma_{B2} cue_{Bi} + \epsilon_i > 0\right).$$

U_{Ai} and U_{Bi} are the utilities that voter i expects to derive from having candidates A and B in office. The locations of voter i and the candidates are given by S_i, A, and B, respectively. P_{*i} is the *generalized proximity difference*, $f\left(S_i - A\right) - f\left(S_i - B\right)$,

where $f(x)$ is a real-valued function that is weakly monotonic in $|x|$, such that greater values of $f(x)$ always indicate greater distances between the voter and the candidate. If $f(x) = |x|$, the generalized proximity difference becomes the simple proximity difference defined above. If $f(x) = x^2$, there is a close correspondence between this model and quadratic utility models of voter choice (e.g., Davis, Hinich, and Ordeshook 1970). If the voter is closer to A than to B, P_{*i} is negative; if he is closer to B, P_{*i} is positive. In either case, I expect γ_1 to be negative, reflecting the assumption that voters are less likely to prefer a candidate as he moves away from their ideal points. The assumption is strongly supported by Figure 1.5.

cue_{Ai} and cue_{Bi} are binary variables scored 1 if the voter knows the candidates' party affiliations and 0 if he does not; γ_1, γ_{A2}, and γ_{B2} are coefficients; and ϵ_i, ϵ_{Ai}, and ϵ_{Bi} are disturbance terms (with $\epsilon_i = \epsilon_{Ai} - \epsilon_{Bi}$) drawn from distributions that are single-peaked and symmetric about zero, e.g, the logistic or Normal distributions. Because voters usually know either both candidates' party affiliations or neither of them, a simplified form of the voting rule is appropriate:

$$p(i \text{ prefers candidate A}) = p(\gamma_1 P_{*i} + \gamma_2 cues_i + \epsilon_i > 0), \tag{1.1}$$

where γ_2 is a coefficient and $cues_i$ is a binary variable scored 1 if the voter has learned both candidates' party affiliations and 0 if he has not.

By this model, the utility that a voter expects to derive from having a candidate in office is a function of his spatial distance from the candidate and his knowledge of the candidate's party.[12] He votes for the candidate from whom he expects to derive

---

[12] Conveniently, this model is agnostic about whether party cues influence vote choice because they change the candidates' perceived policy locations or because they exert an effect—perhaps an emotional effect—on another dimension that voters consider when making decisions. This is an interesting topic but not one that current experiments permit us to learn about. If party cues convey information that seems orthogonal to policy, they can be thought of as signals about the candidates' locations on a "valence dimension": see Groseclose 2001; Londregan 2000; Stokes 1963.

greater utility. If the candidates are equidistant from him ($P_{*i} = 0$) and he does not know the party cues, his choice is determined by ϵ_i, and he is as likely to prefer A as to prefer B.

The effect of cues under this model is

$$p(i \text{ prefers A} \mid P_{*i}, cues_i = 1) - p(i \text{ prefers A} \mid P_{*i}, cues_i = 0)$$

$$= \Lambda\,(\beta_1 P_{*i} + \beta_2) - \Lambda\,(\beta_1 P_{*i})\,,$$

where Λ is the cdf of ϵ_i. The maximum effect of cues therefore occurs when $\beta_1 P_{*i} = -\beta_2/2$, i.e., when the proximity effect ($\beta_1 P_{*i}$) is half the size of the cue effect and of the opposite sign. (See Appendix D for a proof.) In practice, this occurs when P_{*i} is small but counterintuitive, such that i's ideal point is slightly closer to the position of the opposite-party candidate than to the position of his own party's candidate. This is intuitive: if I am a Democrat and one candidate is much closer to my ideal point than the other, I prefer the closer candidate and am unlikely to switch my preference upon learning the party cues. If the closer candidate is only slightly closer, I still prefer him initially and have no reason to switch upon learning that he is a Democrat. But if I learn instead that the slightly-closer candidate is a Republican, I have more cause to switch my vote: this is the scenario under which party cues seem likely to have the greatest effect.

Intuitive as this is, subjects in Experiment 3 do not think this way. For them, the effect of cues does not depend on the positions of the candidates. Figure 1.5 provides the evidence. In the Democratic panels, cue effects—measured by the distance between **N** and **C**—do not vary with the proximity difference. In the Republican panels, the story is the same when the Republican candidate is closer to subjects' ideal points; when the Democratic candidate is closer, the distance

between **N** and **C** does vary, but not systematically. Related evidence appears in Table 1.4, which reports estimates from ordered logistic regression models of candidate preference as a function of cues and the absolute value of the proximity difference. The overall message, given by the cues × |proximity difference| and cues × |proximity difference| × Dem. candidate is closer coefficients, is plain: the effect of cues on the latent scale does not vary with the proximity difference. The only strong exception lies with Republican subjects considering the environment who are closer to the Democratic candidate's ideal point. And for them, proximity matters in an unexpected way: cues have their *smallest* effect on the latent scale when the Democratic candidate is only slightly closer, their greatest effect when he is much closer.[13] The same pattern emerges, albeit weakly, among Republicans considering social services who are closer to the Democratic candidate ($p = .32$). The expected pattern appears only among Democrats considering social services who are *not* closer to the Democratic candidate ($p = .07$). In all other cases, no effect is apparent.[14]

Discussion

The prevailing wisdom tells us that "cues affect results when respondents have no alternative basis for the formation of issue appraisals" (Mondak 1993b, 207). The findings reported here show that an amendment is required. Cues also affect results when respondents have a strong alternative basis—that is, an extraordinary amount of

---

[13] When the Democratic candidate is closer by only one position (|proximity difference| = .167), the effect of cues on the latent scale is $.72 - (.68 \times .167) + (3.05 \times .167) = 1.16$. When he is far closer (|proximity difference| = 1), the effect is $.72 - .68 + 3.05 = 3.09$.

[14] Of course, the regressions reported in Table 1.4 do not directly test the prediction of the formal model: that prediction is about the effect of cues on the probability of supporting a candidate, whereas Table 1.4 indicates the effect of cues on the logit scale. But the interaction effects reported in Table 1.4 show that the effect of cues on the logit scale generally increase as the proximity difference grows, which suggests, albeit indirectly, that the prediction of the formal model is wrong: cue effects on the probability of supporting a candidate are unlikely to be highest when the proximity difference is small.

	social services Dem subjects	environment Dem subjects	aid to blacks Dem subjects
cues	−.81* .22	−.68* .25	−.32 .24
|proximity difference|	1.67* .52	1.01 .62	2.65* .62
Dem. candidate is closer	−.74* .26	−.96* .30	−.66* .30
cues × |proximity difference|	−.92+ .61	−.04 .73	−.53 .73
cues × Dem. candidate is closer	.00 .32	.04 .36	.08 .35
|proximity difference| × Dem. candidate is closer	−3.29* .72	−2.73* .87	−4.21* .87
cues × |proximity difference| × Dem. candidate is closer	.78 .88	.15 1.05	.52 1.04
τprefer A or Democrat | prefer neither	−1.25 .19	−1.17 .21	−.82 .20
τprefer neither | prefer B or Republican	.09 .18	−.08 .21	.55 .20
Log likelihood	−2728.15	−2551.30	−2658.19
Likelihood ratio test	1568.25	1094.45	1148.49
Number of observations	2900	2737	2775

	GOP subjects	GOP subjects	GOP subjects
cues	.57* .29	.72* .29	.18 .30
|proximity difference|	1.47* .64	2.27* .72	1.75* .67
Dem. candidate is closer	−.83* .31	−.40 .32	−1.58* .34
cues × |proximity difference|	.61 .87	−.68 .91	.24 .83
cues × Dem. candidate is closer	.16 .39	−.67 .40	.33 .41
|proximity difference| × Dem. candidate is closer	−4.20* .95	−5.80* .99	−3.71* .93
cues × |proximity difference| × Dem. candidate is closer	.86 1.19	3.05* 1.22	.30 1.13
τprefer A or Democrat | prefer neither	−1.30 .23	−1.02 .23	−1.75 .25
τprefer neither | prefer B or Republican	.10 .22	.12 .23	−.21 .24
Log likelihood	−2120.63	−2098.66	−2032.47
Likelihood ratio test	1072.82	1130.87	1205.28
Number of observations	2324	2345	2350

Table 1.4: Candidate Proximity Does Not Moderate Cue Effects. Cell entries in the top rows of each tier are parameter estimates and standard errors from ordered logistic regressions. Entries in the "likelihood ratio test" rows are χ^2 statistics from tests against a model with no predictors; all are significant at $p < .001$. The dependent variable is the answer to "which candidate better represents your position on this issue?" Larger parameter estimates indicate a higher probability of picking the Republican candidate, who is identified as "Candidate B" in the uncued condition. In the "cues," "Dem. candidate is closer," and "cues × \|proximity difference\|" rows, interesting estimates significant at 95% using a one-tailed test for $H_A>0$ are denoted by *; at 90%, by +. In all other rows, the tests are two-tailed.

"Cues" indicates whether the subject was told the candidates' party affiliations: subjects who did are coded as 1; those who didn't are coded as 0. "Dem. candidate is closer" is a dummy variable scored 1 if the Democratic candidate was closer to the subject's ideal point, 0 otherwise. "\|proximity difference\|" is the absolute value of the proximity difference defined in the text but rescaled to range from 0 to 1. A value of 0 indicates that the candidates are equidistant from the subject's ideal point. A value of 1 indicates that one candidate is much closer than the other to the candidate's ideal point.

policy information—for their attitudes. But whether people are well- or ill-informed, party cue effects are smaller than we often imagine. And when people receive policy information, even in slender amounts, they respond more to it than to party cues.

Why Do We Expect Larger Party Cue Effects?

Cue-based processing usually "predominates" consideration of the information contained in messages (Iyengar and Valentino 2000). This is especially true of political situations (Zaller 1992, ch. 3). And "when voters have both particular information and party stereotypes . . . they use the label rather than policy attributes in drawing inferences" (Rahn 1993, 492). All three of these claims appear to conflict with the findings reported here. But the conflict is more apparent than real.

Begin by noting that the first two claims are about cues writ large rather than about party cues alone. Politicians are prolific and strategic cue-givers (Tessin 2006); party cues are far from the only cues that they offer. It is entirely possible, then, that party cues have small effects while the cumulative effects of different types of cues are large.

A second possibility is that party cues have an indirect effect not captured by these experiments. Party cue information is prevalent and easily understood; people are likely to acquire it before they acquire other kinds of information. And if knowing party cues makes people unlikely to seek out other kinds of information, the effects of party cues will be correspondingly enhanced, because they will be all the information that people have. This idea is explicit but apolitical in Herbert Simon's models of satisficing, implicit and political in Downs (1957) and Popkin (1994). It permits direct experimental tests, e.g., through information board experiments like those used by Lau and Redlawsk (2001). To date, no such tests have been conducted.

A final and related explanation lies with the amount of information that subjects possess in these experiments. All of them, even in Experiment 3, possess at least a modicum of policy information distinct from the party cues. But most people, most of the time, may possess no policy information at all when they make political decisions. This is especially true of policy initiatives, referenda, and the contests for low-level offices that comprise the largest part of any American ballot. In these cases, we should not be surprised to find that party cues have larger effects on voters' choices (as it seems that they may: Ansolabehere et al. 2006; Schaffner and Streb 2002).

Normative Implications

Much of the concern about cues rests on the belief that they mislead citizens into holding views that they would not hold if they were better informed. Undoubtedly, cues of all sorts do have this effect. But there are two reasons to be relatively sanguine about party cues. One is that, as this work shows, the effects of party cues are smaller than we often imagine. Another is that party cues may be more reliable "information shortcuts" than many other types of cues, at least in the American context. While other cues provide information that is at best irrelevant to policy decisions—what is the political relevance of a speaker's attractiveness?—inferences prompted by *party* cues are built on parties' reputations, which are well-established and informative in America (Petrocik 1996; Kingdon 1984; see also Ansolabehere, Snyder, and Stewart 2001)—and only growing more informative as the parties polarize (Poole and Rosenthal 2006). Of course, even reliable reputations can mislead people sometimes. But resting the case against party cues on this observation requires an argument that party reputations will be misleading more often than not. That is a difficult argument to make.

What of the finding that partisans see themselves as unaffected by cues but members of the other party as quite affected? Unlike cue use itself, this is inherently worrisome. It should be especially troubling to readers who subscribe to deliberative theories of democracy, because deliberative theories demand that citizens see each other as committed to advancing the public good rather than any merely sectarian interest (e.g., Cohen 1989; Rawls 1993). And this is precisely how partisans will not view each other if they believe that members of the other party are especially influenced by their party's elites. But one need not be a deliberative democrat—or, indeed, much of a democrat at all—to be troubled by the finding. One need only lament the degree to which discussions of policy are colored by partisan rancor. Wildly disparate beliefs about cue influence probably do not create that rancor, but they are a distressing symptom of it.

Appendix A: Realism and Informational Confounds in Party Cue Experiments

Cue experiments often begin by exposing subjects to newspaper articles in which the cues are embedded (e.g., Cohen 2003; Kam 2005; Schwieder and Quirk 2004). But it is difficult to write even superficially realistic articles. Schwieder and Quirk (2004) identify five types of information common to articles about policies: details of the proposed changes, source cues, horse race information, human interest information that lacks direct political relevance, and contextual information that ties the proposed changes to the problems that they are meant to address. The fabricated articles used in experiments often lack most of these elements. And there are other problems. No real article about a policy change will devote itself to describing the stands of "Democrats" or "Republicans" on a bill without quoting an actual Democrat or Republican. Nor, if the article is about government-provided social services, will it neglect to make at least a brief mention of the costs of the changes. But these omissions are common in party cue experiments.

A graver problem is the failure to describe the status quo in articles about potential policy changes. This is not just a matter of technical infidelity to actual news coverage. Typically, one version of the article used in party cue experiments has Republican legislators supporting a proposed policy and Democratic legislators opposing it. In a second version, the party positions are reversed—this is the manipulation. The policy is described in the same way in both conditions. But if the status quo is not also described and held constant across conditions, a confound may arise: subjects' perceptions of the status quo may vary with party cues. For example, subjects may think the status quo relatively liberal if Democrats oppose the changes but relatively conservative if Republicans oppose them. This is true even if the article

stipulates that the alternative is more liberal or conservative than the status quo. The distance between the status quo and the policy alternative can therefore vary across conditions, just as party cues do. In these cases, the confound makes it impossible to isolate the independent effects of party cues.

The problem is compounded when politicians' arguments vary with party cues. This happens, for example, when Republicans favoring a policy in one condition offer reasons quite different from those given by the Democrats who support the same policy in the next condition. In these cases, the confound is overt. Between-condition differences in the distributions of subjects' policy attitudes may be due to changes in party cues or to changes in politicians' arguments.

Appendix B: Experiment 1 Article Text

All subjects assigned to the liberal information and intuitive cue conditions received this article:

> Gov. David Brady won a key budget battle Thursday as the House sent him a bill authorizing the expansion of Medicaid health coverage for tens of thousands of low-income residents. The House's 87-71 vote came on the same day its Budget Committee was finalizing a roughly $19 billion spending plan that would implement the Medicaid expansion beginning July 1. 80 of 89 House Democrats voted for the bill, while 62 of 69 House Republicans voted against it.
>
> Brady, a Democrat, and Democratic legislative leaders said the expansion is needed to protect the disabled, elderly, and parents who currently lack coverage.
>
> But Republican opponents contend the expansion could lead to reduced school funding, a budget deficit, and higher taxes. They also argued that the expansion could threaten the long-term sustainability of the state's other social welfare services.
>
> The plan would increase health care coverage for nearly 100,000 of Wisconsin's 1 million Medicaid recipients by loosening eligibility standards, and it would add certain services such as dental care for many others. It also would reduce co-payments or premiums for hundreds of thousands of Medicaid enrollees.
>
> Brady praised the Legislature for taking "decisive actions to protect the poorest among us." He said the bill's anti-waste and fraud provisions—such as annual Medicaid eligibility reviews—would "ensure that scarce state resources are going to those in need."

The bill would expand mandatory Medicaid coverage of such things as wheelchairs, artificial limbs and eye care for most adults. It is expected to reduce waiting times for wheelchairs and prostheses. Adult Medicaid recipients would be permitted to receive eye care visits once every year. Recipients are currently permitted one eye care visit every two years.

A late provision added by the House would also expand a program that provides Medicaid coverage to disabled people aged 16 to 64 if they work at least three hours a month. Currently, disabled adults qualify for coverage if they earn less than $1,940 a month. The House bill raises the cutoff to $2,600 a month.

Opponents of the expansion point to the growth of Medicaid. In the past dozen years, the Medicaid rolls doubled while its cost nearly tripled. Yet even without the proposed expansion, Medicaid would cost more than $5.5 billion in state and federal money next fiscal year, consuming nearly 29 percent of Wisconsin's budget.

The expansion is dangerous because "we must ensure the children of our state can be educated, that our most vulnerable are protected, and (that) we do it in such a manner that creates solid footing for the state of Wisconsin," said House Budget Committee Member David Toolan, R-Milwaukee.

But supporters claim the Medicaid expansions would ensure that the most vulnerable receive necessary protections.

Currently, most adult Medicaid recipients are required to make co-payments of between 50 cents and $3, depending on the cost of the service, each time they visit a doctor or hospital. The House bill would eliminate copayments.

The bill also would eliminate monthly premiums of families in the MC+ for Kids program, which provides health care to children whose families earn up to three times the federal poverty level but aren't covered by traditional Medicaid or private insurance. Because some families will join the program if the premiums are eliminated, the Department of Social Services estimates about 23,700 children will gain coverage.

Under the House version, a single parent of two could earn no more than $2,184 a month to qualify for Medicaid. The current cutoff for single parents of two is $1,334 a month.

Rep. Connie Zimmer, D-Mellen, said she gets a $493.50 state mileage check for driving to the Capitol each month.

To qualify for Medicaid under current regulations, "we're telling somebody that they should raise a family of three for less money than any three of us get for gas, and that's hypocritical," she said.

The bill is SB 593.

Liberal Policy Changes	Status Quo	Conservative Policy Changes
Monthly income cutoff: $2184 for a single parent of two	Monthly income cutoff: $1334 for a single parent of two	Monthly income cutoff: $484 for a single parent of two
No copayments for visits to doctor	Copayments for visits to doctor: 50 cents to $3	Copayments for visits to doctor: $4-$10
Coverage for children: eliminate premiums for some families, leading 23,700 children to gain coverage		Coverage for children: require premiums for more families, leading 23,700 children to lose coverage
Expand mandatory coverage of wheelchairs, prostheses, and eye care. Reduce waiting times for wheelchairs and prostheses. Eye care visits once every year.	Coverage of wheelchairs, and prostheses. Eye care visits once every two years.	Repeal mandatory coverage of wheelchairs, prostheses, and eye care. But budget would continue to fund wheelchairs, prostheses, and eye care visits once every three years.
Coverage of temporarily disabled people aged 16 to 64 who earn less than $2,600 per month and work at least three hours per month	Coverage of temporarily disabled people aged 16 to 64 who earn less than $1,940 per month and work at least three hours per month	Eliminates coverage for the temporarily disabled
Expand coverage for 100,000 of the state's one million Medicaid recipients		Reduce coverage for 100,000 of the state's one million Medicaid recipients

Table B1: Policy Details in the Liberal and Conservative Information Conditions of Experiment 1. All subjects in Experiment 1 read a newspaper article that contrasted the status quo with liberal or conservative policy changes that had just been passed by the state House of Representatives.

Appendix C: Cue Effects in Experiment 3

	social services		environment		aid to blacks	
subject's party ID	−.89*	.40	1.15*	.40	−.74*	.40
cues	−.98*	.29	−1.02*	.29	−.22	.28
proximity difference	5.35*	.46	4.83*	.46	5.68*	.45
subject's party ID × cues	3.06*	.49	2.67*	.49	1.39*	.48
subject's party ID × proximity difference	1.22+	.77	2.10*	.78	1.93*	.75
cues × proximity difference	−.82+	.55	.02	.55	−.71+	.54
subject's party ID × cues × proximity difference	−.77	.94	−1.46+	.95	−.49	.90
τprefer A or Democrat | prefer neither	1.65	.24	1.70	.24	2.09	.24
τprefer neither | prefer B or Republican	3.15	.24	2.93	.24	3.62	.24
Log likelihood	−6389.26		−6321.28		−6240.79	
Likelihood ratio test	2112.10		1941.40		2203.29	
Number of observations	6725		6676		6636	

Table C1: Cue Effects in Experiment 3 as a Function of Seven-Category Party ID and the Proximity Difference. Cell entries are parameter estimates and standard errors from ordered logistic regressions. Entries in the "likelihood ratio test" rows are χ^2 statistics from tests against a model with no predictors; all are significant at $p < .001$. The dependent variable is the answer to "which candidate better represents your position on this issue?" Larger parameter estimates indicate a greater tendency to pick Candidate B (in the uncued condition) or the Republican candidate (in the cued condition). Interesting estimates significant at 95% using a one-tailed test for $H_A > 0$ are denoted by *; at 90%, by +.

"Subject's party ID" is a seven-category variable ranging from 0 ("Strong Democrat") to 1 ("Strong Republican"). "Cues" indicates whether the subject was told the candidates' party affiliations: subjects who did are coded as 1; those who didn't are coded as 0. "Proximity difference" is the simple proximity difference defined in the text but rescaled to range from 0 to 1. Values below .5 indicate that the Democratic candidate was closer to the subject's ideal point. Values above .5 indicate that the Republican was closer. A value of .5 indicates that the candidates were equidistant from the subject's ideal point.

Proportions Preferring Republican Candidates
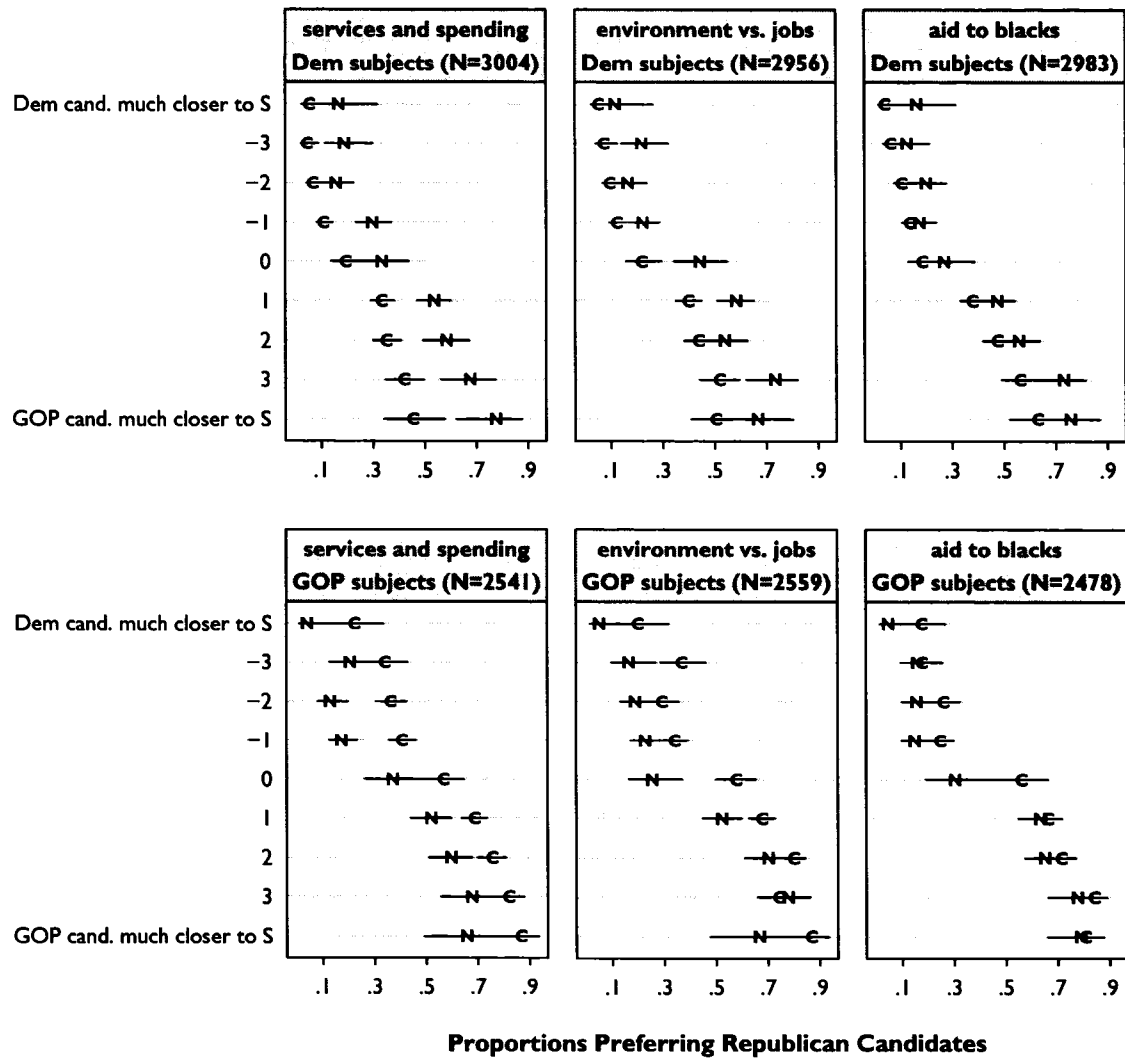
Figure C1: Party Cue Effects on Candidate Preference by Issue and Proximity: Proportions Preferring the Republican Candidates. Each row of each panel plots proportions of subjects (S) preferring the Republican candidate's position. N indicates the proportion in the no-cue condition, in which candidates were identified only as "Candidate A" and "Candidate B." C indicates the proportion in the cued condition. Black lines are 95% confidence intervals.

As in Figure 1.5, three empirical regularities emerge. Party cue effects are usually larger—slightly—in these extremely-low-information experiments than in the previous experiments. Even in these experiments, subjects are more influenced by policy information—the candidates' issue stances—than by party cues. And cue effects vary little with the distances of candidates from subjects' ideal points. (See text for details.)

Appendix D: Proof that Cues Have Their Greatest Effect When $\beta_1 P_i = -\beta_2/2$

Under the model of voter preferences in Equation 1.1 (page 30), the effect of cues on voter preferences is

$$p(i \text{ prefers A} \mid P_{*i}, cues_i = 1) - p(i \text{ prefers A} \mid P_{*i}, cues_i = 0)$$

$$= \Lambda\,(\beta_1 P_{*i} + \beta_2) - \Lambda\,(\beta_1 P_{*i}), \tag{1.2}$$

where Λ is the cdf of ϵ_i. (Recall that the pdf of ϵ_i is single-peaked and symmetric about zero.) With no loss of generality, we assume that Λ is the logistic cdf. We find the maximum effect of cues by taking the derivative of Equation 1.2 with respect to $\beta_1 P_{*i}$, setting it to 0, and solving for $\beta_1 P_{*i}$:

$$\frac{\partial[\Lambda\,(\beta_1 P_{*i} + \beta_2 cues_i) - \Lambda\,(\beta_1 P_{*i})]}{\partial \beta_1 P_{*i}} = \lambda\,(\beta_1 P_{*i} + \beta_2) - \lambda\,(\beta_1 P_{*i})$$

$$\Rightarrow \frac{\exp(\beta_1 P_{*i} + \beta_2)}{(1 + \exp(\beta_1 P_{*i} + \beta_2))^2} = \frac{\exp(\beta_1 P_{*i})}{(1 + \exp(\beta_1 P_{*i}))^2}$$

$$\Rightarrow \frac{\exp(\beta_2)}{(1 + \exp(\beta_1 P_{*i} + \beta_2))^2} = \frac{1}{(1 + \exp(\beta_1 P_{*i}))^2}$$

$$\Rightarrow \exp(\beta_2)\,[1 + 2\exp(\beta_1 P_{*i}) + \exp(2\beta_1 P_{*i})] = 1 + 2\exp(\beta_1 P_{*i} + \beta_2)$$

$$+ \exp(2[\beta_1 P_{*i} + \beta_2])$$

$$\Rightarrow \exp(\beta_2) + \exp(2\beta_1 P_{*i})\exp(\beta_2) = 1 + \exp(2\beta_1 P_{*i})\exp(2\beta_2)$$

$$\Rightarrow \exp(2\beta_1 P_{*i})\exp(\beta_2) - \exp(2\beta_1 P_{*i})\exp(2\beta_2) = 1 - \exp(\beta_2)$$

$$\Rightarrow \exp(2\beta_1 P_{*i}) = \exp(-\beta_2)$$

$$\Rightarrow \beta_1 P_{*i} = -\beta_2/2.$$

Because the logistic pdf is globally concave, this is a unique maximum.

Part 2

Partisanship and the Enduring Importance of False Political Information

Of all the arguments made to martial support for the 2003 invasion of Iraq, the most prominent and successful was that Iraq had and was going to use weapons of mass destruction (Gellman and Pincus 2003; Program on International Policy Attitudes 2003). But contrary evidence was mounting before the invasion, and shortly afterward, it became overwhelming. In the months before the war, Iraq did not possess weapons of mass destruction. Nor was it attempting to make or acquire them (Duelfer 2004, ch. 1; see also Gellman 2004). As students of public opinion, we should wonder: if supporters of the war had been disabused in time—if, after believing that Hussein's regime possessed WMD, they had learned that it did not—would public support for the war have diminished?

The more general question is how citizens respond to political information. Empirically, we expect attitudes to bear some connection to facts and to change as relevant new facts come to light. And normatively, this is what contemporary

47

democratic theory prescribes: that citizens know truths about politics and use them to shape their views. It seems almost redundant to add that people should not be influenced by falsehoods. But while great effort has been spent to determine what people know about politics, little has been spent to determine the extent to which they "know"—i.e., believe—false factual claims about politics. Still less has been spent to examine the possibility that people are influenced by messages that they know to be false.

If anything, a nearly opposite assumption is made: people may be deceived but they will, if undeceived, change their views accordingly. This is why "ad watches" that fact-check political advertisements have become staples of campaign news (Ansolabehere and Iyengar 1995). And it is why studies of the connection between deliberation and attitudes typically involve not just deliberation but exposure to facts in a specific policy area (e.g., Luskin, Fishkin, and Jowell 2002). These efforts are not undertaken because truth is held to have intrinsic value. They are undertaken because learning the truth is supposed to affect attitudes. It may be difficult to correct people's mistaken beliefs, but once that task is accomplished, the assumption is that attitudes immediately related to the beliefs will change accordingly.

Research from the intersection of social and cognitive psychology suggests that the assumption is not always borne out in practice (e.g., Ross, Lepper, and Hubbard 1975). Even when old beliefs are debunked, the attitudes affected by those beliefs sometimes change little or not at all. That said, previous work on this subject has been almost wholly apolitical. In this article, I use real-world cases of political misinformation and a trio of experiments to show that the effects obtain in politics, too: false information influences people's political views even after it is overwhelmingly refuted and perhaps even after it is understood to be false. The strength of the effect depends on the partisanship of the people whose views are at

issue. These findings explain the prevalence of deceptive rhetoric in American politics, even in high-information contexts in which deception is likely to be exposed. From the standpoint of strategic communication, they emphasize the damage that can be done even by baseless slanders, thereby underscoring the importance of protecting one's reputation and the difficulty of repairing it. Most of all, they raise new questions about elites' ability to manipulate voters.

Motivated Reasoning Can Explain Partisan Differences in Reactions to Discredited False Information

People tend to resist arguments and evidence that run counter to their attitudes (McGuire 1964; Lord, Ross, and Lepper 1979; Tetlock 2005). But how much resistance is too much? Some, at some times, is surely sensible, and in real-world situations it is often difficult to say how much attitudes should change in response to contrary new evidence.

At first glance, though, one case does not seem difficult. When new evidence completely discredits old information, attitudes once influenced by the information should no longer be influenced by it. They should become what they would have been if the information had never been encountered at all. In reality, though, people sometimes fall short of this standard. They seem to accept the falsity of a claim but continue to be affected by it. "Belief perseverance" is the name of the phenomenon. In politics, for example, it may occur if voters learn that an attack on a candidate is false but continue to think worse of him because of it.

Why might people be affected by what they know to be false? The cognitive mechanisms that constitute *motivated reasoning* (Kunda 1990) may provide an answer.

When people encounter a message about a politician or a policy that is not absolutely novel, they have an immediate affective reaction—sometimes strong, sometimes quite weak, but always immediate and largely shaped by prior political beliefs (Bargh et al. 1992; Lodge and Taber 2000). The encounter with the message also unconsciously primes related ideas in long-term memory, making them more accessible (Krosnick and Kinder 1990; Higgins 1996). It often sparks a conscious search for related memories, too (Anderson, New, and Speer 1985). And it spurs people not just to recall related memories but to use them to explain the information contained in the message (Kelley 1973; Anderson, Lepper, and Ross 1980). None of these tasks are typically undertaken in an evenhanded fashion. The search of memory, whether conscious or not, is often biased in favor of finding data that support one's prior beliefs. And explanations of the new information are also constructed with an eye to supporting those beliefs.

These cognitive processes therefore tend to produce, highly available in memory, a pattern of evidence supporting one's initial reaction to a new message. And even if one is later persuaded that the message is false, the explanations that he has constructed and the memories that he has recruited will still seem to justify that initial reaction. Consequently, the falsification of a message will not produce a corresponding change in one's attitude toward the subject of the message—even if the attitude once depended on the message for its existence. The attitude quickly becomes independent of the information that engendered it.

This theory can explain why students who receive good scores on ability tests continue to think better of their abilities even when they learn that the scores were fabricated (Ross, Lepper, and Hubbard 1975; see also Lepper, Ross, and Lau 1986). And it can explain why people who view a negative advertisement about a candidate may think worse of the candidate even after learning that the ad is false. It also predicts

partisan differences in reactions to negative ads. To see this, consider Republicans and Democrats who view a false negative ad about a Democrat. Initially, the ad's false claim may cause members of both parties to think worse of the Democrat. But Republicans are more likely to possess related negative ideas about the Democrat and his party, and chiefly for that reason, Republicans will be more likely to retrieve related negative ideas from long-term memory upon seeing the ad. Later, if the ad proves false to all, Republicans will be more likely think worse of the Democrat—not because they still believe the ad's claim, but because the related negative ideas that it has summoned will now seem to justify their lower opinion. *Ceteris paribus*, this should be equally true for Democrats who view a false negative ad about a Republican.

Experimental Designs

Owing to a dearth of data, the political effects of belief perseverance have almost never been studied. They received their most thorough consideration in a *Journal of Politics* article by Kuklinski et al. (2000), who sought to examine the effect of new information on people who were misinformed about welfare. In their experiment, which was embedded in a telephone survey, control-group subjects were asked factual questions about welfare.[1] Most of them far overestimated the generosity of federal welfare programs. And because subjects were randomly assigned to the control group or to the treatment groups, we can assume that treatment-group subjects, too, were misinformed about welfare policy. But interviewers told treatment-group subjects a series of facts about welfare—that was the treatment. In the authors' view, those facts should have corrected their misconceptions about welfare. Later in the experiment, though, all subjects were asked to state their preferences over welfare policy—and the

---

[1] Kuklinski et al. (2000) report two similar experiments. This is a description of the first one, to which the bulk of their paper is devoted.

treatment appeared to make no difference at all. Tentatively, the authors concluded

that they had discovered a case of belief perseverance: their subjects "absorb[ed] the

facts" but nevertheless "failed to change their preferences accordingly" (Kuklinski

et al. 2000, 802-03).

That may be what happened. But as the authors acknowledge, it is also possible

that the facts were not "absorbed" at all. They might have been rejected outright by

the subjects. Consider the predicament of the typical subject in their study: without

advance notice, he receives a phone call from a stranger who purports to be calling

from a university but proceeds to pelt him with factual claims about welfare—claims

that contradict his prior beliefs. What compelling reason does he have to accept these

new claims? Probably none. It may be that Kuklinski et al. uncovered a case of belief

perseverance, but it may also be that their subjects rejected the information they

heard over the phone, attributing it to the partisan motivations of a misinformed or

disingenuous interviewer. In this experiment, there is no decisive way to distinguish

between the two possibilities.

A third possibility also cannot be distinguished from the others. It may be that

subjects believed all the new information provided by the interviewer, considered it

thoroughly, and deemed it irrelevant to their attitudes about welfare or insufficient

reason to change those attitudes. Consider the facts that were given to treated subjects:

in addition to learning the average amount of money given annually to welfare

families, they learned the percentages of families on welfare, of welfare families

that were African-American, of mothers on welfare for more than eight years, of

welfare mothers who had less than a high school education, and of the federal budget

devoted to welfare. There is no logical inconsistency between learning these facts

and continuing to favor welfare cuts—even if hearing the facts makes one realize that

he has been overestimating the generosity of welfare programs. This possibility, too, makes it difficult to determine whether belief perseverance was at work.[2]

Mindful of these difficulties, psychologists embrace different designs to test for belief perseverance. Typically, they deceive subjects about their ability to perform a task and then undeceive them in what seems to be a convincing fashion. The studies in Ross, Lepper, and Hubbard (1975) are instructive: at the beginning, subjects were given pairs of suicide notes and were asked, for each pair, to judge which was authentic and which was fake. After each judgment, they were told that they were correct or incorrect. But toward the end of the experiment, they learned that this feedback was simply made-up. They had been randomly assigned to hear feedback that was mostly positive or mostly negative. The feedback bore no relation to their actual performance at the task. The experimenter emphasized this point. The subjects affirmed that they understood. Yet, when asked how well they performed at the task and how well they would do it in the future, subjects in the "positive feedback" condition gave answers quite different from those in the "negative feedback" condition. This seems to be a case of perseverance: subjects knew that they had received false information but continued to be affected by it.

Consider the two advantages of this design. First, the discrediting of old information is probably much more believable here than in the welfare experiment. In part, this is because it seeks not to destroy a prior belief that may be deeply felt (e.g., belief in the unfairness of welfare policy), but only to convince subjects that a message they heard earlier in the experiment isn't credible. And in part, it is because the discrediting is supplied by the same person who earlier supplied the false information.

---

[2] A separate experiment in Kuklinski et al. (2000) uses a different design that makes this possibility a less likely explanation for the weak effect of the treatment.

When the experimenter tells the subject that he has been lying, the subject cannot easily dismiss him as uninformed.

Second, belief in the discrediting information necessarily entails disbelief in the feedback that was provided earlier. To the extent that subjects accept the discrediting, they are logically committed to believing that the earlier feedback about their performance on the test provides no information about their abilities. (Compare this to the welfare experiment, in which subjects might have believed all that they were told by the interviewer and still—quite logically—not have changed their attitudes toward welfare at all.)

These advantages permit a simple test of perseverance. At the end of the experiment, subjects in the "positive feedback" group and those in the "negative feedback" group should not possess different beliefs about their abilities. The extent to which they do possess different beliefs—easily gauged with a t-test—is the extent of perseverance.

I use an experimental design that borrows elements from both of the designs already described. Following Kuklinski et al., the experiments examine political topics, and the first two do not rely on overt deception by the experimenter. Following Ross, Lepper, and Hubbard, I provide false information during the experiment and then discredit it, instead of trying to disabuse subjects of beliefs that they had before the experiment began. The last experiment presented here follows their design closely, but the signal difference in the first two is that false information is neither provided nor discredited by the experimenter. Instead, I use real news media and advertisements to provide and discredit false political information. I examine treatment-group subjects' attitudes immediately after they receive false information and after the subsequent discrediting, but in all cases, "the treatment" is defined as receipt of both the false information and the discrediting.

Of course, there is no way to be sure that the discreditings of false information in these experiments are trusted completely. Just as in the original perseverance experiments, we cannot peer into subjects' brains to gain certain knowledge of their beliefs. The reader is free to draw his own conclusions about the success of the discreditings; the crucial point is not that they succeeded but that if they failed, real-world discreditings are extremely unlikely to succeed. After all, most political beliefs are stronger and longer-lasting than those that are merely instilled during experiments. Moreover, real-world challenges to beliefs do not arise from authoritative sources: journalists, for example, are extremely reluctant to contradict politicians' factual claims (Jamieson and Waldman 2003, Chapter 7), and there is no analog in ordinary political experience to the experimenter who is credible as the debunker of false information because he is the one who provided it in the first place. Even when the political world does produce decisive refutations of false beliefs, selective exposure to congenial news sources is likely to shield people from them (Gentzkow and Shapiro 2006; Taber and Lodge 2006). In all of these respects, the experiments reported here are conservative tests of the perseverance hypothesis. If the effects of false information persist after the discreditings in these experiments, they are quite likely to persist in real life, where discreditings are much weaker.

Two hypotheses are at stake. The first is that people are affected by false political information even after that information is discredited in what is, by real-world standards, an extremely strong fashion. The second, the *motivation hypothesis*, is that partisanship helps to sustain the effects of false information.

If the theory described in the previous section is correct, all subjects should not be affected in the same way by a false attack on a candidate or a policy. All may initially think worse of the candidate or policy because of the attack, but the extent

to which the attack prompts the retrieval of related negative beliefs should depend on party identification. Democrats who hear attacks on a Republican candidate or policy should be more likely to recall related negative beliefs, and thus be relatively unmoved when they learn that the attacks were not credible. On the other hand, Republicans should be less likely to recall related negative beliefs, and thus more likely to recover when they learn that the attacks are not credible.

Experiment 4: John Roberts and the Abortion Clinic Bombing

On August 8, 2005, NARAL Pro-Choice America (formerly the National Abortion Rights Action League) released a television advertisement that accused Supreme Court nominee John Roberts of "supporting violent fringe groups and a convicted clinic bomber." The ad began by depicting an abortion clinic in ruins, segued to injured women in wheelchairs, and ended with a voiceover admonition that "America can't afford a Justice whose ideology leads him to excuse violence against other Americans" (NARAL Pro-Choice America 2005).

Opponents and supporters of abortion rights were quick to criticize the ad as "blatantly untrue" and "deceptive." They noted that *Bray v. Alexandria*, the case in which Roberts was accused of supporting violent anti-abortion protesters, was not about clinic bombing. (It was about the use of civil rights statutes to prosecute protesters who blockaded abortion clinics.) The only connection to bombing was tenuous: seven years *after* the Supreme Court heard *Bray,* one of its defendants bombed several clinics. As Solicitor General, Roberts argued the Bush administration's position in *Bray* before the Court, which, in a 6-3 decision, agreed with much of that position (Barge 2005, Keenan 2005).

After little more than 24 hours, NARAL Pro-Choice America pulled the
ad from the air. By then, it had played 200 times—almost entirely in the small
television markets of Maine and Rhode Island, the home states of senators whom
the organization hoped to sway (Nielsen Media Research 2005).

Participants, Design, and Procedure

Four hundred and thirty-five adult American citizens were recruited from two pools
of participants—one maintained by Survey Sampling International, the other by
a large private university—to participate in a "study of social attitudes."[3] All had
previously identified with either the Democratic or the Republican Party; in this study,
47% identified as Democrats and 48% as Republicans. 48% were women, 46% had
graduated from a four-year college, and the median age was 46. At the beginning
of the experiment, all read a passage explaining what the Supreme Court is, that
John Roberts had been nominated to fill a vacancy, that Roberts had argued the Bush
administration's position in *Bray*, and that six Justices had sided with him. At this
point, subjects were randomly assigned to the control or treatment group. The success
of randomization was gauged by testing it against the subjects' self-reported party
identification; using a chi-square test, the null hypothesis of independence cannot be
rejected ($\chi^2_2 = .42, p = .81$).

Treatment-group subjects were then asked to read a transcript of the ad. After
answering a series of unrelated questions, they were informed that NARAL had
withdrawn the ad under criticism. They also read specific criticisms from Walter
Dellinger, identified as "an ally of the group that aired the ad and an important attorney
in the Clinton administration," and Arlen Specter, identified as "a Republican senator

---

[3] Subjects were also told that "no knowledge of politics or interest in politics is necessary," and the
open-ended comments that many made at the end of the experiment provide no indication that they were
unusually knowledgeable or interested.

and supporter of abortion rights." These criticisms introduced no new material information about the ad. Instead, they repeated the information that had been provided at the beginning of the experiment.

All subjects were asked to report their attitude toward Roberts "as a Supreme Court Justice" on a seven-point scale ranging from "disapprove strongly" to "neither approve nor disapprove" to "approve strongly." Control-group subjects were asked these questions after reading the introductory text; treatment-group subjects were asked after reading the transcript of the ad and after the discrediting. Subjects were also asked to report how sure they were of their position on a five-point scale ranging from "not at all sure" to "extremely sure" and to report their party identification ("Democrat," "Republican," "Neither").[4]

After finishing the experiment, each subject was told its purpose. Belief perseverance was described, and subjects were given links to a Factcheck.org criticism of the NARAL ad (Barge 2005) and to the NARAL rebuttal of that criticism (Keenan 2005).

Results

What should treated subjects think of John Roberts? Immediately after they read the ad transcript, 50% of them disapproved of Roberts, against only 33% of control-group

---

[4] It is not known whether asking the attitude question immediately after the ad was shown made treated subjects less responsive to the subsequent discrediting. Note, though, that a series of distractor questions intervened between the first attitude question and the discrediting. To the extent that "stickiness" was at work, its effect should have been weakened by those questions. Moreover, a stickiness-based account of perseverance cannot explain the finding of perseverance in studies in which the question of interest was asked only once of each subject (e.g., Ross, Lepper, and Hubbard 1975). Nor, for that matter, can it explain some of the perseverance findings in Experiments 2 and 3 of this paper.

subjects (p = .01, one-tailed).[5] This is unsurprising, for only treated subjects were exposed to the attack on Roberts.

But consider treated subjects' attitudes after they have read the discrediting information. Unlike real-world discreditings, this one cannot be dismissed on partisan grounds. It also eliminates the factual basis for the attack on Roberts. Ideally, it will cause treated subjects' attitudes to revert to their pre-treatment state, in which case they will resemble control-group attitudes. If this does not occur, there is a *perseverance effect*: a difference between the control group and the treatment group after the latter has been exposed to both false information and a discrediting of that information.

The top panel of Figure 2 reveals a perseverance effect for the entire sample: 42% of treated subjects disapproved of Roberts after receiving the discrediting, a rate significantly greater than the 33% in the control group (p = .01). The bottom panels of the figure show that the effect was moderated by partisanship. Initially, the ad increased disapproval among members of both parties: Republican disapproval rates doubled from 11% in the control group to 22% in the treatment group (p < .01); Democratic disapproval rates rose from 56% to 80% (p < .01). But the groups responded quite differently to the subsequent discrediting of the ad. After the discrediting, only 14% of treatment-group Republicans disapproved; the ad's effect among them was statistically indistinguishable from zero (p = .25). But the disapproval rate among Democrats dipped only to 72%, and the difference from the 56% disapproval rate registered by their control-group counterparts was still significant at p = .01. The pattern of results is just what the partisan motivation hypothesis predicts in this case: perseverance among Democrats but not among Republicans.

---

[5] Because there is strong reason to expect differences between the control and the treatment group to take a particular direction, difference-of-means tests in this paper are one-tailed unless otherwise noted.

Percentages Disapproving of John Roberts
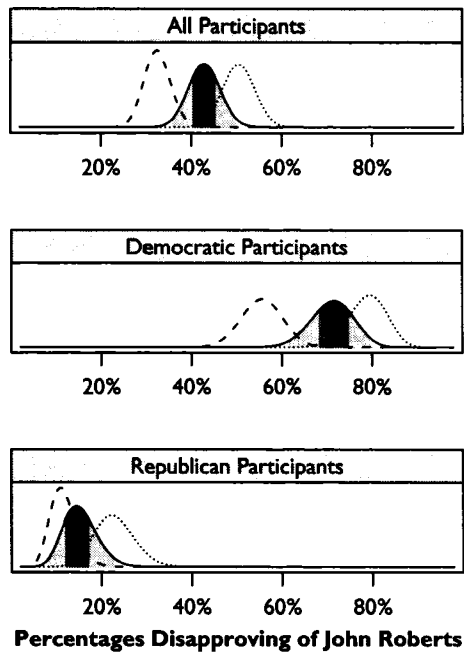
Figure 2.1: Partisanship Moderates Perseverance in the John Roberts Experiment. Each panel depicts estimated densities for mean percentages of subjects disapproving of John Roberts. Dashed lines represent control-group subjects; dotted lines represent treated subjects immediately after reading a transcript after the ad; solid lines represent treated subjects after they receive the discrediting. Light grey areas are 95% highest density regions; dark grey areas are 50% HDRs.

The top panel presents results from all participants and shows a perseverance effect: 42% of treated subjects disapproved of Roberts after exposure to the ad and the discrediting, against only 33% of control-group subjects (who were exposed to neither the ad nor the discrediting). The bottom panels reveal that aggregating over parties masks substantial differences between partisans. Among Democratic participants, the perseverance effect was very large (control = 56%, treatment post-discrediting = 72%, $p = .01$). Among Republicans, it was quite weak (control = 11%, treatment post-discrediting = 14%, $p = .25$).

After all, John Roberts was a Republican, a prominent official in a previous Republican administration, and a Bush nominee. All of those were reasons for Democrats to think badly of him even after learning that the ad was false; Republicans, by contrast, had no obvious reason to think poorly of Roberts after the ad was refuted.

The moderating effect of partisanship is affirmed by the ordered logistic regression model reported in Table 2.1, which also accounts for the polychotomous nature of the approval rating. Estimates from the model, which posits that the approval

rating is a function of the treatment, party ID, and their interaction, reveal that the difference between the Republican and Democratic treatment effects is significant ($p = .09$) and quite substantial. For Democrats, the treatment increased the estimated probability of disapproving by 14%; it decreased the probabilities of remaining neutral and approving by 6% and 8%. For Republicans, all of these quantities changed by less than 1%. The extent to which the treatment effect differs by party can be seen by comparing the bottom two panels of Figure 2, in which the disapproval gap between Democratic and Republican subjects grows from 45% in the control condition to 58% in the treatment condition.

A similar phenomenon is at work in subjects' reports of the confidence they have in their ratings of Roberts. After the ad is discredited, treated subjects have no factual basis for a different level of confidence. Still, we might expect that the experience of having information about Roberts provided and then discredited would shake the confidence that treated subjects reposed in their assessment of Roberts. In reality, it had the opposite effect: treated subjects were *more* sure of their attitude. There are no apparent partisan differences (nor any obvious reason to expect them), but there is a plain treatment effect. 75% of treated subjects were "very sure" or "extremely sure" of their rating of Roberts, against only 64% of control-group subjects ($p = .02$). This is not quite a belief perseverance effect: at issue is not a belief but the confidence with which the belief is held. But it is disconcerting, especially when coupled with evidence about perseverance in subjects' ratings of Roberts. We have already seen that treated subjects differ from the control group even when the advertisement's insinuations are discredited. This finding would be less disturbing if treated subjects were at least less confident of their more disapproving views. But they are more confident of their views, rather than less.

Treatment	−.07	.26
Democrat	−2.42*	.26
Independent	−.89	.55
Treatment × Democrat	−.48+	.35
Treatment × Independent	−1.12	.78
τdisapprove strongly | disapprove somewhat	−3.77	.25
τdisapprove somewhat | disapprove slightly	−2.89	.23
τdisapprove slightly | neither approve nor disapprove	−2.19	.21
τneither approve nor disapprove | approve slightly	−1.01	.19
τapprove slightly | approve somewhat	−.42	.18
τapprove somewhat | approve strongly	.25	.18
Log likelihood	−729.42	
Likelihood ratio test	238.28	<.001
Number of observations	435	

Table 2.1: Belief Perseverance in the John Roberts Experiment. Cell entries in the top rows are parameter estimates and standard errors from an ordered logistic regression. Entries in the "likelihood ratio test" row are the χ^2 statistic and corresponding p-level from a test against a model with no predictors. The dependent variable is approval of the John Roberts as a Supreme Court Justice, a seven-category variable ranging from "disapprove" to "neither disapprove nor approve" to "approve." The treatment, described in the text, consists of reading the transcript of a television advertisement critical of John Roberts and later learning about the ad's retraction following sharp bipartisan criticism of its factual merits. Interesting coefficients statistically significant at 95% using a one-tailed test for $H_A>0$ are denoted by *; interesting coefficients significant at 90% are denoted by +.

The results suggest substantial belief perseverance among Democrats: treated Democrats disapproved of Roberts more than their control-group counterparts, even though the discrediting left them no informational basis for doing so. The difference between the treatment effects for Democrats and Republicans is also substantial, indicating that partisanship moderated the extent of perseverance.

Treated subjects were affected by false claims about John Roberts even after exposure to a discrediting that was powerful by real-world standards. Belief perseverance is one explanation: they accepted the discrediting but were still affected by the false claims. But consider three alternatives. Subjects may have doubted the discrediting because it did not come from the organization that produced the ad. To see the difference that this might have made, we need only imagine that the discrediting had been issued not by Dellinger and Specter but by the president of NARAL herself.

(In this vein, compare the current experiment to the early perseverance experiments, in which the very experimenters who provided new information to subjects later explained that they had fabricated it.) A second alternative lies with priming: although all subjects were told about Roberts' role in *Bray* and that his nomination had been opposed by NARAL, only treated subjects read the advertisement, which may have made their views on abortion a greater influence on their attitude toward Roberts. A third alternative is that NARAL seemed endogenously trustworthy to the treated subjects: even if its ad made false claims, it was a costly signal from an organization that shared subjects' values, and thus a good reason to disapprove of Roberts. By these latter accounts, treated subjects may have accepted the discrediting, altogether stopped believing the false claims, and still had reason to think worse of Roberts than their control-group counterparts. Experiments 2 and 3 were designed to eliminate all three alternative explanations.[6]

Experiment 5: *Newsweek* and the Abuse of Prisoners at Guantánamo Bay

The United States maintains a naval base at Guantánamo Bay, Cuba. On May 1, 2005, *Newsweek* published a two-paragraph article detailing abuses by U.S. interrogators

---

[6] A fourth alternative explanation is that subjects generally put great stock in the views of Dellinger and Specter, and that treated subjects differed from their control-group counterparts because they alone learned that Dellinger and Specter disliked the ad. But this is unlikely. Outside the world of political elites, Dellinger and Specter are almost entirely unknown, and their statements only repeated information that was given to all subjects at the beginning of the experiment. Subjects were not asked whether they recognized either man's name, but consider data from the 2000 ANES, in which respondents were asked to identify four much more famous political figures: Tony Blair, Trent Lott, William Rehnquist, and Janet Reno. Lott was the least well-known of the four; only 9% could name his office, which was Senate Majority Leader at the time of the survey. And Lott was surely many times better-known than Specter or Dellinger.

of Islamic prisoners held at the base.[7] Six words in the article made the charge that interrogators "flushed a Qur'an down a toilet" (Isikoff and Barry 2005). Eleven days later, the claim was publicized in Afghanistan and Pakistan by Arab media outlets. It triggered several days of anti-American rioting in which at least 15 people died and 60 were injured (Kurtz 2005).

Pentagon spokesmen denied the charge. And on May 16, *Newsweek* editor Mark Whitaker issued a retraction: "Based on what we now know, we are retracting our original story that an internal military investigation had uncovered Qur'an abuse at Guantánamo Bay."

Participants, Design, and Procedure

Two hundred and eighty-eight adult American citizens were recruited from a pool of participants maintained by a large private university. 60% identified as Democrats, 20% as Republican, and 20% as independent. 74% were women, and 51% had graduated from a four-year college. The median age was 33. At the beginning of the experiment, all were informed that they would soon read an article about the naval base. They also read two passages containing information that they were told they might need to understand the article. The first was a five-sentence description of the naval base. The second was a description of the Qur'an:

> The *Qur'an* (sometimes spelled *Koran*) is the sacred book of Islam.
> According to Islamic belief, the Qur'an was revealed by God to the
> Prophet Muhammad. Many Muslims believe that it is the literal
> word of God.

The passage was designed to approximate the information about the Qur'an that was provided in news articles about the controversy. (Most provided at least as much information, and some provided more: e.g., Williams and Khan 2005 and

---

[7] Although the issue was published on May 1, it was dated May 9 (Kurtz 2005).

Savage 2005.) It was also designed to increase comparability between the control and treatment groups. As we shall see, the treatment probably primed treated subjects to think of the Qur'an when answering questions that appeared later in the experiment; this passage was administered to all to ensure that control-group subjects were also primed.

Each subject was then randomly assigned to read a version of the *Newsweek* article. Control-group subjects read a version identical to the one that appeared in print, save for the omission of the claim about the Qur'an and one other sentence. Treatment-group subjects read a version that included the claim about the Qur'an but was otherwise identical to the control-group version. (See Appendix E.) One hundred and forty-four subjects were assigned to the control group while another 144 were assigned to the treatment. The success of randomization was gauged by testing it against the subjects' self-reported party identification; using a chi-square test, the null hypothesis of independence cannot be rejected ($\chi^2_2 = 1.44, p = .49$).

After completing a series of unrelated tasks lasting between 8 and 16 minutes, treatment-group subjects learned that *Newsweek* had retracted part of the article. They then read the full text of the *Newsweek* retraction.

All subjects were asked to state whether they approved, disapproved, or neither approved nor disapproved of the treatment of detainees at Guantánamo Bay. Control subjects were asked after reading the article; treatment subjects, after reading the article and after the discrediting. At the end of the experiment, all subjects were also asked to state whether they thought Congress should investigate the reported abuse of prisoners (to which the possible responses were "Yes," "No," and "Unsure") and their party identification ("Democrat," "Republican," "Neither").

At the end, subjects were debriefed as in the previous experiment. Subjects were given links to the actual *Newsweek* article (including the retraction) and to a *Newsweek* article about the controversy.

Results

Reading that a Qur'an had been flushed down a toilet had a moderate initial effect on the treatment group: immediately after reading the article, 60% of treatment-group subjects disapproved of the treatment of detainees, against 51% of control-group subjects ($p = .08$). After reading the retraction, 51% of treated subjects disapproved—the same percentage as in the control group. There seems to be no perseverance effect. But the bottom panels of Figure 2 show that the overall results mask a strong perseverance effect that is conditional on partisanship. Consider first the Democratic subjects: 56% of those in the control group disapproved of detainee treatment, and immediately after reading the article, 78% of treatment-group Democrats did the same. The retraction mollified the treatment-group Democrats, but not nearly as much as the false claim heightened their disapproval in the first place: it lowered their disapproval rate to 68%, still 12 points greater than that of the control group ($p = .05$). On this issue, treated Democrats exhibited a perseverance effect.

Treated Republicans did, too—in the opposite direction. Immediately after reading the claim about the Qur'an, 19% of them disapproved—a rate substantially lower than the 36% disapproval in the control group ($p = .09$). Unexpectedly, the retraction moved treatment-group subjects even further in the same direction: after exposure to it, only 8% of treated Republicans disapproved of detainee treatment. The difference from their previous disapproval rate is $p = .12$. More importantly, the difference from the control-group disapproval rate is significant at $p = .01$. Reading that U.S. interrogators flushed a Qur'an down a toilet made Republican subjects *less*
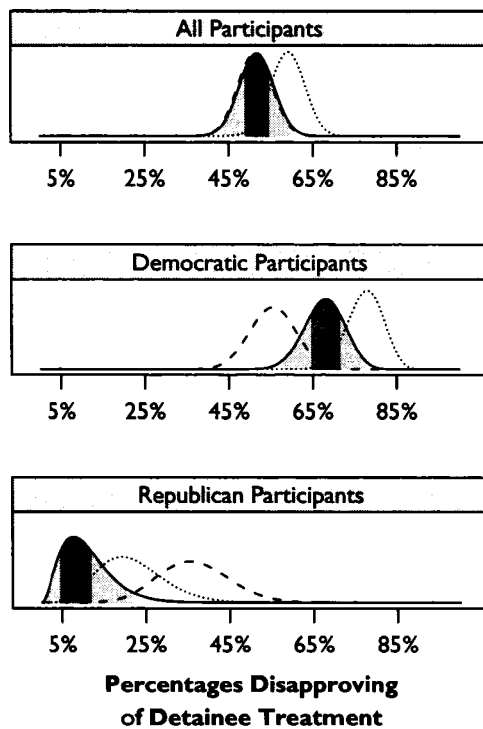
Figure 2.2: Partisanship Moderates Perseverance in the Guantánamo Bay Experiment. Each panel depicts estimated densities for mean percentages of subjects disapproving of the treatment of detainees. Dashed lines represent control-group subjects; dotted lines represent treated subjects immediately after reading a transcript after the ad; solid lines represent treated subjects after they receive the discrediting. Light grey areas are 95% highest density regions; dark grey areas are 50% HDRs.

The dashed and solid lines in the top panel overlap almost perfectly, suggesting no perseverance effect. But the bottom panels reveal that the effect was large. Treated Democrats and Republicans were moved in opposite directions by reading the claim about the Qur'an, and reading *Newsweek's* retraction did not eliminate that effect. (Indeed, it may have heightened the effect among Republicans.) After the retraction, the difference between control- and treatment-group attitudes was significant at $p = .05$ for Democrats, $.01$ for Republicans.

disapproving of the treatment of detainees, and the retraction of that claim did nothing to eliminate its effect. Indeed, it may have heightened the effect—a result that calls to mind Ansolabehere and Iyengar's (1995, 137-42) finding that criticisms of political ads by network news outlets only heightened the effects of the ads.

The reason for the apparent absence of perseverance in the top panel of Figure 2 is now clear. We see no overall effect because the treatment caused two effects—one

Treatment	1.24*	.50
Democrat	−.81*	.40
Independent	−.84	.48
Treatment × Democrat	−1.74*	.59
Treatment × Independent	−.47	.70
Tdisapprove | neither approve nor disapprove	−.58	.34
Tneither approve nor disapprove | approve	1.39	.36
Log likelihood	−263.42	
Likelihood ratio test	58.46	<.001
Number of observations	288	

Table 2.2: Belief Perseverance in the Guantánamo Bay Experiment. Cell entries in the top rows are parameter estimates and standard errors from an ordered logistic regression. Entries in the "likelihood ratio test" row are the χ^2 statistic and corresponding p-level from a test against a model with no predictors. The dependent variable is approval of the treatment of detainees, ranging from "disapprove" to "neither disapprove nor approve" to "approve." The treatment, described in the text, consists of reading about the abuse of the Qur'an at Guantánamo Bay and then learning about the retraction of that claim. Interesting coefficients statistically significant at 95% using a one-sided test for H_A>0 are denoted by *; interesting coefficients significant at 90% are denoted by ⁺.

The results suggest substantial belief perseverance among Democrats and Republicans: treated partisans differed from their control-group counterparts, even though the treatment gave them no informational basis for doing so. The difference between the treatment effects for the two groups is also significant, indicating that partisanship moderated the extent of perseverance.

for Republicans, another for Democrats—in opposite directions. Averaged together, they cancel each other out.

As in the discussion of the previous experiment, the moderating effect of partisanship is affirmed by an ordered logistic regression models in which approval ratings are regressed on the treatment, party ID, and the interaction of those terms. Estimates from the model, reported in Table 2.2, account for the trichotomous nature of the approval rating and thereby offer finer tests of perseverance effects. They reveal that the marginal effect of the treatment is significant for Democrats

(p = .06) and Republicans (p = .01) and that the effect differs by party (p = .002). For Democrats, the treatment increased the estimated probability of disapproving by 11%; the probabilities of remaining neutral and approving declined by 7% and 4%. For Republicans, the treatment caused estimated probabilities of disapproving or taking a neutral stance to fall by 22% and 6%; the probability of approving rose 28%. The extent to which the treatment effect differs by party can be gauged by comparing the bottom two panels of Figure 2, which show that the gap in disapproval rates between Democrats and Republicans grew from 20% in the control condition to 60% in the treatment condition.

Additional perseverance can be found in subjects' end-of-experiment attitudes toward Congressional hearings about the abuse of detainees. In the control group, 69% of subjects called for hearings. In the treatment group, 80% did. The finding was significant overall (p = .02) and for Democrats (control 78%, treatment 86%, p = .09) but not for Republicans (control 65%, treatment 69%, p = .35). The treatment effect thus seems to be twice as great for Democrats (8% vs. 4%), though the small number of Republicans in the sample renders this difference-of-differences estimate imprecise (95% confidence interval: [−23%, 30%]).

Both hypotheses, then, were again borne out. Even after subjects learned that *Newsweek* disavowed its own claim, professed to understand that disavowal and seemed to accept it, they continued to be affected by the claim. And the direction in which they were affected depended on their party ID. Relative to the previous experiment's results, these ones are difficult to dismiss by reference to priming, endogenous trustworthiness of the source making the false claim, or lesser credibility of the source of the discrediting. Both control- and treatment-group subjects were primed to think of the Qur'an; it is unlikely to have played a larger role in the treatment group's evaluations of detainee treatment. And unlike the previous experiment, the

false claim and the discrediting here are issued by the same source. It is not possible

to say that one is more credible than the other or that *Newsweek*'s initial claim about

the Qur'an was a costly signal that should be heeded in spite of its retraction. Still,

consider an important qualification: in light of widely publicized criticisms from

the Pentagon and the White House, treated subjects might have concluded that the

magazine was pressured into issuing an insincere retraction. That would prevent us

from inferring either that they were affected by information that they knew to be false

or that they were unreasonably resisting the discrediting. What happens when subjects

cannot reasonably draw such a conclusion?

Experiment 6: Candidate Evaluation in a U.S. Senate Election

Three hundred and twelve adult American citizens were recruited from two pools of

participants—one maintained by Survey Sampling International, the other by a large

private university. 53% identified as Democrats, 32% as Republicans, and 15% as

independents. 64% were women, and 58% had graduated from a four-year college.

The median age was 34. Upon arriving at the experiment's web site, subjects were told

that much of the study was about "voter guides" published by influential newspapers

before major elections. They were led to believe that they would read about a randomly

selected candidate in a 2004 election for the U.S. Senate. In fact, all were presented

with a voter guide about a fictional Republican candidate for an open Senate seat

in Wyoming. It contained biographical information and the candidate's stances on

political issues. It was formatted to resemble an authentic guide, and the information in

it was adapted (and occasionally taken verbatim) from the positions of real candidates that were published in real voter guides.[8]

Control-group subjects received the short version of the guide, which presented the candidate's centrist views on taxes and economic growth, the Patriot Act, immigration, and payment of dues to the United Nations. Treatment-group subjects received the long version, which added information about other stances: a call to eliminate the Department of Education and most federal involvement in education, and a call to create a market in pollution complete with commercial trade in "pollution credits." These stances, typically associated with dedicated conservatives, were included because they are relatively unpopular even among Republicans. But they are by no means unheard-of: the former stance was part of the national Republican Party platform until 2000 (American Presidency Project 2006), and the latter is a hallmark of the Bush Administration's "Clear Skies" initiative (White House 2002).

One hundred and sixty subjects were assigned to the control group; 152 were assigned to the treatment. The success of randomization was gauged by testing it against the subjects' self-reported party identification; using a chi-square test, the null hypothesis of independence cannot be rejected ($\chi_2^2 = 1.88, p = .39$).

After completing a series of unrelated tasks, treatment-group subjects were debriefed about the information in the voter guide. They were told that the study was actually about their responses to different politicians' stances on education and the environment and that it had therefore been necessary to mislead them by inserting fabricated stances on those issues into an otherwise accurate guide. They were also told that the experimenters did not really know the candidate's stances on those issues. (See Appendix G.)

---

[8] To ensure realism, the guides used in this experiment followed the template of the Voter Guide Toolkit (http://www.vgt2004.org), which was used by more than a dozen high-circulation newspapers during the run-up to the 2004 elections.

All subjects were asked to state whether they approved, disapproved, or neither disapproved nor approved of the candidate. Control-group subjects were asked after reading the voter guide; treated subjects, after reading the guide and after the debriefing. In addition, all subjects answered the party ID and attitude strength questions described in the discussion of Experiment 4.

Results

Consider the predicament of treatment-group subjects just after they read the voter guide. They alone were exposed to the candidate's unpopular stances on education and pollution credit trading. We might expect that they would think worse of the candidate at this point in the experiment, and they did: 52% of them initially disapproved of the candidate, against only 19% of control-group subjects ($p < .01$).

But now consider their attitudes after they read the retraction. It was unambiguous and it could not have come from a more authoritative source. Even so, the top panel of Figure 2.3 reveals a sharp difference between the control and treatment groups. At the end of the experiment, 31% of treated subjects disapproved of the candidate. The difference from the 19% disapproval rating in the control group is significant at $p = .01$: treated subjects continued to be affected by the false information.

Figure 2.3 shows that members of both parties thought worse of the candidate immediately after reading his stances on education and the environment. But only Democrats had an obvious prior reason to think poorly of the candidate, and the motivation hypothesis therefore suggests that Democrats will be more affected by the false information after receiving the discrediting. That is what the bottom panels of Figure 2.3 reveal. At the end of the experiment, treated Republicans were moderately affected by the false information: 20% disapproved of the candidate, against only 12%

All Participants

20%   40%   60%

Democratic Participants

20%   40%   60%

Republican Participants

20%   40%   60%

**Percentages Disapproving
of Senate Candidate**
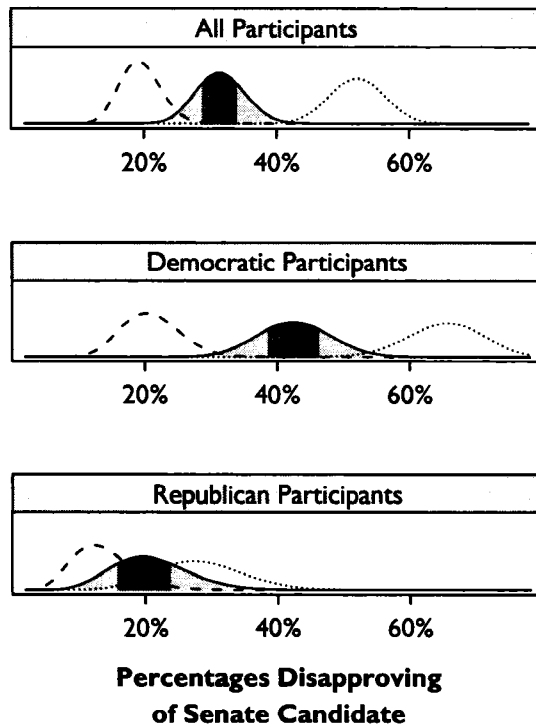
Figure 2.3: Partisanship Moderates Perseverance in the Candidate Evaluation Experiment. Each panel depicts estimated densities for mean percentages of subjects disapproving of the candidate. Dashed lines represent control-group subjects; dotted lines represent treated subjects immediately after reading the voter guide; solid lines represent treated subjects after reading both the voter guide and the discrediting. Light grey areas are 95% highest density regions; dark grey areas are 50% HDRs.

The top panel suggests a moderate overall perseverance effect. But in the by-now-familiar pattern, the overall results obscure sharp partisan differences. As in Experiment 4, the bottom panels reveal that the effect is very strong among Democrats but modest for Republicans.

of their control-group counterparts ($p = .16$). But Democratic subjects exhibited a perseverance effect more than twice as strong. In the control condition, 20% of Democratic subjects disapproved; after the discrediting, 42% of treatment-group Democrats did the same ($p < .01$).

As in the previous experiments, the results are affirmed by an ordered logistic regression model in which the approval rating is regressed on the treatment, party ID, and the interaction of those terms. The model estimates, reported in Table 2.3, show that partisanship moderates belief perseverance even when the trichotomous nature of the dependent variable is taken into account. For Republicans, the treatment was statistically significant ($p = .05$) but modest. It lowered their estimated probability of approval by 15%, increasing their estimated probability of disapproval by 8% and

Treatment	$-.67^*$	.40
Democrat	$-.31$	.37
Independent	$-.78$	.55
Treatment × Democrat	$-.95^*$	.51
Treatment × Independent	.06	.71
$\tau_{\text{disapprove | neither}}$	-2.09	.31
$\tau_{\text{neither | approve}}$	$-.83$	.29
Log likelihood	-298.64	
Likelihood ratio test	58.38	<.001
Number of observations	308	

Table 2.3: Belief Perseverance in the Candidate Evaluation Experiment. Cell entries in the top rows are parameter estimates and standard errors from an ordered logistic regression. Entries in the "likelihood ratio test" row are the χ^2 statistic and corresponding p-level from a test against a model with no predictors. The dependent variable is approval of the U.S. Senate candidate, ranging from "disapprove" to "neither disapprove nor approve" to "approve." The treatment, described in the text, consists of reading a voter guide about the candidate and later learning that some information in the guide was false. Interesting coefficients statistically significant at 95% using a one-sided test for $H_A>0$ are denoted by *. Four subjects did not answer the candidate approval question; they have been omitted from this analysis

The results suggest that treated Democrats disapproved of Roberts more than their control-group counterparts, even though the treatment left them no informational basis for doing so. The difference between the treatment effects for Democrats and Republicans is also substantial, indicating that partisanship moderated the extent of perseverance.

of taking a neutral stance by 7%. For Democrats, though, the marginal effect of the treatment was substantial (and $p < .01$). The treatment decreased their estimated probability of approval by 37%, increasing the probabilities of approving and taking a neutral stance by 31% and 6%. The treatment effect was far greater for Democrats ($p = .03$); the magnitude of the difference can be judged by looking to the bottom two panels of Figure 2.3, in which the gap between percentages of Democratic and Republican subjects disapproving of the candidate expands from 31% in the control group to 42% in the treatment group.

As in the John Roberts experiment, something like belief perseverance was at work in the confidence that subjects repose in their ratings of the candidate. Once again, the treatment provided no obvious basis for a greater level of confidence in the measured attitude. But once again, treated subjects were more confident than control-group subjects at the end of the experiment: 26% of control-group subjects reported that they were "very sure" or "extremely sure" of their attitude toward the candidate, against 32% of the treated subjects. At $p = .12$, the difference is quite unlikely to have occurred by chance.

Pause a moment to consider these findings. Treated subjects knew nothing about the candidate before the experiment began. They apparently believed, and certainly were affected by, the candidate's fabricated stances on two issues. They then read an unambiguous discrediting of that information. But they either put more stock in the fabricated stances than in the discrediting or—more likely—accepted the discrediting but continued to be affected by the made-up information. These results cannot be explained by reference to what subjects knew before the experiment, for subjects could not have known about the candidate or the retraction. (In this respect, the current experiment is distinct from the previous two.) The retraction was not muddled; if anything, it was less ambiguous than the false information that had previously been provided. Its source was not obviously partisan and, more importantly, it had been believed at an earlier point in the experiment. In short, it seems difficult to explain the results by reference to anything like a reasonable dismissal of the discrediting. Of course, subjects may have been unsettled by the unusual experience of having an experimenter provide and then disavow information (though the confidence finding just reported suggests otherwise); in this case, they may have responded by choosing an attitude between the one they had before reading the false information and the one they had immediately after reading it. Perhaps much political deception works

this way: unsure of how to arbitrate between a false claim and its discrediting, citizens respond by choosing an intermediate view. This may seem benign, but it is insidious: it is another way that false claims can withstand evidentiary discreditings.

Discussion

> It never pays for our government to give false impressions to the
> American public with a view to enlisting its support for short-term
> purposes, because this always revenges itself later when it becomes
> necessary to overcome the wrong impression one has created
> (Kennan 1997, 38).

Kennan's conclusion is beyond the purview of this article. But his premise—that it is difficult in politics to defeat wrong impressions—is not. The experiments described here suggest that his premise is correct. We already knew that prior beliefs can be difficult to change under ordinary circumstances (McGuire 1964; Abelson 1986). What these experiments add is an understanding that even under very favorable circumstances—perhaps even when a person fully accepts that his belief is false—related attitudes will not always change accordingly. False beliefs are not just hard to kill. They have an afterlife, too.

These experiments also show, for the first time, that the extent of belief perseverance depends on differences between individuals. In doing so, they offer strong support for the idea that motivated reasoning underpins perseverance. Other explanations can account for the finding of perseverance, but no theory that fails to account for motivation can explain the partisan differences that are manifest here.

One virtue of experiments is that they permit decisive challenges to beliefs that are relatively weak. The beliefs are weak because they have only been instilled at an earlier point in the experiment itself. The challenges are strong because they come

from authoritative sources who cannot easily be dismissed as biased or uninformed. But in ordinary political life, beliefs are often stronger than the ones instilled in these experiments, and challenges to beliefs are weaker because they come from sources who can be dismissed as biased or uninformed. In this sense, the experiments described here are conservative: if there are any political situations in which factual discreditings should render false information powerless, they are the situations created by these experiments. But the effects of false information survived the discreditings here, which suggests that they are likely to do the same in the outside world, and on a greater scale.

There is a second reason why perseverance in the outside world is likely to be stronger than these results suggest. Even when the political world does produce decisive challenges to particular beliefs, selective exposure to congenial news sources is likely to shield people from them (Ross, Lepper, and Hubbard 1975, 891; Taber and Lodge 2006; Gentzkow and Shapiro 2006). This view was challenged in the past by those who argued that people do not consciously seek to reinforce their views through selective exposure (Sears and Freedman 1967; Frey 1986). But selective exposure does not require a reinforcement motive (Katz 1968), and there is reason to believe that heightened sorting of the electorate (Levendusky 2006) and the splintering of the market for news into specialized niches (Prior 2007, Chapter 4) have made it increasingly common (though see Webster 2005).

Future research on this topic may take many paths, but two deserve special attention. The partisan cleavages that appear in each experiment suggest that false claims have enduring effects because they heighten the salience of partisan concepts in people's minds. Future efforts should explore this: for example, do negative claims about a candidate's issue stances prime specific memories of related stances by others in his party, or do they simply promote general negative affect about the party? Future

efforts should also consider strategies for both would-be deceivers and their targets. It may not be optimal for politicians to deceive: once exposed, deception may redound to their discredit even as it continues to affect the public's views. (See Callander and Wilkie 2007.) Perhaps using deceptive surrogates permits politicians to enjoy the benefits of deception without suffering the reputation costs—but this remains to be shown. On the other side of the coin, victims of deceptive attacks might benefit most by heightening legal penalties for deceptive political speech. U.S. courts have been uniformly unwilling to allow this, but their arguments rest heavily on the notion that freedom of speech is a truth-promoting mechanism (*Whitney* v. *California* 1927; *Gertz* v. *Schmidt* 1974, 339-40)—a notion that this article calls into question.

Benign Explanations?

In recent years, the idea that manipulation is a serious problem for democratic polities has been challenged on many fronts. Lupia and McCubbins (1998) argue that manipulation of voters through deception is more difficult than commonly supposed. Druckman and Nelson (2003) and Sniderman and Theriault (2004) argue a similar point about manipulation through issue framing. Druckman (2004) finds the same for classical ("equivalency") framing. Brader (2006, 190-98) cedes the potential danger of manipulation through "fear appeals" in advertisements but suggests that these dangers may be offset by advantages, e.g., the ability of such appeals to increase participation. And Lupia (2006) argues that politicians' ability to manipulate citizens through such appeals may often be weaker than we imagine. To be sure, none of this work suggests that manipulation cannot occur. But all of it suggests that either the ease of manipulation or its normative importance has been overstated. The results presented here seem to tend in the opposite direction—to suggest that citizens' views

on basic political matters can be manipulated by false information in a way that resists correction. Is this right? Or is a more benign interpretation in order?

At a glance, one benign explanation may seem especially attractive: Republicans and Democrats in these experiments differ simply because they have different values and thus prefer different policies. Unfortunately, this is not a cogent explanation. The evidence of perseverance lies not in the differences between Republicans and Democrats but in the the differences between control- and treatment-group subjects. At the end of the experiments, treated Republicans differed from control-group Republicans, even though the two groups had, by virtue of random assignment, approximately identical preference distributions. The same, of course, was true of Democratic participants. Something else must explain why treated subjects differed from their control-group counterparts.

Another explanation is that the findings reported here are the result of an order effect: false information has an enduring effect in these experiments because it is presented first; had true information been presented first, its effects would have survived false attempts to discredit it. This may be true, but as an explanation, it falters conceptually and normatively. Conceptually, order effects are about adding to one's store of data; belief perseverance is about subtracting from it—about challenging the validity of one's initial data rather than augmenting it. Normatively, the order effect explanation does not obviate the importance of the results. The ability of old information to color the interpretation of new information is not at issue. What is important is that old false information, once believed, can affect people's views in spite of subsequent efforts to discredit it that are extremely powerful by political standards.

The most interesting attempt at a benign interpretation starts with the mechanisms of motivated reasoning. Those mechanisms suggest, consistent with these experiments, that false negative information caused treated subjects to summon

related beliefs and attitudes from long-term memory. When the false information was discredited, the retrieved memories still seemed to justify their more negative attitudes. In short, they used other, remembered information to justify their attitudes. What is wrong with bringing more information to bear, rather than less, when justifying a stance toward a politician or a policy?

All else equal, nothing is wrong with it. Priming and remembering are not inherently pernicious. Nor is the instinctive tendency to explain what we observe. But when these phenomena are activated by false information and directed by the desire to shore up our partisan convictions, something *is* amiss. In that case, the stock of considerations that partisans draw on to justify their attitudes is nothing like the one that would result from evenhanded contemplation. Nor is it like the one that they would have if they had never been exposed to the false information at all. False information succeeds in manipulating partisans' views by causing them to draw on a stock of considerations different from, and more biased than, the one that they would otherwise use. And this manner of manipulation is robust because it succeeds even if the false information is discredited. Slanderous advertisements, for example, may achieve their purposes even if they are quickly exposed as false. And the safeguards of political competition and a free speech marketplace may be less helpful than we imagine. Reconciling these facts with a benign explanation is not at all easy.

Appendix E: Guantánamo Bay Experiment

All subjects read a version of this article:

> Investigators probing interrogation abuses at the U.S. detention center at
> Guantánamo Bay have confirmed some infractions alleged in internal FBI
> e-mails that surfaced late last year. Among the previously unreported cases,
> sources tell *Newsweek*: interrogators, in an attempt to rattle suspects, [flushed
> a Qur'an down a toilet and] led a detainee around with a collar and dog leash.
> An Army spokesman confirms that 10 Gitmo interrogators have already been
> disciplined for mistreating prisoners. (New details of abuse are also in a new
> book to be published this week by a former Gitmo translator.)
>
> These findings, expected in an upcoming report by the U.S. Southern
> Command in Miami, could put former Gitmo commander Maj. Gen. Geoffrey
> Miller in the hot seat. Two months ago a more senior general, Air Force Lt.
> Gen. Randall Schmidt, was placed in charge of the SouthCom probe, in part, so
> Miller could be questioned. The FBI e-mails indicate that FBI agents quarreled
> repeatedly with military commanders, including Miller and his predecessor,
> retired Gen. Michael Dunleavy, over the military's more aggressive techniques.
> "Both agreed the bureau has their way of doing business and DOD has their
> marching orders from the SecDef," one e-mail stated, referring to Secretary
> of Defense Donald Rumsfeld. Sources familiar with the SouthCom probe say
> investigators didn't find that Miller authorized abusive treatment. But given
> the complaints that were being raised, sources say, the report will provoke
> questions about whether Miller should have known what was happening—and
> acted to try to prevent it. An Army spokesman declined to comment.

Treatment-group subjects read the article without the text in brackets. Treatment-group
subjects read the article with the text in brackets.

This version is very similar to the original *Newsweek* article, which can be found, with retraction, at http://www.msnbc.msn.com/id/7693014/site/newsweek/.

Appendix F: John Roberts Experiment

At the beginning of the experiment, all subjects were told:

> The Supreme Court is the most powerful court in the country. At any time, it has nine members.
>
> From 1994 through 2004, the Supreme Court had a period of unprecedented stability. The same nine justices served on it for that entire time.
>
> But in 2005, two justices left the Supreme Court and two new ones were nominated by President Bush. One of the new nominees was John Roberts.
>
> In 1991, John Roberts appeared before the Supreme Court to argue the Bush administration's position in *Bray v. Alexandria*. He argued that a federal anti-discrimination law couldn't be used against protesters blockading abortion clinics. He won the case: six justices agreed with him.
>
> Roberts' nomination to the Supreme Court was opposed by NARAL Pro-Choice America and several other organizations. Even so, Roberts was confirmed by the Senate in September. 22 Democrats voted against him, but 78 Senators (including another 22 Democrats) voted for him.

Treatment-group subjects then read this transcript of the NARAL Pro-Choice America ad:

> **Speaking Out**
>
> Announcer: Seven years ago, a bomb destroyed a women's health clinic in Birmingham, Alabama.
>
> (Text on screen: New Woman/All Women Health Clinic; January 28, 1998)

(On screen: Pictures of fire trucks. Pictures of the destroyed clinic. Picture of a hospitalized woman in a wheelchair.)

Emily Lyons: When a bomb ripped through my clinic, I almost lost my life. I will never be the same.

Announcer: Supreme Court nominee John Roberts filed court briefs supporting violent fringe groups and a convicted clinic bomber.

(On screen: Video footage of John Roberts. Picture of a legal brief from *Bray v. Alexandria*.)

(Text on screen: Roberts filed court brief supporting clinic protestors)

Emily Lyons: I'm determined to stop this violence so I'm speaking out.

Announcer: Call your Senators. Tell them to oppose John Roberts. America can't afford a Justice whose ideology leads him to excuse violence against other Americans.

Near the end of the experiment, treatment-group subjects received this information:

Recall the ad transcript you read about earlier in this survey. The ad was strongly criticized by many people, some of whom were prominent supporters of abortion rights.

Walter Dellinger, an ally of the group that aired the ad and an important attorney in the Clinton administration, called the ad "unfair and unwarranted." He added that "it is unfair to suggest that John Roberts, in advancing a somewhat narrow interpretation of [the anti-discrimination law], was supporting 'violent fringe groups and a convicted clinic bomber'—as unfair as it would be to suggest that the six Justices who were part of the majority in *Bray* joined a decision supporting violent fringe groups."

Arlen Specter, a Republican senator and supporter of abortion rights, called the ad "blatantly untrue and unfair."

Stung by these criticisms and many others, NARAL Pro-Choice America withdrew the ad from television.

Appendix G: Candidate Evaluation Experiment

All subjects received the following information about the candidate's stances on issues:

Economy. **With the costs of food, gas and health care rising, how can you make your constituents' lives better?** Two simple words: tax relief. You can't create jobs, and you can't create real opportunity, by taxing your way to a stronger economy. I'm a tax cutter. And I'm proud of it. I believe the people of Wyoming should be able to keep more of their hard-earned money and spend it as they see fit.

To be a tax cutter, you've got to have discipline on the other side. You've got to restrain spending. This is common sense. Everyone from Wyoming must live within a budget. When times get tough, you have to adjust your expenses to live within your income.

Patriot Act. **Would you vote to reauthorize the Patriot Act?** I understand the central role that intelligence gathering plays in law enforcement and national security. The Patriot Act increases law enforcement's ability to collect intelligence. But it goes too far in compromising cherished civil rights for uncertain benefits. Therefore, I support the efforts of a bipartisan group of Senators who are promoting the the SAFE (Security and Freedom Ensured) Act. The SAFE Act makes appropriate amendments to the Patriot Act, such as requiring increased judicial review and higher standards of evidence before law enforcement agencies can obtain secret subpoenas to review medical, business and library records.

Immigration. Immigration is one of the most challenging issues we face today. I support the bipartisan initiative to establish a guest worker program

for immigrants working in the U.S. I believe that legal immigration is good
for our country and good for our economy. But we must be tough on illegal
immigration. The issue reaches well beyond the strain on our economy to
matters of homeland security.

United Nations dues. **Should the United States fully pay the membership
dues recently assessed by the United Nations?** Not until the U.N. undertakes
a program of reform. The U.S. already pays more dues than most members of
the U.N.—combined. And yet we continue to hear of corruption and rampant
misuse of our funds. The misuse of the oil-for-food funds is only the latest
example. We need to demand more for our money from the U.N. before we
give more money.

Treatment group subjects also read about the candidate's much less moderate stances:

Education. Decisions about our children's education are too important to be
left to federal bureaucrats. As Senator, I will work to eliminate the Department
of Education, all federal intervention into education, and all federal subsidies of
education, except those that support veterans. My role as a Senator will be to
get decisionmaking about education out of the hands of the federal government
and into the hands of families and neighborhoods.

Environment. Thirty years of competition between undue alarmism and
unthinking skepticism have confused environmental issues in the minds of most
Americans.

The first thing that we have to realize is that property rights and free
markets are essential protectors of a clean, sustainable environment. The

phenomenon was most pronounced in Eastern Europe during the heyday of the Soviet Union, but it is also discernible in America: Government is the biggest polluter and the biggest facilitator of pollution.

If we are going to preserve and redeem our environment, we must begin a commercial trade in "pollution credits"—a quantified, qualified "right to pollute." Pollution, properly understood, is an offense against the property rights of those whom it affects, and should be treated as an actionable tort to be adjudicated by the legal system.

At the end of the experiment, treated subjects read that the candidate's stances on education and the environment were not really his. Instead, they had been fabricated:

At the beginning of this survey, you read a voter guide about a candidate. Not all of the information in that guide is known to be true.

We—researchers at [university name deleted]—made up information about the candidate's stances on education and the environment. We do not really know where he stood on those two issues. Our intention is to see how different positions on those issues affect people's views of different candidates. To do this, we present a lot of true information from a voter guide, adding a few made-up stances on several issues.

In this case, only the candidate's stances on education and the environment were fabricated. All of the other information in the voter guide was accurate, and was presented to you as it initially appeared in print.

Part 3

Bayesian Updating of Political Beliefs: Normative and Descriptive Properties

Social scientists increasingly use Bayes' Theorem as a normative standard of rational political thinking (Bartels 2002; Gerber and Green 1998, 1999; Tetlock 2005; Steenbergen 2002) and as a tool to describe how people actually think (Bartels 1993; Achen 1992, 2002; Husted, Kenny, and Morton 1995; Grynaviski 2006; Lohmann 1994; Neill 2005). Models based on the Theorem are attractive because they offer a way to account for the weight that people place on old beliefs and new influences when revising their political ideas. They are also formal models, and as such, they bring the benefits of mathematical exposition to topics that have usually lacked it (Lupia 2002; see also Luce 1995). But uncertainty remains about the basis of their normative appeal and about whether they can accommodate everyday features of political cognition.

This essay clarifies those matters. After explaining the Theorem's appeal as a standard of rationality, I show that four important features of political thought—increased uncertainty in response to surprising information, selective

89

perception, attitude polarization, and enduring disagreement—are inconsistent with the

most widely-used Bayesian updating model but quite consistent with other Bayesian

models. Bayes' Theorem proves capable of capturing many features of real-world

political thinking. But this flexibility is the Theorem's downside: precisely because it is

consistent with so many different ways of thinking about politics, it is inadequate as a

standard of rationality.

Bayes' Theorem and Bayesian Updating Models

Most of the matters that interest political scientists—public opinion toward candidates,

implications of new policies, the probability of terrorist attacks—can be thought of

as probability distributions. Like most distributions, they have means and variances,

and the task that we set for ourselves is to learn about these parameters. A politician's

ability to manage the economy, for example, may oscillate over time around a fixed

but unknown mean. Learning about politics becomes a matter of learning about

probability distributions—a task to which Bayesian statistics is especially well-suited.

The bedrock of Bayesian statistics is Bayes' Theorem, an equation that relates

conditional and marginal probabilities:

$$p(S|E) = \frac{p(E|S)p(S)}{p(E)} = \frac{p(E|S)p(S)}{\int p(E|S)p(S)\,dS}, \tag{3.1}$$

where S and E are events in a sample space and $p(\cdot)$ is a probability distribution

function. In words, the Theorem indicates that $p(S|E)$, the probability that S occurs

conditional on E having occurred, is a function of the conditional probability $p(E|S)$

and the marginal probabilities $p(E)$ and $p(S)$. Stated thus, the Theorem is merely an

accounting identity. But a change in terminology draws out its significance. This time,

let S be a statement about politics and $p(S)$ be a *belief* about S, i.e., a probability

distribution indicating someone's estimate of the extent to which S is true. E is evidence bearing on the belief. In this version of Bayes' Theorem, the estimated probability of S before observing E is given by $p(S)$; it is often called the *prior probability of S*, or simply the "prior." The estimated probability that S is true after observing E is $p(S|E)$, often called the *posterior probability of S*. And $p(E|S)$ is the likelihood function that one assigns to the evidence; it reflects a person's guess about the probability distribution from which the data are drawn. Understood in this way, the Theorem tells us how to revise any belief after receiving relevant evidence and subjectively estimating its likelihood. It is most often applied to beliefs about future events (Tetlock 2005), but it is fundamentally a tool for calculating probabilities, and it applies with equal force to all ideas that can be described in probabilistic terms. See Figure 3.1; for applications of Bayesian models to political attitudes and evaluations, see Bartels (1993, 2002) and Gerber and Green (1999).[1]

Bayes' Theorem is attractive as a normative standard of belief updating because it can be derived from two fundamental axioms of probability:

$$p(S \cap E) = p(E \cap S) \tag{3.2a}$$

$$p(S|E) = \frac{P(S \cap E)}{p(E)}. \tag{3.2b}$$

---

[1] If it seems confusing to think of attitudes in probabilistic terms, consider that attitudes are merely beliefs that objects are good or bad in some way, often accompanied by affective responses to those objects (Zanna and Rempel 1988; Abelson 1986). If I like John McCain, I have assigned a high probability to the hypothesis that he belongs to a category of objects that I like. And if reviewing new evidence causes me to like McCain less, I assign a lower probability to that hypothesis. This definition of attitude is in keeping with the view that much mental categorization is probabilistic (e.g., Smith and Medin 1981; Smith 1990).

**Median Household Income
in America**

below   above
$50,000 $50,000
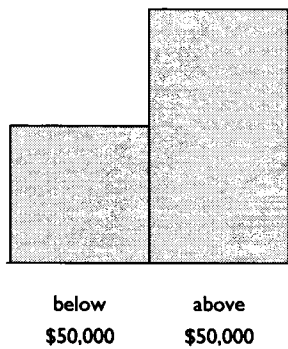
**Probability that Gingrich
Wins the Presidency**

0   .2   .4   .6   .8   I
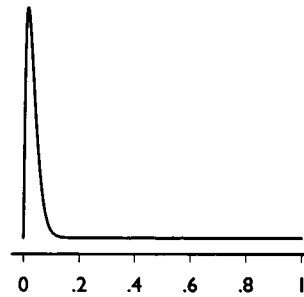
**Attitude toward
John Edwards**

neither   like   like
like nor   a little   a lot
dislike
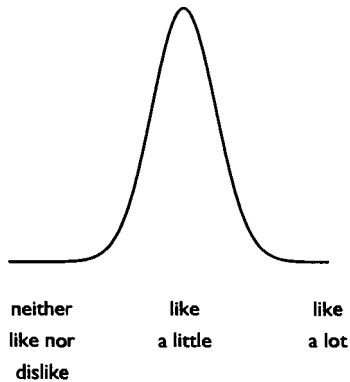
**Bill Clinton as
Steward of the Economy**

incompetent   not very competent   neither   very competent   extremely competent
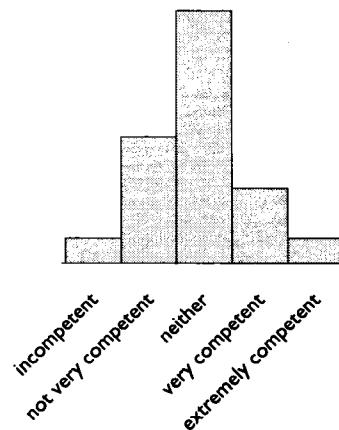
Figure 3.1: Attitudes, Evaluations, and Factual Beliefs Are Probability Distributions. All of political cognition can be conceived in terms of probability distributions, and in doing so we win for political science the vast body of knowledge about subjective probability theory, the chief element of which is Bayes' Theorem. The upper left-hand panel depicts a factual belief about household income in the U.S.: the person holding this belief estimates that there is a 35% chance that the median household income is below $50,000 and a 65% chance that it is greater than that. This is just a two-category discrete probability distribution. The upper right-hand panel depicts a belief about a future matter, Newt Gingrich's chance of winning the Presidency in 2008. The belief is a beta distribution: continuous, asymmetric, and bounded between 0 and 1 (Paolino 2001; Jackman 2008, ch. 2). The lower left-hand panel depicts a positive but somewhat ambivalent attitude about John Edwards: it is a normal distribution. The lower right-hand panel is an ambivalent voter's evaluation of Bill Clinton as a manager of the national economy—a discrete probability distribution with five categories.

Equation 3.2a says that the probability that S and E both occur equals the probability that E and S both occur. Equation 3.2b is a definition of conditional probability.[2] No one consciously rejects either axiom, and following both of them requires that beliefs be updated according to Bayes' Theorem: $p(S|E) = \frac{p(E \cap S)}{p(E)}$, and $p(E \cap S) = p(E|S)p(S)$, so $p(S|E) = \frac{p(E|S)p(S)}{p(E)}$. By contrast, updating that does not correspond to Bayes' Theorem constitutes an implicit rejection of either or both of the axioms.[3]

Because the denominator of Equation 3.1 only serves to ensure that the posterior density integrates to one (and is therefore a proper probability density), Bayes' Theorem is more commonly expressed as

$$p(S|E) \propto p(E|S)p(S).$$

In words, "the posterior belief is proportional to the prior belief times the likelihood." People are Bayesian if their posterior beliefs are determined in this fashion. Importantly, Bayes' Theorem says nothing about what one's priors should be, what evidence one should use to update, or how one should interpret the evidence that one does use—a point to which we shall return.

In political science, one Bayesian updating model is far more common than others: the "normal-normal" model, so-called because it assumes both that people's priors are normally distributed and that they perceive new information to be normally distributed (e.g., Achen 1992; Bartels 1993, 2002; Gerber and Green 1999; Husted, Kenny, and Morton 1995; Gerber and Jackson 1993; Zechman 1979). Suppose that a voter is trying to learn about μ, a politician's level of honesty. Initially, her belief

---

[2] Although conditional probability is usually presented as an axiom, Bernardo and Smith (1994, ch. 2) show that it can be derived from simpler axioms.

[3] This is the simplest valid treatment of a complex topic. For elaborate efforts to root Bayesian statistics in an axiomatic framework, see Savage (1954) and Pratt, Raiffa, and Schlaifer (1964).

about his honesty is normally distributed: $\mu \sim N(\mu_0, \sigma_0^2)$. Later, she encounters a

new message, x, that contains information about his level of honesty. She assumes

that the message is a draw from a distribution with a mean equal to the parameter of

interest.[4] The normal distribution is usually a sensible assumption: if the message can

theoretically assume any real value, and if error or "noise" is likely to be contributed

to it by many minor causes, the central limit theorem suggests that it is likely to be

normal. We write $x \sim N(\mu, \sigma_x^2)$. The variance of this distribution, σ_x^2, captures how

definitive the new information is. If it is communicated directly from a highly credible

source, the signal it sends is clear and the variance is quite small. But if it is merely

a rumor that the voter spots in a tabloid, the signal is only slightly informative and its

variance will be high. Similarly, the variance of a prior or posterior belief is a measure

of the confidence with which it is held: the higher the variance, the less confidence one

places in one's estimate of μ. (See Figure 3.2.)

---

[4] Assuming that the mean of the message distribution is μ is tantamount to assuming that the
message comes from an unbiased source. If the voter believes that the message comes from a biased
source, she needs to adjust for that bias before updating. This is no obstacle to Bayesian updating (e.g.,
Jackman 2005), but it is not part of the normal-normal model.

strong belief   weak belief
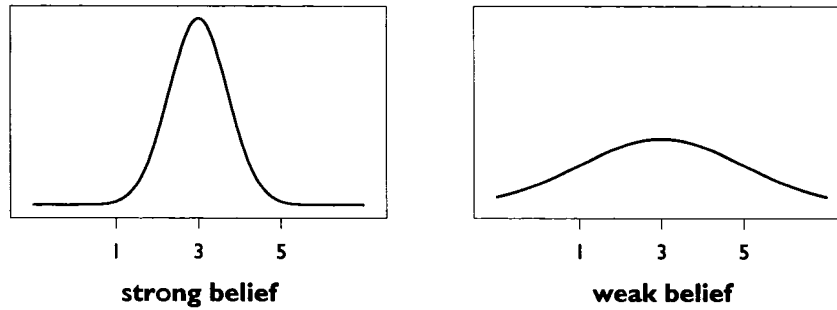
Figure 3.2: The Variance of a Belief Indicates Its Strength. Both panels depict beliefs that are normal distributions with means of 3. The distribution in the left-hand panel has a variance of .25: the person who holds this belief is quite confident that the parameter of interest is about 3. By contrast, the distribution in the right-hand panel has a variance of 4. The person who holds this belief is not confident that the parameter of interest is close to 3; to him, it could easily be around 1 or 5 or some value even more distant from 3.

By a common result (e.g., Box and Tiao 1973), a voter with a normal prior belief who updates according to Bayes' Theorem in response to x will have posterior belief $\mu|x \sim N(\mu_1, \sigma_1^2)$, where

$$\mu_1 = \frac{\sigma_0^2 x + \sigma_x^2 \mu_0}{\sigma_0^2 + \sigma_x^2}$$

$$= \mu_0 \left(\frac{1/\sigma_0^2}{1/\sigma_0^2 + 1/\sigma_x^2} \right) + x \left(\frac{1/\sigma_x^2}{1/\sigma_0^2 + 1/\sigma_x^2} \right), \quad \text{and} \tag{3.3a}$$

$$\sigma_1^2 = \frac{\sigma_0^2 \sigma_x^2}{\sigma_0^2 + \sigma_x^2} = \frac{1}{1/\sigma_0^2 + 1/\sigma_x^2}. \tag{3.3b}$$

The posterior mean, μ_1, is a weighted average of the mean of the prior belief and the new message. The weights are determined by the *precisions*, i.e., the reciprocals of the variances of the prior belief and the new message. This is a fantastically convenient result, as it permits us to compute posterior means without multiplying the prior probability distribution by the likelihood. And this convenience helps to account for

the popularity of the normal-normal model. But the model has shortcomings that make it normatively unattractive and descriptively unrealistic:

1. *Surprise is impossible.* The model implies that people always become more certain of their beliefs over time.

2. *Polarization is unthinkable.* Under widely assumed conditions, the model implies that people who initially disagree are literally incapable of holding posterior beliefs that are less alike than their priors.

3. *Agreement is inevitable.* The model implies that people will always disagree less as they learn more. Furthermore, learning enough will always cause their beliefs to converge to agreement.[5]

All of these shortcomings are peculiar to the normal-normal model; they are not inherent properties of Bayesian updating. In the remainder of this article, I elaborate each of these shortcomings and define other Bayesian updating models that surmount them.

---

[5] These two statements are not equivalent: it is possible for an updating model to imply ever-diminishing disagreement without implying eventual agreement. But the normal-normal model implies both.

The Normal-Normal Model Implies that People Always Become More Certain, But Other Bayesian Updating Models Do Not

A desirable property of any learning model is that surprising new evidence can cause people to become less sure of their beliefs. A distressing property of the normal-normal model is that surprising new evidence always makes people *more* sure of their beliefs. Formally, note that the variance of a posterior belief in the normal-normal model can be expressed as

$$\sigma_1^2 = \sigma_0^2 \left(\frac{\sigma_x^2}{\sigma_0^2 + \sigma_x^2} \right).$$

If σ_x^2 is finite, $\frac{\sigma_x^2}{\sigma_0^2 + \sigma_x^2} < 1$, and therefore $\sigma_1^2 < \sigma_0^2$: updating will always make people more certain of their belief. Moreover, the extent to which new information is surprising has no bearing on the extent to which it changes the certainty of one's beliefs. This shortcoming of the normal-normal model is often noted (Leamer 1978; Gerber and Green 1998; Bartels 1993, 2002; Grynaviski 2006), even though it has not dented the model's popularity. Fortunately, the problem is anything but endemic to Bayesian updating. It is an artifact of an unrealistic assumption of the normal-normal model: that we know the variance of the distribution that we are trying to learn about, and need only estimate its mean. This situation is as rare in politics as it is in any other domain. A model in which we simultaneously learn about the mean and variance of the unknown distribution is both more realistic on its face and capable of accommodating cases in which people become less certain over time.

Updating with Unknown Mean and Unknown Variance

Instead of trying to learn only about the mean of a distribution, we are now trying

to learn about both its mean, μ, and its variance, σ^2. Our prior belief about these

parameters is a bivariate joint distribution, $p(\mu, \sigma^2)$. It can be expressed as the product

of a conditional prior belief about the mean, $p(\mu|\sigma^2)$, and a marginal prior belief about

the variance, $p(\sigma^2)$. Many different densities might be used to model these priors,

but two of the most obvious choices are a normal distribution for the mean and an

inverse-Gamma distribution for the variance. (The inverse-Gamma distribution is

attractive for modeling variances because it is continuous, flexible, and has a lower

bound of 0 but no upper bound.) Specifically, the normal/inverse-Gamma prior

distribution can be expressed as the product of two densities,

$$\mu|\sigma^2 \sim N\left(\mu_0, \frac{\sigma^2}{n_0}\right)$$

$$\sigma^2 \sim \text{inverse-Gamma}\left(\frac{\nu_0}{2}, \frac{\nu_0 \sigma_0^2}{2}\right),$$

where

- μ_0 is the mean of the prior belief about μ

- σ^2/n_0 is the variance of the prior distribution for μ, conditional on σ^2. n_0 is
  interpretable as "prior sample size": the smaller it is, the larger the variance
  of the prior belief, reflecting the fact that a prior based on less information
  (or fewer "prior observations") is less precise than a prior based on more
  information

- $\nu_0 > 0$ is a prior shape parameter, i.e., a prior "degrees of freedom" parameter

- $v_0 \sigma_0^2$ is a prior scale parameter, equivalent to the sum of squared residuals one would obtain from a previously observed dataset of size v_0 in which each observation came from a $N(\mu_0, \sigma_0^2)$ dataset.

The marginal inverse-Gamma prior distribution has a χ^2 shape, and indeed, the inverse-χ^2 distribution is a special case of the inverse-Gamma distribution. (Gelman et al. 2004 discuss updating with normal/inverse-χ^2 priors; see Grynaviski 2006 for an application.)

Suppose that we encounter messages $\mathbf{x} = (x_1, \ldots, x_n)$ that we believe bear directly on the parameters about which we are trying to learn; i.e., we believe that $x_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2) \; \forall \; i \in (1, \ldots, n)$. If our prior belief about μ and σ^2 is a normal/inverse-Gamma distribution with parameters μ_0, n_0, v_0, and σ_0^2, our posterior belief will also be a normal/inverse-Gamma distribution:

$$\mu | \sigma^2, \mathbf{x} \sim N\left(\mu_1, \frac{\sigma^2}{n_0 + n}\right) \text{ and} \tag{3.4a}$$

$$\sigma^2 | \mathbf{x} \sim \text{inverse-Gamma}\left(\frac{v_1}{2}, \frac{v_1 \sigma_1^2}{2}\right), \tag{3.4b}$$

where $\mu_1 = \frac{n_0 \mu_0 + n\bar{x}}{n_0 + n}$, $v_1 = v_0 + n$, and $v_1 \sigma_1^2 = v_0 \sigma_0^2 + \sum_{i=1}^{n} (x_i - \bar{x})^2 + \frac{n_0 n}{n_0 + n} (\mu_0 - \bar{x})^2$.

Usually, interest focuses on the marginal posterior distribution of μ, which is a t distribution:

$$\mu | \mathbf{x} = t_{v_1}\left(\mu_1, \sqrt{\sigma_1^2 / (n_0 + n)}\right).$$

Now, the marginal prior distribution of μ is $t_{\nu_0}\left(\mu_0, \sqrt{\sigma_0^2/n_0}\right)$. If $\sigma_0^2/n_0 < \sigma_1^2/(n_0 + n)$, the posterior belief about μ has a higher variance than the prior. This occurs when

$$\frac{\sigma_0^2}{n_0} < \frac{\sigma_1^2}{n_0 + n}$$

$$\sigma_0^2\left(\frac{n_0 + n}{n_0}\right) < \frac{\nu_0\sigma_0^2 + \sum_{i=1}^{n}(x_i - \bar{x})^2 + \frac{n_0 n}{n_1}(\mu_0 - \bar{x})^2}{\nu_0 + n}$$

$$\sigma_0^2\left(\frac{n_0 + n}{n_0}\right)(\nu_0 + n) - \nu_0\sigma_0^2 < \sum_{i=1}^{n}(x_i - \bar{x})^2 + \frac{n_0 n}{n_0 + n}(\mu_0 - \bar{x})^2.$$

$$\frac{\sigma_0^2}{n_0}\left(n\nu_0 + nn_0 + n^2\right) + \frac{\sigma^2}{n_0}(n_0\nu_0) - \nu_0\sigma_0^2 < \sum_{i=1}^{n}(x_i - \bar{x})^2 + \frac{n_0 n}{n_0 + n}(\mu_0 - \bar{x})^2$$

$$\frac{\sigma_0^2}{n_0}[n(\nu_0 + n_0 + n)] < \sum_{i=1}^{n}(x_i - \bar{x})^2 + \frac{n_0 n}{n_0 + n}(\mu_0 - \bar{x})^2. \quad (3.5)$$

Equation 3.5 establishes that people who update according to Equations 3.4a and 3.4b will become less sure if they learn from data that are sufficiently surprising ($(\bar{x} - \mu_0)^2$ is large enough) or sufficiently vague ($\sum_{i=1}^{n}(x_i - \bar{x})^2$ is large enough). Thus, in contrast to the normal-normal model, a model that presumes that both a mean and a variance are unknown permits new information to shake the confidence that people repose in their beliefs.

Bayesian Updating Models Can Accommodate Biased Interpretation of Political Information

It is common to hear or read that "Bayesian updating requires independence between priors and new evidence" (Taber and Lodge 2006, 767; see also Ottati 1990, 160; Fischle 2000). The notion underpinning the claim is, presumably, that Bayes' Theorem demands that beliefs correspond to some objective conception of reality. But there

is not even a germ of truth to such claims. Nothing in Bayes' Theorem—nothing in the other writing of Reverend Bayes—nothing in the writing of his contemporary, Laplace—nothing in the whole of Bayesian statistics past or present warrants such a claim. "Objective perception" of political information may belong in a standard of rationality—and this is a point to which I shall return—but it has no place in a framework for belief updating. Indeed, the entire history of Bayesian scholarship militates against the notion that understanding of probability requires or profits from objective perception of the world (de Finetti 1974; Savage 1964; Jeffrey 2004).[6] Still, there are some for whom the dream will never die, and Bayesian updating models are quite able to accommodate them.

Partisanship is often thought to influence political views through *selective perception*: the assimilation, from the evidence at hand, of only or chiefly those details that support one's prior beliefs (Campbell et al. 1960, esp. Chapter 6; Bartels 2002; Taber and Lodge 2006; Jacobson 2006; Gaines et al. 2007; see also Bullock 2006). Nothing in the normal-normal model is incompatible with selective perception; like Bayes' Theorem, it makes no prescription about the way in which evidence is to be interpreted. Still, researchers who have a standard of objective perception may prefer a model that explicitly distinguishes between selective and objective perception. The normal-normal model can be adapted to this task.

We begin by distinguishing "good" messages that comport with one's prior and "bad" messages that do not. Formally, let the former set of messages be $x_g = (x_1, \ldots, x_t)$ and the latter set be $x_b = (x_{t+1}, \ldots, x_T)$. $x_i \overset{iid}{\sim} N(\mu, \sigma_x^2) \ \forall \ i \in (1, \ldots, T)$.

---

[6] Of course, this is not to deny that there are objective probabilities; although many contemporary "subjectivist Bayesians" (e.g., de Finetti, Savage, Diaconis) deny that there are, Bayes and Laplace did not. Note that there is a school of thought that goes by the name "Objective Bayes," but the "objectivity" that its members favor amounts to the rejection of certain types of prior beliefs as inappropriate—not to the assumption that there are objectively correct likelihoods for data, least of all in the ambiguous world of politics. On both counts, see Press 2003.

$\mathbf{x_g} \equiv x_g \sim N\left(\mu, \sigma_x^2/t\right); \mathbf{x_b} \equiv x_b \sim N\left(\mu, \sigma_x^2/(T-t)\right)$. If one's prior belief is

$\mu \sim N\left(\mu_0, \sigma_0^2\right)$, and one updates according to the normal-normal model, one's posterior

is $\mu|\mathbf{x_g}, \mathbf{x_b} \sim N\left(\mu_1, \sigma_1^2\right)$, where

$$\mu_1 = \mu_0\left(\frac{\sigma_1^2}{\sigma_0^2}\right) + \bar{x}_g\left(\frac{\sigma_1^2}{\sigma_x^2/t}\right) + \bar{x}_b\left(\frac{\sigma_1^2}{\sigma_x^2/(T-t)}\right),$$

$\bar{x}_g = \sum_{i=1}^{t} x_i/t$, and $\bar{x}_b = \sum_{i=t+1}^{T} x_i/(T-t)$. As each x_i is subjectively determined, so

too are $\bar{x}_g$ and $\bar{x}_b$: different people interpreting the same evidence may come up with

different values of these "sample means." But if we, as researchers, have a standard

of objective interpretation, we can make this subjective assessment explicit by using

a modified model. The formula for the mean of the posterior is given by Gerber and

Green (1999):

$$\mu_1 = \mu_0\left(\frac{1/\sigma_0^2}{1/\sigma_0^2 + t/\sigma_x^2 + (T-t)/\sigma_x^2}\right)$$

$$+ \alpha_g\bar{x}_{g*}\left(\frac{t/\sigma_x^2}{1/\sigma_0^2 + t/\sigma_x^2 + (T-t)/\sigma_x^2}\right) + \alpha_b\bar{x}_{b*}\left(\frac{(T-t)/\sigma_x^2}{1/\sigma_0^2 + t/\sigma_x^2 + (T-t)/\sigma_x^2}\right).$$

$$(3.6)$$

α_g and α_b are the *selection weights*; they indicate the extent to which favorable and

unfavorable messages are misinterpreted. $\bar{x}_{g*}$ and $\bar{x}_{b*}$ are the objective sample means

of the favorable and unfavorable messages, i.e., the sample means in the eyes of the

researcher. In the absence of selective perception, $\alpha_g = \alpha_b = 1$. If higher values

of x_i are preferable, and selective perception consists of giving an unduly favorable

interpretation to bad news, α_b will be greater than 1. If selective perception consists of

exaggerating the good news provided by favorable information, α_g will be greater than

one.

A closely related form of political bias is at work when people attribute more or less credibility to a news source than it deserves: when Communist Party officials exalt the *People's Daily*, perhaps; or when Republicans take Rush Limbaugh at face value. This calls for a slightly different model:

$$\mu_1 = \mu_0 \left(\frac{1/\sigma_0^2}{1/\sigma_0^2 + t/\sigma_x^2 + (T-t)/\sigma_x^2} \right)$$

$$+ \bar{x}_g \left(\frac{\alpha_g t/\sigma_x^2}{1/\sigma_0^2 + \alpha_g t/\sigma_x^2 + (T-t)/\sigma_x^2} \right) + \bar{x}_b \left(\frac{\alpha_b(T-t)/\sigma_x^2}{1/\sigma_0^2 + t/\sigma_x^2 + \alpha_b(T-t)/\sigma_x^2} \right).$$

$$(3.7)$$

Here, α_g and α_b apply not to the content of new information but to its credibility. If one overrates the credibility of favorable messages, α_g is greater than 1: it it is as though he is responding to more favorable messages than he really received. If one underrates the credibility of unfavorable messages, α_b is less than 1: it is as though he is responding to fewer favorable messages than he has received.

Convergence and Polarization of Public Opinion under Bayesian Updating

No issue in Bayesian analysis of public opinion is more disputed than the implications of Bayesian updating for disagreement among people with different prior beliefs. Gerber and Green (1999, 203-05) maintain that if Republicans and Democrats are Bayesian, they will agree neither more nor less as they update in response to new evidence. Empirically, Gerber and Green find just this patterning of presidential approval over time and adduce it as evidence that many people are Bayesians whose views are unaffected by partisan bias. Bartels (2002) cites the same public opinion data as evidence that people are biased Bayesians or not Bayesian at all:

unbiased application of Bayes' Theorem, he writes, implies convergence of public

opinion. Achen (2005, 334) agrees, and Grynaviski (2006, 331) claims that Bartels

"formally proved" that Bayesian updaters will "inexorably come to see the world

in the same way." (He didn't.) A closely related dispute is about polarization of

public opinion: Gerber and Green (1999) write that attitude polarization (Lord,

Ross, and Lepper 1979) is incompatible with Bayesian updating except under very

unusual circumstances, while Steenbergen (2002, 7-8) concludes exactly the opposite.

There is at least a little truth to all of these positions: Bayesian updating does imply

convergence of public opinion under some conditions, but these are more numerous

and more stringent than the discussions to date have acknowledged.

Before embarking on a series of proofs, it will help to distinguish between

three kinds of convergence, all of which are depicted in Figure 3.3. *Convergence*

to agreement occurs when prior beliefs converge to the same belief after updating.

Convergence to signal occurs when people's beliefs converge to the mean of the

distribution of messages that they are using to update; if beliefs converge to the

same signal, they also converge to agreement. This kind of convergence has been

widely discussed in Bayesian statistics, where it is subsumed by the broader topic

of consistency of Bayes estimates. In that literature, convergence to signal has

been proved to hold under general conditions for a wide variety of prior and data

distributions (e.g., Diaconis and Freedman 1986; Strasser 1981); I focus here on the

case of normal priors and data because of the ubiquity of these assumptions in political

science. *Convergence to truth* occurs when beliefs converge to the true parameter of

interest. Often, convergence to signal implies convergence to truth, but not always. If

beliefs are updated in response to messages from a biased news source, convergence to

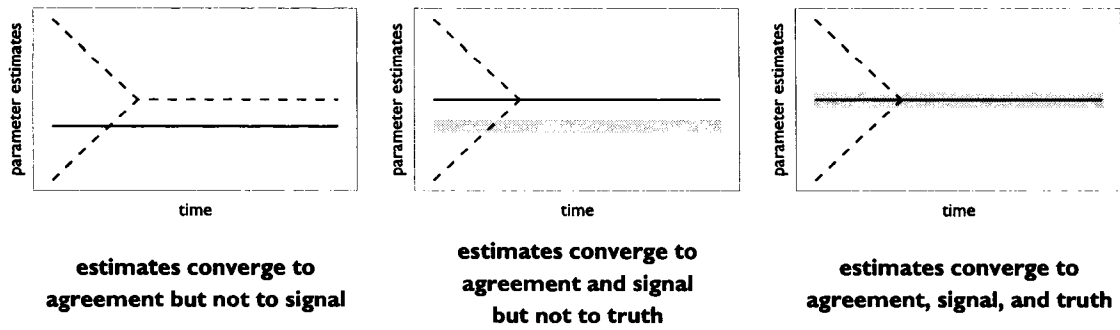signal implies that beliefs are not converging to the truth.

estimates converge to
agreement but not to signal

estimates converge to
agreement and signal
but not to truth

estimates converge to
agreement, signal, and truth

Figure 3.3: Convergence to Agreement, to Signal, and to Truth. The dashed lines depict parameter estimates by two people. They converge in each panel; this is *convergence to agreement*. The solid black line in each panel represents the mean of the distribution of information that people are using to update their estimates, e.g., the distribution from which news articles are drawn. In the leftmost panel, the parameter estimates do not converge to the mean of this distribution. In the middle and rightmost panels, they do: this is *convergence to signal* as well as convergence to agreement.

The grey line in the last two panels indicates the true parameter value. In the middle panel, it differs from the mean of the information distribution. This occurs whenever the information that people use to update their beliefs is biased on average. In the rightmost panel, the mean of the information distribution is also the true parameter value: here we have *convergence to truth* as well as to agreement and to signal.

Convergence of Public Opinion Under the Normal-Normal Model

Proposition 1. A person's prior belief is $\mu \sim N(\mu_0, \sigma_0^2)$. He updates according to Equation 3.3a in response to x, a sample of t messages that he perceives to have mean $\bar{x}$, with

$$x_i \overset{iid}{\sim} N(\mu, \sigma_x^2) \; \forall \; i \in (1, \ldots, t).$$ If t is large enough, his belief will converge to $\bar{x}$.

Proof. By a result shown in the appendix, the t messages are equivalent to a single message $\bar{x}$ from distribution $N(\mu, \sigma_x^2/t)$. Suppose $\epsilon > 0$ and $T = \frac{\sigma_x^2(|\mu_0 - \bar{x}| - \epsilon)}{\sigma_0^2 \epsilon}$. Then $t > T$ implies

$$|\mu|x - \bar{x}| = \left| \frac{\sigma_0^2 \bar{x} + (\sigma_x^2/t)\mu_0 - \bar{x}(\sigma_0^2 + \sigma_x^2/t)}{\sigma_0^2 + \sigma_x^2/t} \right|$$

$$= \left| \frac{(\sigma_x^2/t)(\mu_0 - \bar{x})}{\sigma_0^2 + \sigma_x^2/t} \right|$$

$$= \frac{\sigma_x^2 |\mu_0 - \bar{x}|}{t\sigma_0^2 + \sigma_x^2}$$

$$< \frac{\sigma_x^2 |\mu_0 - \bar{x}|}{T\sigma_0^2 + \sigma_x^2}$$

$$= \frac{\sigma_x^2 |\mu_0 - \bar{x}|}{\frac{\sigma_0^2[\sigma_x^2(|\mu_0 - \bar{x}| - \epsilon)]}{\sigma_0^2 \epsilon} + \sigma_x^2}$$

$$= \frac{\sigma_x^2 |\mu_0 - \bar{x}|}{\frac{\sigma_x^2(|\mu_0 - \bar{x}| - \epsilon)}{\epsilon} + \sigma_x^2}$$

$$= \frac{\epsilon \sigma_x^2 |\mu_0 - \bar{x}|}{\sigma_x^2 (|\mu_0 - \bar{x}| - \epsilon) + \epsilon \sigma_x^2}$$

$$= \epsilon.$$ □

Discussion. The proof reveals conditions under which any Bayesian updater's belief will converge to a subjectively defined $\bar{x}$. To some (e.g., Grynaviski 2006; Bartels 2002), it seems a short step to infer that Bayesians who initially disagree will come to agree with each other. In fact, the conditions set forth in the proof are quite stringent, and the conditions required for convergence to agreement are more stringent still.

First among the requirements for convergence to agreement is simply that people are Bayesian or that they adopt non-Bayesian updating rules that nevertheless permit convergence. This is a point too often elided by those who take nonconvergence as proof of partisan bias (e.g., Bartels 2002). There is ample laboratory evidence

that people are not Bayesian updaters, and while most documented non-Bayesian tendencies do nevertheless permit convergence (Phillips and Edwards 1966; Tversky and Kahneman 1971), some may not (Chapman and Chapman 1959; Hamm 1993).

The second requirement is that people with different priors perceive the new information in the same way: technically, they must agree on the value of $\bar{x}$. This rules out selective perception (unless people with different prior selectively perceive evidence in the same way—an unlikely circumstance, given that priors influence the strength and direction of selective perception). It also rules out cases in which people are updating evaluations that simply reflect different values. For example, if a Republican president's economic policies are consistent with Republican values and inconsistent with Democratic ones, we should not expect convergence even if members of both parties are Bayesians who have the same understanding of new economic information.

A third requirement likely to be violated is that people update exclusively on the basis of the same set of messages. If they do not—if, for example, i and j both update on the basis of x, but i also updates on the basis of messages that seem have a different sample mean—there is no reason to expect updating to cause agreement. This rules out *selective exposure*, whereby people with different views may systematically expose themselves chiefly to congenial news sources (Taber and Lodge 2006; Gentzkow and Shapiro 2006). It was once argued that selective exposure is uncommon—especially in politics—because people do not consciously seek to reinforce their views through their choice of media (Sears and Freedman 1967; Frey 1986). But selective exposure does not require a reinforcement motive (Katz 1968), and the heightened sorting of the electorate (Levendusky 2006) and splintering of the market for news into specialized niches (Prior 2007, Chapter 4) may have made it increasingly common.

A fourth requirement is that the parameter about which people are updating is constant over time. As I show below, weakening this assumption allows nonconvergence and polarization when people update their beliefs, even if they are updating in response to the same information and are interpreting that information in the same way.

Finally, complete convergence to agreement—the assumption that Bayesians will "inexorably come to see the world in the same way" (Grynaviski 2006, 331)—can only occur if updaters are responding to a set of messages so powerful that it causes them to completely ignore their prior beliefs. (Formally, the precision of the set of new messages must be infinitely greater than the precision of the prior beliefs.) Even relative to the other conditions, this is unrealistic. Bartels (1993) argues forcefully that people's beliefs about candidates at the start of Presidential campaigns are far stronger than we usually imagine. And it is widely known that most Americans are exposed to only meager amounts of political news (Campbell et al. 1960; Zaller 1992; Delli Carpini and Keeter 1996; Prior 2007). The assumption may hold in the extremely long run, but by then, as Keynes noted, we'll be dead.

Almost no interesting political predicaments satisfy all of these conditions. This suggests that the recent focus on convergence has been misplaced: the failure of Republicans and Democrats to evaluate the President in the same way or to otherwise share the same beliefs may be evidence of selective perception, but it may also be due to any of several other, quite likely factors. On the other hand, convergence becomes more likely as more of these conditions are met, and it would be quite surprising if Bayesian updating did *not* imply convergence of public opinion when all of them are met. That is why the next two proofs are interesting: they show that convergence may not occur under Bayesian updating if these conditions are only slightly relaxed.

Nonconvergence and Polarization Under Bayesian Updating with Selective Perception

Proposition 2: Nonconvergence and Polarization Under Bayesian Updating with Selective Perception. Let voters i and j have prior belief $\mu \sim N(\mu_0, \sigma_0^2)$. Both update in response to $\mathbf{x}$, a set of T messages, under the presumption that $x_i \overset{\text{iid}}{\sim} N(\mu, \sigma_x^2)$ $\forall\ i \in$ $(1, \ldots, T)$. As in Equation 3.6, $\mathbf{x}$ can be partitioned into $\mathbf{x_g}$, a set of t messages that favor the voters' prior, and $\mathbf{x_b}$, a set of $(T - t)$ messages that contradict the voters' prior. Assume $t \geq 1$ and $(T - t) \geq 1$. The objective mean of $\mathbf{x_g}$ is $\bar{x}_{g*}$ and the objective mean of $\mathbf{x_b}$ is $\bar{x}_{b*}$; again, these "objective" means are stipulated by the researcher. Assume $\bar{x}_{g*} \neq \bar{x}_{b*}$. Voter i updates by Equation 3.6 with selection weights α_{gi} and α_{bi}. Voter j updates by Equation 3.6 with selection weights α_{gj} and α_{bj}. If $\alpha_{gi} \neq \alpha_{gj}$ or $\alpha_{bi} \neq \alpha_{bj}$, convergence to agreement cannot occur unless

$$\left(\alpha_{gi} - \alpha_{gj}\right)\left(\frac{t\bar{x}_{g*}}{T}\right) = \left(\alpha_{bi} - \alpha_{bj}\right)\left(\frac{(T - t)\bar{x}_{b*}}{T}\right),$$

and that will only occur by chance.

Proof. By Appendix A, updating in response to the messages in $\mathbf{x_g}$ and $\mathbf{x_b}$ is equivalent to updating in response to a single draw

$$\bar{x}_* = \frac{\alpha_g \bar{x}_{g*}\sigma_x^2/(T - t) + \alpha_b \bar{x}_{b*}\sigma_x^2/t}{\sigma_x^2/t + \sigma_x^2/(T - t)}$$

$$= \frac{\alpha_g \bar{x}_g \sigma_x^2/(T - t) + \alpha_b \bar{x}_{b*}\sigma_x^2/t}{\sigma_x^2 T/n(T - t)}$$

$$= \left(\alpha_g \bar{x}_g \sigma_x^2/(T - t) + \alpha_b \bar{x}_{b*}\sigma_x^2/t\right)\frac{n(T - t)}{\sigma_x^2 T}$$

$$= \frac{t\alpha_g \bar{x}_{g*} + (T - t)\alpha_b \bar{x}_{b*}}{T}$$

from a normal distribution with variance σ_x^2/T. By Equation 3.3a, the posterior mean is

$$\mu|\mathbf{x} = \mu_0\left(\frac{1/\sigma_0^2}{1/\sigma_0^2 + T/\sigma_x^2}\right) + \bar{x}_*\left(\frac{T/\sigma_x^2}{1/\sigma_0^2 + T/\sigma_x^2}\right).$$

Assume $\epsilon > 0$. By Proposition 1, $|\mu|\mathbf{x} - \bar{x}_*| < \epsilon$ for $T > \frac{\sigma_x^2(|\mu_0 - \bar{x}_*| - \epsilon)}{\sigma_0^2 \epsilon}$, i.e., $\mu|\mathbf{x}$ will converge to $\bar{x}_*$. It only remains to show that $\bar{x}_*$ takes on different values for i and j when they have different selection weights:

$$\bar{x}_{*i} - \bar{x}_{*j} = \frac{t\alpha_{gi}\bar{x}_{g*} + (T-t)\alpha_{bi}\bar{x}_{b*} - t\alpha_{gj}\bar{x}_{g*} + (T-t)\alpha_{bj}\bar{x}_{b*}}{T}$$

$$= \left(\alpha_{gi} - \alpha_{gj}\right)\left(\frac{t\bar{x}_{g*}}{T}\right) - \left(\alpha_{bi} - \alpha_{bj}\right)\left(\frac{(T-t)\bar{x}_{b*}}{T}\right).$$

where $\bar{x}_{*i}$ is the value of $\bar{x}_*$ using i's selection weights and $\bar{x}_{*j}$ is the value of $\bar{x}_*$ using j's selection weights. By assumption, t and $T - t$ are positive, and $\bar{x}_{g*} \neq \bar{x}_{b*}$, so the posterior beliefs of i and j will never converge to agreement unless

$$\left(\alpha_{gi} - \alpha_{gj}\right)\left(\frac{t\bar{x}_{g*}}{T}\right) = \left(\alpha_{bi} - \alpha_{bj}\right)\left(\frac{(T-t)\bar{x}_{b*}}{T}\right). \qquad \square$$

Discussion. The proof shows that Bayesians with the same prior beliefs who update under selective perception will generally have different prior beliefs; it is therefore not just a nonconvergence result but a polarization result, too. Of course, nonconvergence is no less likely when i and j have different prior beliefs: it depends entirely on the difference between $\bar{x}_{*i}$ and $\bar{x}_{*j}$, and not at all on the difference between prior

beliefs. And even when i and j have different priors, selective perception will lead to polarization when

$$\left|\mu_{0i} - \mu_{0j}\right| < \left|\left(\alpha_{gi} - \alpha_{gj}\right)\left(\frac{t\bar{X}_{g*}}{T}\right) - \left(\alpha_{bi} - \alpha_{bj}\right)\left(\frac{(T-t)\bar{X}_{b*}}{T}\right)\right|.$$

Convergence and Polarization under Kalman-Filter Updating

The models considered to this point assume that people are trying to learn about μ, a quality of the political environment that does not change. Achen (1992), for example, assumes that the Democratic and Republican Parties offer benefits to each voter that oscillate over time around a mean benefit level that never changes, and he uses the assumption to justify his use of the normal-normal model to study changes in party identification. Such fixed-mean assumptions are apt when we are trying to learn about history and perhaps when we are trying to update our beliefs over the short term. But they are inappropriate when the parameter of interest changes over time. *Pace* Achen, the net benefit that I derive from a party changes as my views or economic status change. My preferences over policies change as new proposals are placed on the table or taken off of it. And candidates may improve during their time in office or fall increasingly under the sway of constituents whose views I oppose. In all of these cases, the constant-parameter assumption is an approximation at best.

Suppose that a Bayesian is trying to learn about a parameter that changes according to the rule

$$\alpha_t = \gamma\alpha_{t-1} + \epsilon_\alpha, \quad \epsilon_\alpha \sim N(0, \sigma_\alpha^2) \tag{3.8}$$

where t is the current period, γ is a known autoregressive parameter, and ϵ_α is a disturbance term with known, finite, nonzero variance σ_α^2. Let $x_1, \ldots, x_{t-1}, x_t$ denote the observed values of the parameter of interest at times $1, \ldots, t-1, t$. The relationship between x_t and α_t is

$$x_t = \alpha_t + \epsilon_x, \quad \epsilon_x \sim N(0, \sigma_x^2) \tag{3.9}$$

where σ_x^2 is a known, finite, nonzero variance term. (Following Gerber and Green 1998 and Green, Gerber, and de Boef 1999, γ, σ_α^2, and σ_x^2 are held constant for simplicity of exposition. But the model described in this section can easily accommodate the case in which they change over time. See Meinhold and Singpurwalla 1983 for an example.) His initial belief is

$$\left(\alpha_0 \mid \gamma, \sigma_\alpha^2, \sigma_x^2\right) \sim N(\hat{\alpha}_0, P_0).$$

Looking forward to period 1, his prior belief about the parameter of interest is governed by Equation 3.8:

$$\left(\alpha_1 \mid \gamma, \sigma_\alpha^2, \sigma_x^2\right) \sim N(\gamma\hat{\alpha}_0, \gamma^2 P_0 + \sigma_\alpha^2).$$

But after receiving message x_1, he updates. Because both his prior belief and the perceived likelihood of x_1 are normal, he updates according to Equations 3.3a and 3.3b:

$$\left(\alpha_1 \mid \gamma, \sigma_\alpha^2, \sigma_x^2, x_1\right) \sim N(\hat{\alpha}_1, P_1), \text{ where}$$

$$\hat{\alpha}_1 = \gamma\hat{\alpha}_0 \left[\frac{1/(\gamma^2 P_0 + \sigma_\alpha^2)}{1/(\gamma^2 P_0 + \sigma_\alpha^2) + 1/\sigma_x^2}\right] + x_1 \left[\frac{1/\sigma_x^2}{1/(\gamma^2 P_0 + \sigma_\alpha^2) + 1/\sigma_x^2}\right]$$

$$P_1 = \frac{1}{1/(\gamma^2 P_0 + \sigma_\alpha^2) + 1/\sigma_x^2}.$$

And generally,

$$(\alpha_t \mid \gamma, \sigma_\alpha^2, \sigma_x^2, \mathbf{x_{t-1}}, x_t) \sim N(\hat{\alpha}_t, P_t), \text{ where}$$

$$\hat{\alpha}_t = \gamma\hat{\alpha}_{t-1} \left[\frac{1/(\gamma^2 P_{t-1} + \sigma_\alpha^2)}{1/(\gamma^2 P_{t-1} + \sigma_\alpha^2) + 1/\sigma_x^2}\right] + x_t \left[\frac{1/\sigma_x^2}{1/(\gamma^2 P_{t-1} + \sigma_\alpha^2) + 1/\sigma_x^2}\right]$$

$$(3.10)$$

$$P_t = \frac{1}{1/(\gamma^2 P_{t-1} + \sigma_\alpha^2) + 1/\sigma_x^2},$$

$$(3.11)$$

and where $\mathbf{x_{t-1}}$ is the vector of messages $x_1, \ldots, x_{t-1}$. Equations 3.10 and 3.11 are known as the *Kalman filter algorithm* after Kalman (1960) and Kalman and Bucy (1961), who show that Equation 3.10 yields the expected value of α_t under the assumption of normal errors. If the normality assumption is relaxed, the Kalman filter estimator of α_t remains best (i.e., least-squares-minimizing) among all linear estimators. Harvey (1989) and Beck (1990) contain extensive descriptions of the Kalman filter and its properties. Goussev (2004) argues from a neurological perspective that the human brain unconsciously uses it to update probabilities. Meinhold and Singpurwalla (1983) provide a lucid introduction to it from a Bayesian point of view. Gerber and Green (1998) note that the normal-normal model is a special

case of the Kalman filter model in which $\gamma = 1$ and $\sigma_\alpha^2 = 0$. And given that the Kalman filter model is a Bayesian updating model, one of its surprising implications is that people's beliefs may diverge even if they are updating in response to the same messages and interpreting those messages in the same way.

Proposition 3A: Polarization Can Occur under Kalman Filtering Even in the Absence of Selective Perception and Selective Exposure. Assume that α_t, σ_α^2, γ, x_t, and σ_x^2 are as described above. Let $K_t = \frac{1/\sigma_x^2}{1/(\gamma^2 P_{t-1}+\sigma_\alpha^2)+1/\sigma_x^2}$. Voter i's belief about α_0 is $N(\hat{\alpha}_{i0}, P_0)$. Voter j's belief about α_0 is $N(\hat{\alpha}_{j0}, P_0)$. They are exposed to one message at each stage t; the entire set of messages is x_t. They update their beliefs according to Equations 3.10 and 3.11. If $\hat{\alpha}_{i0} \neq \hat{\alpha}_{j0}$, their beliefs diverge from time t to time $t + 1$ if and only if $(1 - K_{t+1})|\gamma| > 1$.

Proof. We begin with a lemma: the Kalman filter estimator $\hat{\alpha}_t$ can be written as a linear function of $\hat{\alpha}_0$,

$$\hat{\alpha}_t = c_t \hat{\alpha}_0 + \mathbf{f}_t' \mathbf{x_t}, \tag{3.12}$$

where $c_t = \prod_{i=1}^{t}(1 - K_i)\gamma$, $\mathbf{f}_t'$ is a row vector, and $\mathbf{x_t}$ is the column vector of messages $x_1, \ldots, x_t$. (See the appendix for a proof.)

Proof by contradiction: assume some t such that beliefs at $t + 1$ are not more polarized than beliefs at t even though $(1 - K_{t+1}) |\gamma| > 1$. Note that $0 < 1 - K_t < 1$ $\forall$ t. Then

$$\left| \hat{\alpha}_{it} - \hat{\alpha}_{jt} \right| - \left| \hat{\alpha}_{i,t+1} - \hat{\alpha}_{j,t+1} \right| \geq 0$$

$$\Rightarrow \left| c_t \hat{\alpha}_{i0} + \mathbf{f}_t' \mathbf{x}_t - (c_t \hat{\alpha}_{j0} + \mathbf{f}_t' \mathbf{x}_t) \right| - \left| (c_{t+1} \hat{\alpha}_{i0} + \mathbf{f}_{t+1}' \mathbf{x}_{t+1}) - (c_{t+1} \hat{\alpha}_{j0} + \mathbf{f}_{t+1}' \mathbf{x}_{t+1}) \right| \geq 0$$

$$\Rightarrow |c_t| - |c_{t+1}| > 0$$

$$\Rightarrow \prod_{i=1}^{t} (1 - K_i) |\gamma| - (1 - K_{t+1}) |\gamma| \prod_{i=1}^{t} (1 - K_i) |\gamma| \geq 0$$

$$\Rightarrow (1 - K_{t+1}) |\gamma| \leq 1,$$

which is a contradiction. This establishes that divergence occurs between times t and $t + 1$ if $(1 - K_{t+1}) |\gamma| > 1$. Now assume some t such that beliefs at $t + 1$ are less alike than beliefs at t even though $(1 - K_{t+1}) |\gamma| \leq 1$. Then

$$\left| \hat{\alpha}_{it} - \hat{\alpha}_{jt} \right| - \left| \hat{\alpha}_{i,t+1} \hat{\alpha}_{j,t+1} \right| < 0$$

$$\Rightarrow \prod_{i=1}^{t} (1 - K_i) |\gamma| - (1 - K_{t+1}) |\gamma| \prod_{i=1}^{t} (1 - K_i) |\gamma| < 0$$

$$\Rightarrow (1 - K_{t+1}) |\gamma| > 1,$$

which is a contradiction. This establishes that divergence occurs only if $(1 - K_{t+1}) |\gamma| > 1$. □

Proposition 3B: Convergence to Agreement Occurs Eventually under the Kalman Filter. Assume the conditions of Proposition 3A. At some period t, K_t will reach a steady state. After that point, polarization will not be possible.

Proof. We begin with a lemma: K_t will gradually converge to

$$K = \frac{\gamma^2 - c - 1 + \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2}}{2\gamma^2},$$

where $c = \sigma_\alpha^2/\sigma_x^2$ (Gerber and Green 1998; see the appendix for a proof). By

Proposition 3A, divergence occurs if and only if $(1 - K_{t+1})|\gamma| > 1$, but this is not

possible when $K_{t+1} = K$.

Proof by contradiction. $(1 - K_{t+1})|\gamma| > 1$ implies $|\gamma| > 1$, because $0 < K_t < 1 \; \forall \; t$.

If $\gamma > 1$, divergence in the steady state implies

$$\left(1 - \frac{\gamma^2 - c - 1 + \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2}}{2\gamma^2}\right)\gamma > 1$$

$$\Rightarrow \gamma - \frac{\gamma^2 - c - 1 + \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2}}{2\gamma} > 1$$

$$\Rightarrow -\left(\gamma^2 - c - 1 + \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2}\right) > (1 - \gamma)2\gamma$$

$$\Rightarrow c + 1 - \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2} > 2\gamma - \gamma^2$$

$$\Rightarrow c + 1 - \sqrt{\gamma^4 + 2c\gamma^2 - 2\gamma^2 + c^2 + 2c + 1} > 2\gamma - \gamma^2$$

$$\Rightarrow \gamma^2 - 2\gamma + c + 1 > \sqrt{\gamma^4 + 2c\gamma^2 - 2\gamma^2 + c^2 + 2c + 1}$$

$$\Rightarrow \gamma^4 - 4\gamma^3 + 2c\gamma^2 + 6\gamma^2 - 4c\gamma - 4\gamma + c^2 + 2c + 1 > \gamma^4 + 2c\gamma^2 - 2\gamma^2 + c^2 + 2c + 1$$

$$\Rightarrow -4\gamma^3 + 6\gamma^2 - 4c\gamma - 4\gamma > -2\gamma^2$$

$$\Rightarrow -4\gamma^3 + 8\gamma^2 - 4\gamma c - 4\gamma > 0.$$

This implies $c < -\gamma^2 + 2\gamma - 1$. But $-\gamma^2 + 2\gamma - 1 \le 0 \; \forall \; \gamma$, and c must be positive because

it is a ratio of positive variances. Contradiction.

If $\gamma < -1$,

$$\left(1 - \frac{\gamma^2 - c - 1 + \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2}}{2\gamma^2}\right)|\gamma| > 1$$

$$\Rightarrow |\gamma| - \frac{\gamma^2 - c - 1 + \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2}}{2|\gamma|} > 1$$

$$\Rightarrow -\left(\gamma^2 - c - 1 + \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2}\right) > (1 - |\gamma|)2|\gamma|$$

$$\Rightarrow -\left(\gamma^2 - c - 1 + \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2}\right) > 2|\gamma| - 2\gamma^2$$

$$\Rightarrow c + 1 - \sqrt{(-\gamma^2 + c + 1)^2 + 4c\gamma^2} > 2|\gamma| - \gamma^2$$

$$\Rightarrow \gamma^2 - 2|\gamma| + c + 1 > \sqrt{\gamma^4 + 2c\gamma^2 - 2\gamma^2 + c^2 + 2c + 1}$$

$$\Rightarrow \gamma^4 - 4|\gamma|^3 + 6\gamma^2 + 2\gamma^2 c - 4|\gamma| - 4|\gamma|c + c^2 + 2c + 1 > \gamma^4 + 2c\gamma^2 - 2\gamma^2 + c^2 + 2c + 1$$

$$\Rightarrow -4|\gamma|^3 + 8\gamma^2 - 4|\gamma|c - 4|\gamma| > 0.$$

This implies $c < -\gamma^2 - 2\gamma - 1$. But $-\gamma^2 - 2\gamma - 1 \leq 0 \ \forall \ \gamma$, and c must be positive because it is a ratio of variances. Contradiction. □

Discussion. The proof of Proposition 3A shows that polarization can happen even in the absence of selective exposure or perception: it occurs when updating causes people who disagree to weight their prior beliefs more heavily than they did before. And this occurs when $(1 - K_{t+1})|\gamma| > 1$. Because K_t lies between 0 and 1 for all values of t, $(1 - K_{t+1})$ must also lie between 0 and 1, and $|\gamma| > 1$ is therefore a necessary condition for polarization. It may seem, moreover, that bigger values of $|\gamma|$ produce more polarization. But this is not generally so, because γ enters into the definition of K_t, too, and the algorithm that determines K_t quickly "catches up" to offset the size of $|\gamma|$ in $(1 - K_t)|\gamma|$. Proposition 3B shows that the algorithm catches up completely when K_t reaches its steady state K: it is impossible for $(1 - K)|\gamma|$ to be greater than 1.
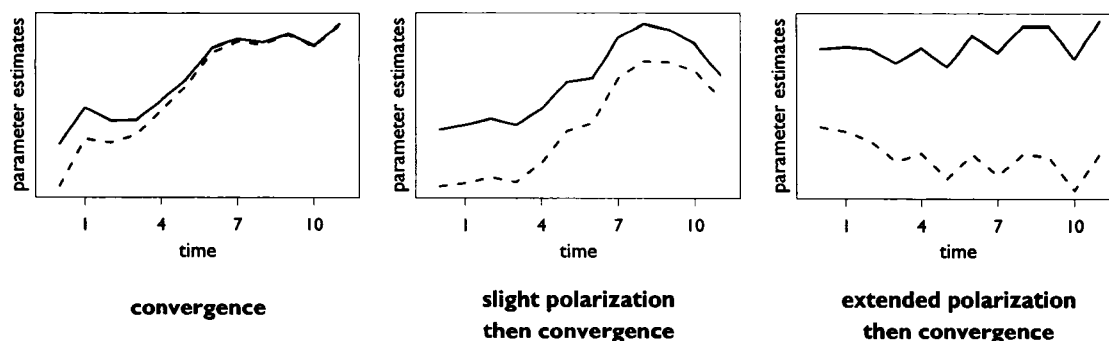
Figure 3.4: Convergence and Polarization under Kalman Filter Updating. All panels depict simulated belief updating by two hypothetical voters when $\gamma = 1.1$, $\alpha_0 = 1$, the voter represented by the solid line has prior belief $N(1,1)$, and the voter represented by the dashed line has prior belief $N(-1,1)$. The scale of the y axes differs across panels to make differences between the voters apparent.

In the leftmost panel, $\sigma_\alpha^2 = .7$ and $\sigma_x^2 = 4$. We see the intuitive pattern of belief updating under Bayes' Theorem: subjects who receive the same messages and interpret them in the same way draw closer to agreement every time they update. The convergence is represented by the vertical distance between the solid and dashed line, which diminishes in each period.

The middle panel depicts updating when $\sigma_\alpha^2 = .5$ and $\sigma_x^2 = 25$. Even though voters are receiving the same messages and interpreting them in the same way, their beliefs diverge (slightly) from period 0 to 1 and again from period 1 to period 2. Not until period 4 is the distance between their beliefs smaller than the distance between their priors. This pattern is exacerbated in the rightmost panel, where $\sigma_\alpha^2 = .19$ and $\sigma_x^2 = 100$. Here, voters' views diverge continuously in each of the first eleven periods; not until the 22nd period (unshown) is the distance between their beliefs smaller than the distance between their priors.

In practice, polarization is sustained longest when $|\gamma|$ is barely greater than 1 and when σ_x^2 is far smaller than σ_α^2. (See Figure 3.4.) The former condition occurs in politics whenever a parameter of interest is trending slowly away from zero.[7] Almost all political parameters of interest trend away from zero sometimes, and some (e.g., population, racial tolerance, per capita income in developed countries) seem to trend slowly away from zero for very long periods of time. The second condition, $\sigma_x^2 \ll \sigma_\alpha^2$, exists whenever the true variation in a parameter is slight but the quality of the

---

[7] For parameters that have no natural scale, zero is arbitrarily defined. For example, in the case of the President's honesty, zero might correspond to a survey answer of "neither honest nor dishonest."

messages that we receive about it is poor. This is likely when secretive or totalitarian states exert control over the press. Schumpeter (1942) argues that it is also a condition endemic to democratic politics: although there are knowable political truths, the feedback that we receive about our political decisions is so poor as to verge on useless. Presume, for example, that we want to know whether the President is a good steward of the economy. How are we to know? If we think that we have observed improvement in the economy, it might be attributed to the President's policies, or the lagged effects of his predecessor's policies, or the actions of incumbents at other levels of government, or the lagged effects of their predecessors, or the Federal Reserve, or wholly apolitical factors. It is rarely easy to tell who is responsible. And this elides the difficulty of simply knowing when the economy has improved. Even a free and robust media will not be of much help in these cases: no matter how precise the signals they send about the state of the economy, signals about who deserves credit and blame are unavoidably vague.[8]

Bartels (2002, 123) tells us that "it is *failure* to converge that requires explanation within the Bayesian framework," and he fingers selective perception by Democrats and Republicans as the culprit in their failure to agree on a host of actual matters. He may be right, but the case is far from cinched, and his characterization of Bayesian updating is too strong. Bayesian updating models imply only partial convergence of only some beliefs, and only under fantastically rare conditions do they imply that people who disagree will "inexorably come to see the world in the same way." The logic of Bayesian updating alone provides no reasonable expectation of

---

[8] Indeed, $\sigma_x^2 \ll \sigma_a^2$ in the case of attributions of praise and blame is the root of Schumpeter's (1942, 262) contention that "the typical citizen drops down to a lower level of mental performance as soon as he enters the political field." The problem is not that political man is stupid but that causality in politics is so complex and feedback about political decisions is so poor. Man's "lower level of mental performance" is due to his failure to sufficiently adjust for the poor quality of information at his disposal.

convergence—not during a campaign, not even over a lifetime. Of course, agreement may occur. But enduring disagreement is no proof that people are not Bayesian. Still less is it proof of partisan bias.

Discussion: Bayes' Theorem as a Normative Standard of Belief Updating

As a way of describing how citizens actually update their beliefs, Bayesian updating models have come in for a wealth of criticism. Much laboratory research shows that people do not update as Bayesians (Phillips and Edwards 1966; Tversky and Kahneman 1971; though see Koehler 1996). And some political scientists believe that the Theorem cannot accommodate ordinary features of public opinion about politics (Taber and Lodge 2006; Fischle 2000). This essay shows, to the contrary, that important features of public opinion about politics are readily accommodated by Bayesian updating. Political events are often surprising, and Bayesian updating can reflect that surprise by causing people to hold their views less confidently. Partisan bias affects people's views through selective perception or misjudgments of the credibility of media outlets, and Bayesian updating can easily accommodate this. Most importantly, people who disagree about politics are rarely moved to agreement by even a flood of evidence. Exposure to the evidence may even cause their views to draw further apart. Contrary to what some have written, Bayes' Theorem can easily accommodate enduring disagreement and polarization, too. None of this means that Bayesian updating models perfectly capture every facet of political decision-making—no models do—but it does mean that they are better than many suppose.

But it is precisely this flexibility that makes Bayesian updating inadequate as a standard of rational thinking about politics. One need only consider all that it permits. There is no limit to the amount of evidence that a Bayesian may ignore or misconstrue, for Bayes' Theorem applies only to updating, not to the collection or interpretation of evidence. And even in the absence of perceptual biases—even if people are interpreting and using new information in exactly the same way—Bayes' Theorem permits their views to polarize. No standard of rational belief revision should be so permissive.

The problem is not that the Theorem is irrational; as we have seen, it entails abiding by axioms of probability so fundamental that they should be a component of any standard of rational thinking. The problem is that the Theorem is not restrictive enough: it permits what we should reject as irrational. Political scientists interested in constructing standards of rational political thinking will do well to couple it with restrictions on the interpretation of political evidence. Bayes' Theorem should be part of any normative criteria by which we judge rational political thinking—but only a part.

Appendix H: Proofs of Several Bayesian Updating Results

Updating in Response to Many Normal Messages is Equivalent to Updating in Response to a Single, More Precise Normal Message

Updating sequentially in response to many messages is equivalent to updating once in response to a single, more precise message. I show this result for normal-normal updating—first for the case in which all the messages are drawn from the same distribution, then for the case in which different messages are drawn from distributions with different variances.

Formally, assume (as in the normal-normal model) that we are trying to learn the mean of a distribution whose variance is known. Let the prior belief about the mean be $\mu \sim N\left(\mu_0, \sigma_0^2\right)$. A random sample of messages $\mathbf{x} = (x_1, \ldots, x_n)$ is drawn independently from a distribution that is presumed to be $N(\mu, \sigma_x^2)$. Because Bayes' Theorem requires that the posterior be proportional to the prior times the likelihood of the new messages, updating sequentially requires that

$$p(\mu|\mathbf{x}) \propto \text{prior} \times \mathcal{L}(x_1|\mu) \times \cdots \times \mathcal{L}(x_n|\mu)$$

$$= \text{prior} \times \exp\left[-\frac{(x_1-\mu)^2}{2\sigma_x^2}\right] \times \cdots \times \exp\left[-\frac{(x_n-\mu)^2}{2\sigma_x^2}\right]$$

$$= \text{prior} \times \exp\left[-\frac{n}{2\sigma_x^2}\left(\mu^2 - 2\mu\bar{x} + \frac{x_1^2 + \cdots + x_n^2}{n}\right)\right]$$

$$= \text{prior} \times \exp\left[-\frac{n}{2\sigma_x^2}\left(\mu^2 - 2\mu\bar{x} + \bar{x}^2\right)\right] \times \exp\left[-\frac{n}{2\sigma_x^2}\left(\frac{x_1^2 + \cdots + x_n^2}{n} - \bar{x}^2\right)\right].$$

The part that does not depend on μ is constant. Absorbing it into the proportionality constant, we get

$$p(\mu|\mathbf{x}) \propto \text{prior} \times \exp\left[-\frac{1}{2}\left(\frac{(\bar{x} - \mu)^2}{\sigma_x^2/n}\right)\right].$$

The rightmost term is the likelihood of a single observation from a normal distribution with mean μ and variance σ_x^2/n.

The case of updating with messages from distributions with different variances is very similar. Suppose that messages $\mathbf{x_1} = (x_1, \ldots, x_n)$ have mean $\bar{x}_1$ and are presumed to be drawn independently from the $N(\mu, \sigma_{x_1}^2)$ distribution, while messages $\mathbf{x_2} = (x_{n+1}, \ldots, x_N)$ have mean $\bar{x}_2$ and are presumed to be drawn independently from the $N(\mu, \sigma_{x_2}^2)$ distribution. By the previous result, this is equivalent to updating on the basis of just two messages:

$$p(\mu|\mathbf{x_1}, \mathbf{x_2}) \propto \text{prior} \times \exp\left[-\frac{1}{2}\left(\frac{(\bar{x}_1 - \mu)^2}{\sigma_{x_1}^2/n}\right)\right] \times \exp\left[-\frac{1}{2}\left(\frac{(\bar{x}_2 - \mu)^2}{\sigma_{x_2}^2/(N-n)}\right)\right].$$

Let $\sigma_{1*}^2 = \sigma_{x_1}^2/n$ and $\sigma_{2*}^2 = \sigma_{x_2}^2/(N-n)$. Then

$$p(\mu|\mathbf{x_1}, \mathbf{x_2}) \propto \text{prior} \times \exp\left[-\frac{1}{2}\left(\frac{(\bar{x}_1 - \mu)^2}{\sigma_{1*}^2} + \frac{(\bar{x}_2 - \mu)^2}{\sigma_{2*}^2}\right)\right]$$

$$= \text{prior} \times \exp\left[-\frac{1}{2}\left(\frac{(\bar{x}_1 - \mu)^2 \sigma_{2*}^2 + (\bar{x}_2 - \mu)^2 \sigma_{1*}^2}{\sigma_{1*}^2 \sigma_{2*}^2}\right)\right]$$

$$= \text{prior} \times \exp\left[-\frac{1}{2}\left(\frac{\left(\bar{x}_1^2 - 2\mu\bar{x}_1 + \mu^2\right)\sigma_{2*}^2 + \left(\bar{x}_2^2 - 2\mu\bar{x}_2 + \mu^2\right)\sigma_{1*}^2}{\sigma_{1*}^2 \sigma_{2*}^2}\right)\right]$$

$$= \text{prior} \times \exp\left[-\frac{1}{2}\left(\frac{\mu^2(\sigma_{1*}^2 + \sigma_{2*}^2) - 2\mu(\bar{x}_1\sigma_{2*}^2 + \bar{x}_2\sigma_{1*}^2) + \bar{x}_1^2\sigma_{2*}^2 + \bar{x}_2^2\sigma_{1*}^2}{\sigma_{1*}^2 \sigma_{2*}^2}\right)\right]$$

$$= \text{prior} \times \exp\left[-\frac{1}{2}\left(\mu^2 - \frac{2\mu(\bar{x}_1\sigma_{2*}^2 + \bar{x}_2\sigma_{1*}^2)}{\sigma_{1*}^2 + \sigma_{2*}^2} + \frac{\bar{x}_1^2\sigma_{2*}^2 + \bar{x}_2^2\sigma_{1*}^2}{\sigma_{1*}^2 + \sigma_{2*}^2}\right)\left(\frac{\sigma_{1*}^2 + \sigma_{2*}^2}{\sigma_{1*}^2 \sigma_{2*}^2}\right)\right].$$

Absorbing the part that does not depend on μ into the constant of proportionality, we get

$$p(\mu|x_1, x_2) \propto \text{prior} \times \exp\left[-\frac{1}{2(\sigma_{1*}^2\sigma_{2*}^2)/(\sigma_{1*}^2 + \sigma_{2*}^2)}\left(\mu - \frac{\bar{x}_1\sigma_{2*}^2 + \bar{x}_2\sigma_{1*}^2}{\sigma_{1*}^2 + \sigma_{2*}^2}\right)^2\right].$$

The rightmost term is the kernel of a normal density with mean $\frac{\bar{x}_1\sigma_{2*}^2 + \bar{x}_2\sigma_{1*}^2}{\sigma_{1*}^2 + \sigma_{2*}^2}$ and variance $\frac{\sigma_{1*}^2\sigma_{2*}^2}{\sigma_{1*}^2 + \sigma_{2*}^2}$.

Under Kalman Filtering, the Posterior Estimate of the Mean is a Linear Function of the Prior Belief and Messages Received

Equation 3.3a shows that when a normally distributed prior belief is updated in response to information that is perceived to be normal, the Bayes estimate of the posterior mean is a linear combination of the prior estimate of the mean and the new information. Diaconis and Ylvisaker (1979) show that this is not peculiar to updating with normal priors and likelihoods: whenever the prior belief has a distribution in the exponential family and is conjugate to the distribution of the new information, the Bayes estimate of the posterior mean is a convex combination of the prior estimate of the mean and the new information. In this appendix, I calculate the weights of the linear combination for the Kalman filter estimator $\hat{\alpha}_t$ of random variable α_t, which are defined on pages 111-113.

The Kalman filter estimator $\hat{\alpha}_t$ can be written as a linear function of $\hat{\alpha}_0$,

$\hat{\alpha}_t = c_t\hat{\alpha}_0 + \mathbf{f}_t'\mathbf{x}_t$, where $c_t = \prod_{i=1}^t(1 - K_i)\gamma$, K_t is the "Kalman gain" defined on page 114, $\mathbf{x}_t$ is the vector of messages $(x_1, \ldots, x_t)$, and $\mathbf{f}_t'$ is a row vector of weights on the components of $\mathbf{x}_t$.

Proof by induction. The claim is true for $t = 1$:

$$\hat{\alpha}_1 = \gamma\hat{\alpha}_0 + K_1(x_1 - \gamma\hat{\alpha}_0)$$

$$= (\gamma - K_1\gamma)\hat{\alpha}_0 + K_1 x_1$$

$$= c_1\hat{\alpha}_0 + \mathbf{f}_1'\mathbf{x}_1,$$

with $c_1 = (\gamma - K_1\gamma) = \prod_{i=1}^{1}(1 - K_i)\gamma$ and $\mathbf{f}_1 = K_1$. Assuming the claim is true for t, it is also true for $t + 1$:

$$\hat{\alpha}_{t+1} = \gamma\hat{\alpha}_t + K_{t+1}(x_{t+1} - \gamma\hat{\alpha}_t)$$

$$= (1 - K_{t+1})\gamma\hat{\alpha}_t + K_{t+1}x_{t+1}$$

$$= (1 - K_{t+1})\gamma[c_t\hat{\alpha}_0 + \mathbf{f}_t'\mathbf{x}_t] + K_{t+1}x_{t+1}$$

$$= [(1 - K_{t+1})\gamma c_t]\hat{\alpha}_0 + (1 - K_{t+1})\gamma\mathbf{f}_t'\mathbf{x}_t + K_{t+1}x_{t+1}$$

$$= c_{t+1}\hat{\alpha}_0 + \mathbf{f}_{t+1}'\mathbf{x}_{t+1},$$

with $c_{t+1} = (1 - K_{t+1})\gamma c_t$ and $\mathbf{f}_{t+1}$ a vector in which the first t elements are given by $(1 - K_{t+1})\gamma\mathbf{f}_t$ and in which the $(N + 1)_{\text{th}}$ element is K_{t+1}. $\square$

This proof follows the similar proof in Gerber and Green (1998). They are not identical because (perhaps due to a copyediting error) the earlier works reports $c_t = \prod_{i=1}^{t}(1 - K_i)\gamma^t$, which is not quite right.

K Has a Steady State

According to Gerber and Green (1998), the Kalman weight K_t in Kalman-filter updating stabilizes at a unique value:

$$K = \frac{-[c + (1 - \gamma^2)] + \sqrt{[c + (1 - \gamma^2)]^2 + 4c\gamma^2}}{2\gamma^2}$$

where $c = \sigma_\alpha^2/\sigma_x^2$.

Proof. $K_t = P_t/\sigma_x^2$, so behavior of P_t implies behavior of K_t. To find the steady state of P_t (and thus of K_t), we need to find the value for which $P_t = P_{t+1}$:

$$P_t = \sigma_x^2 K_{t+1}$$

$$= \sigma_x^2(\gamma^2 P_t + \sigma_\alpha^2)/(\gamma^2 P_t + \sigma_\alpha^2 + \sigma_x^2)$$

$$P_t(\gamma^2 P_t + \sigma_\alpha^2 + \sigma_x^2) = P_t\gamma^2\sigma_x^2 + \sigma_x^2\sigma_\alpha^2$$

$$P_t^2\gamma^2 + P_t(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2) - \sigma_x^2\sigma_\alpha^2 = 0.$$

This is just a quadratic equation, so

$$P = \frac{-(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2) \pm \sqrt{(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2)^2 - 4\gamma^2(-\sigma_x^2\sigma_\alpha^2)}}{2\gamma^2}.$$

The denominator is positive. The numerator must also be positive, then, because P must be positive.

σ_α^2, σ_x^2, and γ^2 must all be positive. The numerator will thus only be positive if it is $-(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2) + \sqrt{(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2)^2 - 4\gamma^2(-\sigma_x^2\sigma_\alpha^2)}$. In other words, we replace $\pm$ with $+$:

$$P = \frac{-(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2) + \sqrt{(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2)^2 - 4\gamma^2(-\sigma_x^2\sigma_\alpha^2)}}{2\gamma^2}$$

$$K = \frac{-[\sigma_\alpha^2 + \sigma_x^2(1 - \gamma^2)] + \sqrt{(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2)^2 + 4\gamma^2\sigma_x^2\sigma_\alpha^2}}{\sigma_x^2 2\gamma^2}$$

$$= \frac{-[c + (1 - \gamma^2)]}{2\gamma^2} + \frac{\sqrt{(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2)^2 + 4\gamma^2\sigma_x^2\sigma_\alpha^2}}{\sigma_x^2 2\gamma^2}$$

$$= \frac{\gamma^2 - c - 1}{2\gamma^2} + \frac{\sqrt{\frac{(\sigma_\alpha^2 + \sigma_x^2 - \gamma^2\sigma_x^2)^2}{(\sigma_x^2)^2} + 4c\gamma^2}}{2\gamma^2}$$

$$= \frac{\gamma^2 - c - 1}{2\gamma^2} + \frac{\sqrt{\frac{(\sigma_\alpha^2)^2 + (\sigma_x^2)^2 + \gamma^4(\sigma_x^2)^2 + 2\sigma_\alpha^2\sigma_x^2 - 2\sigma_\alpha^2\gamma^2\sigma_x^2 - 2\gamma^2(\sigma_x^2)^2}{(\sigma_x^2)^2} + 4c\gamma^2}}{2\gamma^2}$$

$$= \frac{\gamma^2 - c - 1 + \sqrt{(c - \gamma^2 + 1)^2 + 4c\gamma^2}}{2\gamma^2}.$$

That is, after some stage t, the newest observation always receives the same weight (K) when updating. By extension, one's prior always receives the same weight ($1 - K$) when updating.

This result is contrary to the normal-normal model, in which information received at previous stages is always reflected in the prior. As time passes, the weight placed on the prior under the normal-normal model always increases, because the prior always reflects more information than it did in the past. Consequently, the weight placed on new messages always decreases.

The Kalman filter is more realistic because it implies that very old information is either forgotten or discounted: when the parameter of interest is changing over time, a very old message is not as important as a new message, even if both come from

equally credible sources. That is why the weight placed on the prior when updating

under the Kalman filter model does not always increase over time. See Gerber and

Green (1998) for an extensive discussion of this difference between the normal-normal

and Kalman-filter models.

Bibliography

Abelson, Robert P. 1986. "Beliefs Are Like Possessions." *Journal for the Theory of Social Behaviour* 16 (October): 224-50.

Achen, Christopher H. 1992. "Social Psychology, Demographic Variables, and Linear Regression: Breaking the Iron Triangle in Voting Research." *Political Behavior* 14 (September): 195-211.

Achen, Christopher H. 2002. "Parental Socialization and Rational Party Identification." *Political Behavior* 24 (June): 151-70.

Achen, Christopher H. 2003. "Toward a New Political Methodology: Microfoundations and ART." *Annual Review of Political Science* 5 (June): 423-50.

Achen, Christopher H. 2005. "Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong." *Conflict Management and Peace Science* 22 (Winter): 327-39.

American Presidency Project. 2006 03 01. "Political Party Platforms." http://www.presidency.ucsb.edu/platforms.php.

Anderson, Craig A., B. Lynn New, and James R. Speer. 1985. "Argument Availability as a Mediator of Social Theory Perseverance." *Social Cognition* 3 (3): 235-49.

Anderson, Craig A., Mark R. Lepper, and Lee Ross. 1980. "The Perseverance of Social Theories: The Role of Explanation in the Persistence of Discredited Information." *Journal of Personality and Social Psychology* 39 (December): 1037-49.

Ansolabehere, Stephen, James M. Snyder, Jr., and Charles Stewart, III. 2001. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science* 45 (January): 136-59.

Ansolabehere, Stephen, and Shanto Iyengar. 1995. *Going Negative: How Political Advertisements Shrink and Polarize the Electorate*. New York: Free Press.

Ansolabehere, Stephen, Shigeo Hirano, James M. Snyder, Jr., and Michiko Ueda. 2006. "Party and Incumbency Cues in Voting: Are They Substitutes?" *Quarterly Journal of Political Science* 1 (March): 119-37.

Barge, Matthew. 9 August 2005. "NARAL Falsely Accuses Supreme Court Nominee Roberts." http://www.factcheck.org/article340.html.

Bargh, John A., Shelly Chaiken, Rajen Govender, and Felicia Pratto. 1992. "The Generality of the Automatic Attitude Activation Effect." *Journal of Personality and Social Psychology* 62 (June): 893-912.

Bartels, Larry M. 1993. "Messages Received: The Political Impact of Media Exposure." *American Political Science Review* 87 (June): 267-85.

Bartels, Larry M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24 (June): 117-50.

Beck, Nathaniel. 1990. "Estimating Dynamic Models Using Kalman Filtering." *Political Analysis* 1: 121-56.

Bernardo, José. M., and A. F. M. Smith. 1994. *Bayesian Theory*. New York: Wiley.

Bochner, Stephen, and Chester A. Insko. 1966. "Communicator Discrepancy, Source Credibility, and Opinion Change." *Journal of Personality and Social Psychology* 4 (December): 614-21.

Bound, John, David A. Jaeger, and Regina M. Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak." *Journal of the American Statistical Association* 90 (June): 443-50.

Box, George E. P., and George C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley.

Brader, Ted. 2006. *Campaigning for Hearts and Minds: How Emotional Appeals in Political Ads Work*. Chicago: University of Chicago Press.

Bullock, John G. 2006. "Partisanship and the Enduring Effects of False Political Information." Presented at the Annual Conference of the Midwest Political Science Association, Chicago.

Callander, Steven, and Simon Wilkie. 2007. "Lies, Damned Lies, and Political Campaigns." *Games and Economic Behavior*.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. Chicago: University of Chicago Press.

Chaiken, Shelly. 1979. "Communicator Physical Attractiveness and Persuasion." *Journal of Personality and Social Psychology* 37 (August): 1387-97.

Chapman, Loren J., and Jean P. Chapman. 1959. "Atmosphere Effect Re-examined." *Journal of Experimental Psychology* 58 (September): 220-26.

Chong, Dennis, and James N. Druckman. 2006. "Democratic Competition and Public

    Opinion." Presented at the Annual Meeting of the American Political Science

    Association, Philadelphia.

Cohen, Geoffrey L. 2003. "Party Over Policy: The Dominating Impact of Group

    Influence on Political Beliefs." *Journal of Personality and Social Psychology* 85

    (November): 808-22.

Cohen, Joshua. 1989. "Deliberation and Democratic Legitimacy." In *The Good Polity*,

    ed. Alan Hamlin and Phillip Pettit. Oxford: Blackwell.

Conover, Pamela Johnston, and Stanley Feldman. 1989. "Candidate Perception in an

    Ambiguous World: Campaigns, Cues, and Inference Processes." *American Journal*

    of Political Science 33 (November): 912-40.

Davis, Otto A., Melvin J. Hinich, and Peter C. Ordeshook. 1970. "An Expository

    Development of a Mathematical Model of the Electoral Process." *American Political*

    Science Review 64 (June): 426-48.

de Finetti, Bruno. 1974. *Theory of Probability: A Critical Introductory Treatment.*

    Trans. Antonio Machi and Adrian F. Smith. New York: Wiley.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about*

    Politics and Why It Matters. New Haven, CT: Yale University Press.

Diaconis, Persi, and David Freedman. 1986. "On the Consistency of Bayes Estimates."

    Annals of Statistics 14 (March): 1-26.

Diaconis, Persi, and Donald Ylvisaker. 1979. "Conjugate Priors for Exponential

    Families." *Annals of Statistics* 7 (March): 269-81.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: HarperCollins.

Druckman, James N. 2001. "The Implications of Framing Effects for Citizen Competence." *Political Behavior* 23 (September): 225-56.

Druckman, James N. 2004. "Political Preference Formation: Competition, Deliberation, and the (Ir)relevance of Framing Effects." *American Political Science Review* 98 (November): 671-86.

Druckman, James N., and Kjersten R. Nelson. 2003. "Framing and Deliberation: How Citizens' Conversations Limit Elite Influence." *American Journal of Political Science* 47 (October): 729-45.

Duelfer, Charles. 2004. "Comprehensive Report of the Special Advisor to the Director of Central Intelligence on Iraq's Weapons of Mass Destruction." https://www.cia.gov/cia/reports/iraq_wmd_2004/.

Epley, Nicholas, and David Dunning. 2000. "Feeling 'Holier than Thou': Are Self-Serving Assessments Produced by Errors in Self- or Social Prediction?" *Journal of Personality and Social Psychology* 79 (December): 861-75.

Fischle, Mark. 2000. "Mass Response to the Lewinsky Scandal: Motivated Reasoning or Bayesian Updating?" *Political Psychology* 21 (March): 135-59.

Freedman, David A. 1991. "Statistical Models and Shoe Leather." *Sociological Methodology* 21: 291-313.

Frey, Dieter. 1986. "Recent Research on Selective Exposure to Information." *Advances in Experimental Social Psychology* 19: 41-80.

Gaines, Brian J., James H. Kuklinski, Paul J. Quirk, Buddy Peyton, and Jay Verkuilen. 2007. "Same Facts, Different Interpretations: Partisan Motivation and Opinion on Iraq." *Journal of Politics*.

Gellman, Barton. 2004 January 7. "Iraq's Arsenal Was Only on Paper; Since Gulf War, Nonconventional Weapons Never Got Past the Planning Stage." *Washingon Post* p. A1.

Gellman, Barton, and Walter Pincus. 2003 August 10. "Depiction of Threat Outgrew Supporting Evidence." *Washington Post* p. A1.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 2004. *Bayesian Data Analysis*. 2nd ed. Boca Raton, FL: Chapman & Hall.

Gentzkow, Matthew, and Jesse M. Shapiro. 2006. "What Drives Media Slant? Evidence from U.S. Daily Newspapers." University of Chicago. Manuscript.

Gerber, Alan, and Donald Green. 1999. "Misperceptions about Perceptual Bias." *Annual Review of Political Science* 2: 189-210.

Gerber, Alan, and Donald P. Green. 1998. "Rational Learning and Partisan Attitudes." *American Journal of Political Science* 42 (July): 794-818.

Gerber, Elisabeth R., and John E. Jackson. 1993. "Endogenous Preferences and the Study of Institutions." *American Political Science Review* 87 (September): 639-56.

Gertz v. *Schmidt*. 1974. 418 U.S. 323.

Goussev, Valeri. 2004. "Does the Brain Implement the Kalman Filter?" *Behavioral and Brain Sciences* 27 (June): 404-05.

Green, Donald P., Alan S. Gerber, and Suzanna L. de Boef. 1999. "Tracking Opinion over Time: A Method for Reducing Sampling Error." *Public Opinion Quarterly* 63 (Summer): 178-92.

Groseclose, Tim. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science* 45 (October): 862-86.

Grynaviski, Jeffrey D. 2006. "A Bayesian Learning Model with Applications to Party Identification." *Journal of Theoretical Politics* 18 (July): 323-46.

Hamm, Robert M. 1993. "Explanations for Common Responses to the Blue/Green Cab Probabilistic Inference Word Problem." *Psychological Reports* 72 (1): 219-42.

Harvey, Andrew C. 1989. *Forecasting, Structural Time Series and the Kalman Filter.* New York: Cambridge University Press.

Higgins, E. Tory. 1996. "Knowledge Activation: Accessibility, Applicability, and Salience." In *Social Psychology: Handbook of Basic Principles*, ed. E. Tory Higgins and Arie W. Kruglanski. New York: Guilford Press.

Hovland, Carl I., and Muzafer Sherif. 1957. "Assimilation and Contrast Effects in Reactions to Communication and Attitude Change." *Journal of Abnormal Social Psychology* 55 (September): 244-52.

Husted, Thomas A., Lawrence W. Kenny, and Rebecca B. Morton. 1995. "Constituent Errors in Assessing Their Senators." *Public Choice* 83 (June): 251-71.

Isikoff, Michael, and John Barry. 2005 May 9. "Gitmo: SouthCom Showdown." *Newsweek* p. 10. http://www.msnbc.msn.com/id/7693014/site/newsweek/.

Iyengar, Shanto, and Nicholas A. Valentino. 2000. "Who Says What? Source Credibility as a Mediator of Campaign Advertising." In *Elements of Reason,*

ed. Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin. New York: Cambridge University Press.

Jackman, Simon. 2005. "Pooling the Polls Over an Election Campaign." *Australian Journal of Political Science* 40 (December): 499-517.

Jacobson, Gary C. 2006. *A Divider, Not a Uniter: George W. Bush and the American People*. Upper Saddle River, NJ: Pearson.

Jamieson, Kathleen Hall, and Paul Waldman. 2003. *The Press Effect: Politicians, Journalists, and the Stories that Shape the Political World*. New York: Oxford University Press.

Jeffrey, Richard. 2004. *Subjective Probability: The Real Thing*. New York: Cambridge University Press.

Jessee, Stephen A., and Douglas Rivers. 2007. "Political Judgment Under Uncertainty: Heuristics and Biases." Stanford University. Manuscript.

Jones, Edward E., and Richard E. Nisbett. 1972. "The Actor and the Observer: Divergent Perceptions of the Causes of the Behavior." In *Attribution: Perceiving the Causes of Behavior*, ed. Edward E. Jones, David E. Kanouse, Harold H. Kelley, Richard. E. Nisbett, Stuart Valins, and Bernard Weiner. Morristown, NJ: General Learning Press.

Kalman, Rudolph E. 1960. "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering, Series D* 82: 35-45.

Kalman, Rudolph E., and Richard S. Bucy. 1961. "New Results in Linear Filtering and Prediction Theory." *Journal of Basic Engineering, Series D* 83 (March): 95-108.

Kam, Cindy D. 2005. "Who Toes the Party Line? Cues, Values, and Individual Differences." *Political Behavior* 27 (June): 163-82.

Katz, Elihu. 1968. "On Reopening the Question of Selectivity in Exposure to Mass Communication." In *Theories of Cognitive Consistency: A Sourcebook*, ed. Robert P. Abelson, Elliot Aronson, William J. McGuire, Theodore M. Newcomb, Milton J. Rosenberg, and Percy H. Tannenbaum. Chicago: Rand McNally.

Keenan, Nancy. 11 August 2005. "Rebuttal to FactCheck.org Analysis of NARAL Pro-Choice America's TV Ad."
http://www.factcheck.org/UploadedFiles/Rebuttal-letter-from-Nancy-Keenan-to-FactCheck.org-8-11-05.pdf.

Kelley, Harold H. 1973. "The Processes of Causal Attribution." *American Psychologist* 28 (February): 107-28.

Kennan, George F. 1997. *George F. Kennan and the Origins of Containment, 1944-1946: The Kennan-Lukacs Correspondence*. Columbia, Missouri: University of Missouri Press.

Kerkhof, Peter. 1999. "Applying the Unimodel to Political Persuasion." *Psychological Inquiry* 10 (2): 137-40.

Kingdon, John W. 1984. *Agendas, Alternative, and Public Policies*. Boston: Little, Brown.

Knight, Jack, and James Johnson. 1997. "What Sort of Political Equality Does Deliberative Democracy Require?" In *Deliberative Democracy*, ed. James Bohman and William Rehg. Cambridge: MIT Press.

Koehler, Jonathan J. 1996. "The Base Rate Fallacy Reconsidered: Descriptive, Normative, and Methodological Challenges." *Behavioral and Brain Sciences* 19 (March): 1-53.

Krosnick, Jon A., and Donald R. Kinder. 1990. "Altering the Foundations of Support for the President through Priming." *American Political Science Review* 84 (June): 497-512.

Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F. Rich. 2000. "Misinformation and the Currency of Democratic Citizenship." *Journal of Politics* 62 (August): 790-816.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108 (November): 480-98.

Kurtz, Howard. 16 May 2005. "*Newsweek* Apologizes: Inaccurate Report on Koran Led to Riots." *Washington Post* p. A01.

Lau, Richard P., and David P. Redlawsk. 2001. "Advantages and Disadvantages of Heuristics in Political Decision Making." *American Journal of Political Science* 45 (October): 951-71.

Leamer, Edward E. 1978. *Specification Searches*. New York: Wiley.

Leamer, Edward E. 1983. "Let's Take the Con out of Econometrics." *American Economic Review* 73 (1): 31-43.

Lepper, Mark R., Lee Ross, and Richard R. Lau. 1986. "Persistence of Inaccurate Beliefs about the Self: Perseverance Effects in the Classroom." *Journal of Personality and Social Psychology* 50 (March): 482-91.

Levendusky, Matthew S. 2006. "Sorting in the U.S. Mass Electorate." Ph.D. diss. Stanford University.

Lieb, David A. 2005 April 7. "Missouri Medicaid Cuts Sent to Blunt." *Associated Press*.

Lodge, Milton, and Charles Taber. 2000. "Three Steps toward a Theory of Motivated Political Reasoning." In *Elements of Reason*, ed. Arthur Lupia, Mathew D. McCubbins, and Samuel L. Popkin. New York: Cambridge University Press.

Lohmann, Suzanne. 1994. "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-91." *World Politics* 47 (October): 42-101.

Londregan, John B. 2000. *Legislative Institutions and Ideology in Chile*. New York: Cambridge University Press.

Lord, Charles. G., Lee Ross, and Mark R. Lepper. 1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." *Journal of Personality and Social Psychology* 37 (November): 2098-2109.

Luce, R. Duncan. 1995. "Four Tensions Concerning Mathematical Modeling in Psychology." *Annual Review of Psychology* 46: 1-26.

Lupia, Arthur. 2002. "Who Can Persuade Whom? Implications from the Nexus of Psychology and Rational Choice Theory." In *Thinking about Political Psychology*, ed. James H. Kuklinski. New York: Cambridge University Press.

Lupia, Arthur. 2006. "When Can Politicians Scare Citizens Into Supporting Bad Policies? A Theory of Incentives with Fear-Based Content." University of Michigan at Ann Arbor. Typescript.

Lupia, Arthur, and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* New York: Cambridge University Press.

Lupia, Arthur, Mathew D. McCubbins, and Samuel L. Popkin, eds. 2000. *Elements of Reason*. New York: Cambridge University Press.

Luskin, Robert C., James S. Fishkin, and Roger Jowell. 2002. "Considered Opinions: Deliberative Polling in Britain." *British Journal of Political Science* 32 (July): 455-87.

McGuire, William J. 1964. "Inducing Resistance to Persuasion: Some Contemporary Approaches." *Advances in Experimental Social Psychology* 1: 191-229.

McGuire, William J. 1969. "The Nature of Attitudes and Attitude Change." In *Handbook of Social Psychology*, ed. Gardner Lindzey and Elliot Aronson. 2nd ed. Reading, MA: Addison-Wesley.

Meinhold, Richard J., and Nozer D. Singpurwalla. 1983. "Understanding the Kalman Filter." *American Statistician* 37 (May): 123-27.

Mondak, Jeffery J. 1993a. "Public Opinion and Heuristic Processing of Source Cues." *Political Behavior* 15 (June): 167-92.

Mondak, Jeffery J. 1993b. "Source Cues and Policy Approval: The Cognitive Dynamics of Public Support for the Reagan Agenda." *American Journal of Political Science* 37 (February): 186-212.

NARAL Pro-Choice America. 2005. "Speaking Out." Television advertisement. http://www.factcheck.org/video/Speaking-Out.wmv.

Neill, Daniel B. 2005. "Cascade Effects in Heterogeneous Populations." *Rationality and Society* 17 (2): 191-241.

Newell, Allen, J.C. Shaw, and Herbert A. Simon. 1958. "Chess-Playing Programs and the Problem of Complexity." *IBM Journal of Research and Development* 2 (October): 320-35.

Nielsen Media Research. 18 August 2005. "Supreme Court Political Ads Virtually Absent From Television." Press release.

Ottati, Victor C. 1990. "Determinants of Political Judgments: The Joint Influence of Normative and Heuristic Rules of Inference." *Political Behavior* 12 (June): 159-79.

Page, Benjamin I., and Richard A. Brody. 1972. "Policy Voting and the Electoral Process: The Vietnam War Issue." *American Political Science Review* 66 (September): 979-95.

Petrocik, John R. 1996. "Issue Ownership in Presidential Elections, with a 1980 Case Study." *American Journal of Political Science* 40 (August): 825-50.

Petty, Richard E., and John T. Cacioppo. 1986. *Communication and Persuasion: Central and Peripheral Routes to Attitude Change.* New York: Springer-Verlag.

Phillips, Lawrence D., and Ward Edwards. 1966. "Conservatism in a Simple Probability Inference Task." *Journal of Experimental Psychology* 72 (September): 346-54.

Poole, Keith T., and Howard L. Rosenthal. 2006. *Ideology and Congress.* Piscataway, NJ: Transaction Publishers.

Popkin, Samuel L. 1994. *The Reasoning Voter: Communication and Persusasion in Presidential Campaigns.* 2nd ed. Chicago: University of Chicago Press.

Pratt, John W., Howard Raiffa, and Robert Schlaifer. 1964. "The Foundations of Decision under Uncertainty: An Elementary Exposition." *Journal of the American Statistical Association* 59 (June): 353-75.

Press, S. James. 2003. *Subjective and Objective Bayesian Statistics.* 2nd ed. Hoboken, NJ: Wiley.

Prior, Markus. 2007. *Post-Broadcast Democracy*. New York: Cambridge University Press.

Program on International Policy Attitudes. 2003 February 21. "Americans on Iraq and the U.N. Inspections II." http://www.pipa.org/OnlineReports/Iraq/IraqUNInsp2_Feb03/IraqUNInsp2%20Feb03%20rpt.pdf.

Rahn, Wendy M. 1993. "The Role of Partisan Stereotypes in Information Processing about Political Candidates." *American Journal of Political Science* 37 (May): 472-96.

Rawls, John. 1993. *Political Liberalism*. New York: Columbia University Press.

Ross, Lee, and Andrew Ward. 1996. "Naive Realism in Everyday Life: Implications for Social Conflict and Misunderstanding." In *Values and Knowledge*, ed. T. Brown, E. Reed, and E. Turiel. New Jersey: Lawrence Erlbaum.

Ross, Lee, Mark R. Lepper, and Michael Hubbard. 1975. "Perseverance in Self-Perception and Social Perception: Biased Attributional Processes in the Debriefing Paradigm." *Journal of Personality and Social Psychology* 32 (May): 880-92.

Savage, Charlie. 2005 June 4. "Pentagon Details Koran Mistreatment." *Boston Globe* p. A1.

Savage, Leonard J. 1954. *The Foundations of Statistics*. New York: Wiley.

Savage, Leonard J. 1964. "The Foundations of Statistics Reconsidered." In *Studies in Subjective Probability*, ed. Henry E. Kyburg, Jr. and Howard E. Smokler. New York: Wiley.

Schaffner, Brian F., and Matthew J. Streb. 2002. "The Partisan Heuristic in Low-Information Elections." *Public Opinion Quarterly* 66 (Winter): 559-81.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context.* San Diego: Academic Press.

Schumpeter, Joseph A. 1942. *Capitalism, Socialism, and Democracy.* New York: Harper & Brothers.

Schwieder, David W., and Paul J. Quirk. 2004. "Missed Cues: Judgment Heuristics and Citizen Competence." Presented at the Annual Conference of the Midwest Political Science Association, Chicago.

Sears, David O., and Jonathan L. Freedman. 1967. "Selective Exposure to Information: A Critical Review." *Public Opinion Quarterly* 31 (Summer): 194-213.

Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69 (February): 99-118.

Simon, Herbert A. 1959. "Theories of Decision-Making in Economics and Behavioral Science." *American Economic Review* 49 (June): 253-83.

Simon, Herbert A., and Allen Newell. 1958. "Heuristic Problem Solving: The Next Advance in Operations Research." *Operations Research* 6 (January-February): 1-10.

Smith, Edward E. 1990. "Categorization." In *Thinking: An Invitation to Cognitive Science,* ed. Daniel N. Osherson and Edward E. Smith. Cambridge, MA: MIT Press.

Smith, Edward E., and Douglas L. Medin. 1981. *Categories and Concepts.* Cambridge, MA: Harvard University Press.

Sniderman, Paul M., and Sean M. Theriault. 2004. "The Structure of Political Argument and the Logic of Issue Framing." In *Studies in Public Opinion: Attitudes, Nonattitudes, Measurement Error, and Change*, ed. Willem E. Saris and Paul M. Sniderman. Princeton, New Jersey: Princeton University Press.

Steenbergen, Marco R. 2002. "Political Belief Updating: An Experimental Investigation Within a Bayesian Framework." Presented at the Annual Scientific Meeting of the International Society of Political Psychology, Berlin.

Stokes, Donald E. 1963. "Spatial Models of Party Competition." *American Political Science Review* 57 (June): 368-77.

Strasser, Helmut. 1981. "Consistency of Maximum Likelihood and Bayes Estimates." *Annals of Statistics* 9 (September): 1107-1113.

Taber, Charles S., and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50 (July): 755-69.

Tessin, Jeff. 2006. "Cues Given, Cues Received: How Candidates Use Shortcuts When Voters Need Them Most." Presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

Tetlock, Philip E. 2005. *Expert Political Judgment*. Princeton: Princeton University Press.

Tversky, Amos, and Daniel Kahneman. 1971. "Belief in the Law of Small Numbers." *Psychological Bulletin* 76 (August): 105-10.

van Houweling, Robert P., and Paul M. Sniderman. 2004. "The Political Logic of a Downsian Space." Presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

Webster, James G. 2005. "Beneath the Veneer of Fragmentation: Television Audience Polarization in a Multichannel World." *Journal of Communication* 55 (June): 366-82.

White House. 2002 February 14. "President Announces Clear Skies & Global Climate Change Initiatives." Speech. http://www.whitehouse.gov/news/releases/2002/02/20020214-5.html.

Whitney v. *California*. 1927. 274 U.S. 357.

Williams, Daniel, and Kamran Khan. 2005 May 28. "Muslims Rally Over Koran Report." *Washington Post* p. A14.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Zanna, Mark P., and John K. Rempel. 1988. "Attitudes: A New Look at an Old Concept." In *The Social Psychology of Knowledge*, ed. Daniel Bar-Tal and Arie W. Kruglanski. New York: Cambridge University Press.

Zechman, Martin J. 1979. "Dynamic Models of the Voter's Decision Calculus: Incorporating Retrospective Considerations into Rational-Choice Models of Individual Voting Behavior." *Public Choice* 34 (September): 297-315.