

Social Influence on Political Judgments: The Case of Presidential Debates

Steven Fein

Williams College

George R. Goethals

University of Richmond

Matthew B. Kugler

Princeton University

Four experiments investigated the extent to which judgments of candidate performance in presidential debates could be influenced by the mere knowledge of others' reactions. In Experiments 1 and 2 participants watched an intact version of a debate or an edited version in which either "soundbite" one-liners or the audience reaction to those soundbites were removed. In Experiment 3 participants saw what was supposedly the reaction of their fellow participants on screen during the debate. Participants in Experiment 4 were exposed to the reactions of live confederates as they watched the last debate of an active presidential campaign. In all studies, audience reactions produced large shifts in participants' judgments of performance. The results illustrate the power of social context to strongly influence individuals' judgments of even large amounts of relevant, important information, and they support the categorization of presidential debates as ambiguous stimuli, fertile ground for informational social influence.

KEY WORDS: presidential debates, conformity, ambiguity, informational social influence, political communication

Ronald Reagan leans back and enjoys the moment. It is the second of his two presidential debates with Walter Mondale as they near the finish line to the 1984 U.S. presidential election. He has just delivered his famous line: "I will not make age an issue of this campaign. I am not going to exploit, for political purposes, my opponent's youth and inexperience." The question had been a difficult one, asking

about Reagan's advanced age and the concerns raised by his performance in the first debate, but his answer has satisfied the audience and they are cheering. The moderator, Henry Trewhitt, says "I'd like to head for the fence and try to catch that one before it goes over." The newspapers soon report that Reagan's response has erased concerns over his age. Reagan's declining poll numbers quickly returned to their previously high levels. He wins reelection in 1984 in a 49-state landslide.

The above events could all be seen as independently stemming from Reagan's answer. The audience heard it, and applauded; the moderator heard it, and praised him; the newspapers heard it, and stopped discussing the age issue; the public heard it, and voted for him. But these events could also be seen as describing a causal path. The audience heard it and they applauded, the moderator heard the applause and praised Reagan's answer, and so on. The difference between these two ways of looking at what happened is that one, the latter, does not give the actors much credit. After the initial applause, everyone is part of the herd, running smoothly down preset tracks, conforming to the example set by a small studio audience that was neither better nor wiser than they. In American culture, great value is set on independence. One of the primary sources of enthusiasm for debates is the belief that they help the public makes up its own mind by allowing them to cut through the superficial. Our second, uncharitable, interpretation of these events is quite worrisome from this individualistic perspective, but that does not mean it is wrong.

While many studies have examined conformity in basic laboratory settings, none exists that explores how it may affect assessments of such complex and important real world events as presidential debates. Millions of people watch debates, discuss them, and read and watch analyses of the candidates' performances. They have been shown to play a role in the outcome of every campaign in which they have been held, and they may have been decisive in several (e.g., 1960, 1980, 1992, and 2000). Is conformity an important determinant of people's reactions to them?

While we do not know whether potential voters simply conform to the reactions of other viewers, we do know that contextual features of debates—for example, what network commentators say about them and whether citizens watch them on television or listen to them on the radio—make a difference (Kaid & Bystrom, 1999; Schroeder, 2000). An overall conclusion from Kraus's (1962) volume on *The Great Debates* of 1960 between John F. Kennedy and Richard M. Nixon was that the debates mattered whether people watched them or not. There was so much discussion, especially of Kennedy's performance in the first of the four debates, that a social reality took shape holding that Kennedy had "won" the debates, and this perception very likely was an important factor in the outcome of a very close election. Similarly, media commentary following the Ford-Carter foreign policy debate in 1976 contributed substantially to both the salience and negative evaluation of Ford's statement that there was no Soviet domination of Eastern Europe (Steeper, 1978). McKinnon, Tedesco, and Kaid (1993) showed that network commentary after the 1992 debates between George Bush, Bill

Clinton, and Ross Perot lessened Clinton's perceived margins of victory, and McKinnon and Tedesco (1999) showed that network commentary following the 1996 debates raised perceivers' assessments of both Bill Clinton and Bob Dole. While it is not entirely predictable how commentators' appraisals might affect voter opinion, it seems clear that they can and do.

Considering the influence of audience reaction specifically, some elements of the social influence literature suggest that they would produce little conformity in judging presidential debates. First, social influence often produces public compliance (Kiesler & Kiesler, 1969) through normative social influence, where people conform essentially to be liked by their peers. Debate evaluations poorly fit that mold. There is little normative pressure to conform to debate audiences; viewers are not threatened for failing to cheer when their fellows do, especially not when the fellows in question are a studio audience on the other side of the country. And debates are important in that they can affect private voting behavior on Election Day, when people will be answering to no one but themselves. When considering voting behavior, we must limit ourselves primarily to informational social influence, conformity that works on a person's desire to be right, not their desire to be liked (Deutsch & Gerard, 1955).

Debate audiences supply little fodder for informational social influence. Their reactions do little more than provide knowledge of others' judgments. They do not present arguments or reasons for those judgments. This stands in contrast to the separate processes of pre- and postdebate spin, which do provide at least some argumentation and analysis. Thus audience reactions provide information about the judgments of other people, but not the reasons for their judgments, and, as noted above, they apply little normative pressure to conform to or to adopt those judgments. In such situations, the actions of the others provide only information about the typical or predominant response to the situation and, therefore, what the appropriate response for the individual might be.

How likely is it that this seemingly limited form of social influence can affect a real-world judgment, one that we assume not only is important but also that most people believe is important? While we are intuitively drawn to doubt that people would conform on important tasks, conformity research does show that not only trivial laboratory judgments about the length of lines (Asch, 1951) or the movement of lights (Sherif, 1936), but also substantially more important, real-world questions, such as whether New Jersey is really being invaded by Martians (Cantril, 1940), can be heavily influenced by other people's reactions. Still, most of the research on social influence concerns judgments made in the laboratory and its external validity is often unknown. And even laboratory experiments show that conformity is often reduced when judgments are made more consequential (Baron, Vandello, & Brunsman, 1996). In a somewhat related vein, theory and research supporting the Elaboration Likelihood Model of persuasion indicate that individuals are less likely to be influenced by factors peripheral to the content of a message when the issue is an important one (Petty & Cacioppo, 1986).

With these considerations in mind, the literature suggests that conformity to studio audiences in debates should be weak. As noted, there is little interpersonal or normative pressure, no explanations are provided to explain the audience behavior, and the judgment is important, giving the participant every reason to work for accuracy. Is there any reason to think this limited form of social influence could affect such important judgments?

The results of McKinnon et al. (1993)'s research concerning the 1992 presidential debates suggested that viewers may typically process the debates at a relatively shallow, heuristic level. To the extent that this is the case, individuals are more likely to be susceptible to social influence and other factors peripheral to the content of the debate. Another reason is suggested by a finding in the Baron et al. study cited above. They showed that importance reduces conformity on clear-cut judgments, but not when the judgment is ambiguous. Are presidential debates clear stimuli? In reconstructing debates, people often see them as having clear outcomes. Commentators tell narrative-style stories explaining how candidates won or lost: Michael Dukakis as the unnatural "Iceman" in 1988, George H. W. Bush as the disconnected elitist in 1992, John Kennedy as the unexpected professional in 1960. After the tales are told enough times, it becomes hard to imagine that things could be seen any other way.

It is not obvious, though, that we should accept this retrospective consensus as good evidence that debates are clear stimuli. As noted above, audience and media reactions that accompany debates are a major part of the unfolding political stories. While it could be that the debates really are clear, it could also be that perceptions of debates are instead largely a function of how studio audiences, debate panelists, and various professional commentators respond and the narratives they construct. Would Lloyd Bentsen telling Dan Quayle that he was "no Jack Kennedy" in the 1988 vice-presidential debate (Germond & Witcover, 1989) have been as memorable if the audience hadn't reacted loudly? Polls show, for example, that Al Gore in 2000 (Jamieson & Waldman, 2002) and Gerald Ford in 1976 (Sears & Chaffee, 1979) both won debates in their immediate aftermath but lost them in the postdebate spin wars, Ford to Eastern European immigrants who felt quite sure that their homelands were under Soviet domination and Gore to parodies of him on the television show *Saturday Night Live*.

A critical question, then, is whether presidential debates constitute an ambiguous stimulus. Certainly debates provide a rich and substantial physical reality. Therefore individuals might not be much influenced by other people's reactions. Furthermore, people may have strong candidate preferences going into a debate and, often, strong expectations that their candidate will win, with the result that they have clear perceptions of exactly who won (their candidate) and little susceptibility to social influence. Sigelman and Sigelman's study (1984) on the Carter-Reagan debates found many people do exactly that, as did Sears and Chaffee's (1979) review of studies on the 1976 debates. On the other hand, the complex physical reality of debates might be difficult to interpret, leaving participants open to outside

“help” as has been seen in studies of pre- and postdebate spin (Fein, Frost, Goethals, & Kassin, 1994; Kugler & Goethals, 2005; Norton & Goethals, 2004). This perspective suggests a high potential for conformity effects. Because of the influence of debates in our modern political arrangements, it seems important to find out.

Four experiments are reported below. In the first three experiments, participants watched parts or all of the second 1984 Mondale-Reagan debate, approximately a decade (from 9 to 12 years) after the debate occurred. In the first of these studies, participants watched parts of the debates under one of three conditions: a pure, unaltered format; with two “soundbites” deleted; or with the soundbites remaining but the audience reaction to them deleted. With this design, we could isolate the effects of the content of the soundbites versus the effects of the audience reaction to those soundbites. Experiment 2 used a similar design, but in this study participants saw the entire 90-minute debate, rather than excerpts from the debate. Participants in Experiment 3 watched the excerpts from this debate, but saw superimposed on the video screen alleged real-time tracking of other people’s reactions. The fourth experiment we report was actually conducted before the first three, in October of 1992 on the night of the third debate between George Bush, Bill Clinton, and Ross Perot. It is reported last because it addresses a key “external reality” question lurking in the background of the first three. That is, can real voters be influenced by the reactions of others when watching a debate in real time, when it is actually broadcast, when there are real levels of engagement and resistance? In this study, participants watched the debate either in the presence of several confederates who cheered for pro-Bush remarks, in the presence of several confederates who cheered for pro-Clinton remarks, or with no confederates.

In all four experiments the main dependent variables were participants’ judgments of each candidate’s overall performance and personal qualities. Because we believe that presidential debates are, in fact, rather ambiguous stimuli and ripe for social influence, we predicted for all these experiments that the cues from the studio audience or from the participants’ peers around them would significantly influence participants’ judgments and that this social influence would emerge despite the large amount of relevant content expressed by the candidates during the debate.

Experiment 1

With recent presidential debates including the presence of a live studio audience, political commentators have suggested that too often debates turn on simple-minded soundbites, one-liners, slogans, and canned mini-speeches designed to elicit resounding applause (Sigelman, 1992; Zarevsky, 1992). In 1980, Ronald Reagan buried Jimmy Carter in a televised debate the moment he shook his head and, with a tone of sorrow in his voice, delivered the fatal blow, “There you go again.” Four years later, with concerns being raised about his age and mental acumen, Reagan turned around his faltering campaign with his one-liner noted

above: "I am not going to exploit . . . my opponent's youth and inexperience." Then in 1988, Lloyd Bentsen stopped Dan Quayle in his tracks by responding to Quayle's self-comparison to John F. Kennedy with the unforgettable line, "You're no Jack Kennedy." In all instances, live audiences loudly erupted with some combination of cheers, laughter, boos, and applause.

Our first experiment addressed two questions. First, how much impact do such memorable moments really have on people's judgments of a candidate's overall performance? It is part of the political lore that debate outcomes, and perhaps elections as well, may turn on these isolated moments. However, it may be that because one-liners constitute only a tiny fraction of what transpires in a debate, they actually have little effect on overall evaluations. To answer this question our first experiment presented a 40-minute debate segment either with or without one-liners and the audience reaction they generated. Second—and in the context of issues about social influence, perhaps a more interesting question—is it the *content* of the one-liners themselves, or just the audience *reaction* to them, that is responsible for any impact the soundbite moments might have? To answer this question we included a condition in which the candidate's remarks, but not the reaction to them, were presented. Thus we compared an unedited 40-minute segment with two edited variations—one that deleted the soundbite exchanges in their entirety and one that deleted just the audience reaction to the relevant one-liners.

Method

Participants and Design

Fifty-three introductory psychology students participated in exchange for course credit. By random assignment, participants saw one of three versions of a debate. One tape included two soundbites and the positive reactions they had drawn from the audience. In a second version, the soundbites and reactions were both deleted from the tape. In a third version, the soundbites were included but the audience reactions to them were not.

Procedure

Participants completed a brief predebate questionnaire in which they described their own party affiliation, level of involvement, and political orientation.¹ Next they were shown a 40-minute segment of the second (and final) 1984

¹ In each of the four studies, we analyzed the main and interactive effects of a variety of individual differences, including participants' gender, political party affiliation, degree of involvement and interest in politics, and so forth. In these studies, there were approximately equal numbers of men and women (all were undergraduates at Williams College). In each study, the majority of the participants identified themselves as Democrats (ranging from 51% of the sample to 57%; the percentage identifying themselves as Republican ranged from 12% to 26%, with the remaining identifying

debate between Reagan and Mondale. The focus of this debate was on foreign policy. Participants were randomly assigned to watch one of three versions of the videotape. In the *control* condition, the tape included two remarks by Reagan, both of which drew loud cheers and laughter from members of the live audience (each remark-reaction segment lasted for approximately 30 seconds). The first remark concerned a TV ad in which Mondale projected an image of strength by standing on the deck of the Aircraft Carrier Nimitz. Reagan quipped, "If he had had his way when the Nimitz was being planned, he would have been deep in the water out there, because there wouldn't have been any Nimitz to stand on. He was against it." Later, in response to an expected question concerning his age and mental competence, Reagan delivered the memorable remark with which we opened this paper: "I will not exploit, for political purposes, my opponent's youth and inexperience." In the *soundbite-deleted* condition both one-liners—as well as the reactions they had elicited—were deleted. In the *reaction-deleted* condition, Reagan's remarks were included but the audience's reactions were deleted.²

Following the tape, participants rated the two candidates for their overall performance on 0–100 point scales. They were also asked to indicate which candidate (or neither) they thought won the debate. Participants then answered additional questions concerning each candidate's handling of key substantive issues (Central America policy, the Middle East, and national defense) and personal character (sense of humor, intelligence, likeability, strength, competence, sincerity, and leadership). Finally, participants were asked if they were transported back in time to the 1984 campaign, who they thought they would have voted for.

Results and Discussion

Overall Performance

A series of one-way analyses of variance (ANOVAs) on the ratings of each candidate's performance, as well as the difference between these ratings, revealed significant effects as a function of condition. As can be seen in Figure 1, participants in the control condition tended to rate Reagan's performance much more positively than they rated Mondale's, reflecting the popular opinion of the debate

themselves as Independent or having no party affiliation). In less than half the analyses across the four studies, a significant main effect for party affiliation emerged, but in the slight majority of analyses, there was no significant difference as a function of party. No significant interactions between party affiliation and any of our independent variables emerged. In each study, there were no reliable gender differences, nor reliable differences as a function of any of the other measures of individual differences that we took. We therefore do not report the analyses as a function of these subject variables in the Results sections of these studies.

² There was occasional, very brief audience applause sprinkled in with the candidates' responses during other parts of the debate, but the reaction was always mild and fleeting. No sustained applause, jeering, or any other reactions could be heard anywhere on the tape, with the exception of the two soundbites in the control condition.

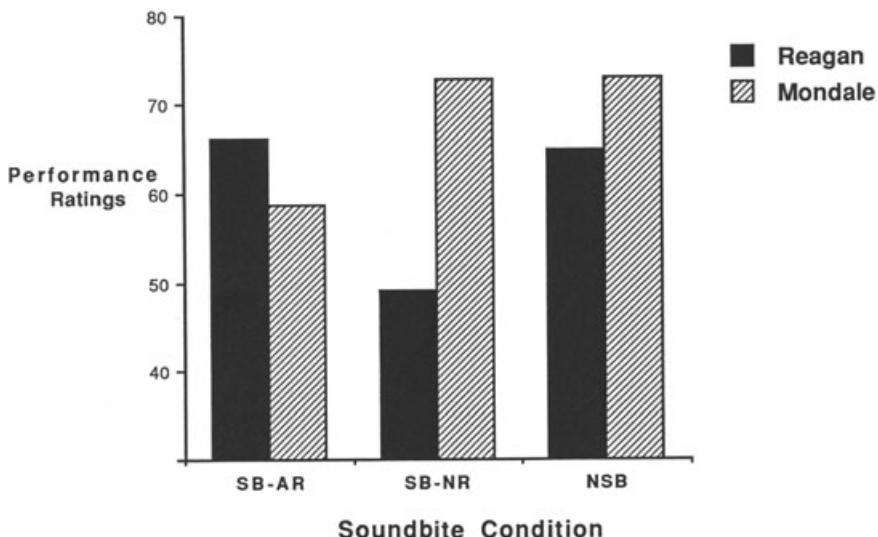


Figure 1. Post-debate performance ratings of Reagan and Mondale in Experiment 1 as a function of whether the clips contained the soundbites and the audience reaction (SB-AR), the soundbites with no audience reaction (SB-NR), or no soundbites (NSB).

back in 1984. Indicating the importance of the two brief soundbites, however, Figure 1 illustrates that this advantage enjoyed by Reagan disappeared and was even slightly reversed when the soundbites were deleted (Reagan's 7.50-point advantage in the control condition switched to a 7.63 disadvantage in the soundbite-deleted condition), $F(1, 50) = 6.82, p < .02$. Most interestingly, the reversal is even more pronounced in the condition in which the content of the soundbites was left in, but the audience reaction to them was deleted (Reagan's average rating was 23.57 points below Mondale's), $F(1, 50) = 24.29, p < .0001$. The results switched from an apparent victory for Reagan in the control condition to a decisive victory for Mondale in the condition in which the *content* was identical to the control condition but the audience reaction to the soundbites was eliminated.

Examining the ratings of each candidate independently, ratings of Reagan's performance varied significantly as a function of our manipulation, $F(2, 50) = 4.71, p < .02$. It is interesting to note that relative to the ratings in the control condition of Reagan's performance ($M = 66.25$), the ratings in the soundbite-deleted condition did not drop significantly ($M = 65.40, F < 1$), but the ratings in the reaction-deleted condition did decrease significantly ($M = 49.29, F(1, 50) = 7.89, p < .008$). It was the audience reaction to Reagan's soundbites, rather than the soundbites themselves, that made the critical difference in judgments of his performance.

Although the soundbites were made by Reagan, evaluations of Mondale were also significantly affected by our manipulations; in this case with participants rating Mondale more positively in the soundbite-deleted and reaction-deleted conditions than in the control condition ($M_s = 73.03, 72.86$, and 58.75 , respectively), $F(2, 50) = 8.59, p < .001$. The effects of the manipulation on ratings of Mondale's performance were similar conceptually to results from other research illustrating the implicit effects on a candidate whose opponent is the target of spin (Kugler & Goethals, 2005; Norton & Goethals, 2004).

Participants in this study also responded to the question of which candidate they thought won the debate. Here again the manipulation had a significant effect on participants' responses, $\chi^2 (2, N = 53) = 17.75, p < .002$. A majority of the participants in the control condition thought Reagan won the debate (60% versus 15% for Mondale and 25% judging that neither won), whereas most participants in the soundbite-deleted condition thought either neither won (42.11%) or Mondale won (42.11%), and, even more striking, a large majority in the reaction-deleted condition thought Mondale won (71.43% versus 7.14% for Reagan and 21.43% for neither).

For the question of who they would have voted for, there was not a significant difference between the control and the soundbite-deleted conditions, but there was a dramatic difference between these conditions and the reaction-deleted condition, with the percentage of participants indicating they would have voted for Reagan dropping from 51.28% to 14.29%, and those indicating they would have voted for Mondale increasing from 33.33% to 71.43% (and the remaining participants indicating no preference), $\chi^2 (2, N = 53) = 6.85, p < .04$.

Candidates' Qualities

The presence of the soundbites and audience reactions also affected perceptions of each candidate's personal character. On a composite measure that combined mean ratings of intelligence, likeability, leadership, sense of humor, strength, competence, and sincerity ($\alpha = .91$), Reagan was viewed less favorably in the reaction-deleted condition than in the control and soundbite-deleted conditions ($M_s = 61.22$ compared to 70.25 and 72.23 , respectively), $F(2, 50) = 2.81, p < .07$ —a pattern of results that was strongest on ratings of Reagan's sense of humor ($M = 83.50$ in the control condition, 75.79 in the soundbite-deleted condition, 63.21 in the reaction-deleted condition; $F(2, 50) = 6.11, p < .005$) and intelligence ($M_s = 67.75$ versus 71.16 versus 55.71 , respectively; $F(2, 50) = 4.67, p < .02$). As with overall performance ratings, evaluations of Mondale's personal character ($\alpha = .87$) were also significantly affected by soundbites and audience reactions, $F(2, 50) = 4.56, p < .02$ —particularly on ratings of his leadership, $F(2, 50) = 5.30, p < .005$. Remarkably, participants perceived Mondale as having significantly less leadership ability when they watched the original control debate

($M = 62.75$) than when they saw the tape without Reagan's soundbites ($M = 75.53$) or without positive audience reactions ($M = 75.00$).

Summary

Participants' ratings of the candidates' overall performance and specific qualities were affected significantly by our manipulation. Reagan was seen as the clear victor in the debate in the unedited, control condition, but this victory was lost with the deletion of the two brief, but apparently critical, soundbites. Deleting the reaction to the soundbites while keeping the soundbites intact tended to have a much bigger effect on participants' relative judgments of the candidates, illustrating the relative importance of the social context over the content of the soundbite.

Relative to the control conditions, ratings of Reagan tended to be affected much more by the deletion of the audience reaction than the deletion of the soundbite itself, whereas ratings of Mondale tended to be affected equally by these two conditions. Although it is only speculation, perhaps the content of the soundbites affected ratings of Mondale more because he was the target of them, particularly of the one in which Reagan mocked Mondale's record on defense. That soundbite may have presented a strong negative for Mondale, although it may not have counted for ratings of Reagan himself all that much. In any case, and most importantly, it is clear that the relative ratings of Reagan and Mondale were influenced strongly by both independent manipulations, and that, overall, the deletion of the audience reaction had a particularly strong effect. The content of Reagan's responses may have hurt perceptions of Mondale, but it was only with the strong validation of the responses by the audience and moderator that they made participants see Reagan much more positively.

Experiment 2

The results of Experiment 1 indicated that viewers of the 1984 debate between Reagan and Mondale were highly influenced not only by two remarks in the context of a substantive 40-minute debate excerpt, but also by others' reactions to those remarks. When the remarks were heard but the audience reactions taken out, the validation of these remarks as a significant "home run" for Reagan appeared to evaporate.

In light of the provocative implications of these results, a second experiment was designed to replicate these results and to address further three questions. One question is whether these effects would emerge in the context of the full 90-minute debate. With so much information presented in a full 90 minutes of debate, would the effects of two brief soundbites, or the reactions to these soundbites, be strong enough to influence participants' ultimate judgments? Moreover, we felt that using the full debate would increase the study's external validity. A second question

addressed in this experiment was whether participants perceived the soundbites to be memorable and noteworthy in the absence of audience reactions. That is, was the content of Reagan's remarks really effective and meaningful in the context of the entire debate, or did they become so in part because of the snowballing reaction of the moderator and audience? To examine this question, we added an item to the postdebate questionnaire asking the participants to list several key moments in the debate.

Third, we worried that having participants sit through an entire 90-minute-long debate would cause the participants to get antsy and stop paying attention after a while. We thought it might help to interrupt the tape halfway through and ask the participants some questions about the debate, and then resume the tape after that break. This both would provide a break and would remind the participants that we would be asking them questions about the debate. We worried, though, that this could reduce the external validity of the study, and, perhaps, have some carry-over effects on the final measures. To allay these concerns, we randomly assigned half the participants to watch the entire debate without interruption, whereas the other half were asked questions approximately midway through the debate.³

Method

Participants and Design

One-hundred and seven introductory psychology students participated in this experiment for course credit. Participants were randomly assigned to one of six cells produced by a 3 (control versus no soundbite-deleted versus no reaction-deleted) \times 2 (mid-debate ratings versus postdebate ratings only) factorial design.

Procedure

The procedure was similar to that of Experiment 1 with four exceptions. First, as explained in footnote 3, participants filled out the Need for Cognition Scale as part of their predebate questionnaire. Second, participants were shown a videotape

³ We originally had a fourth question in mind when designing this experiment: Would individual differences in need for cognition moderate our effects? Prior research suggested that audience reaction is a peripheral cue, influencing only participants who are low in involvement or in the need for cognition (Axsom, Yates, & Chaiken, 1987). This same study, however, did not find such a limitation on effects concerning participants' perceptions of a persuasive speaker, as distinct from that speaker's message. We administered Cacioppo and Petty's (1982) Need for Cognition Scale in Experiment 2 in order to determine whether individual differences in need for cognition would moderate our effects. As with the other measures of individual differences that we used in the various studies, we did not find any significant relationship between participants' scores on this measure and their ratings of the candidates.

of the entire 90-minute debate rather than the 40-minute excerpt previously used. Third, for half the participants, we stopped the debate at the halfway point and had the participants make ratings of the candidates' performance up to that point, explaining to them that these ratings were only their sense of how the debate was going thus far and that their final judgments may or may not be quite different from them. After a brief break, the rest of the debate was shown, after which participants filled out the complete set of ratings. Fourth, to examine whether participants were aware of the impact of the soundbites (i.e., for self-report purposes), they were asked on the final questionnaire to describe two or three "highlights" of the debate. We also shortened the final questionnaire a bit from the first study, in part to give the participants more time to complete this new question.

Results and Discussion

The manipulation of whether or not participants completed mid-debate ratings had no effect on any of the results. We therefore report only the main effects of the manipulation of the presence or absence of the soundbites and audience reaction.

Overall Performance

As in our previous study, the postdebate ratings of each candidate's performance constituted the primary dependent measure. Consistent with the results of Experiment 1, the relative ratings of the candidates in this study varied as a function of our manipulation. As can be seen in Figure 2, in this study Reagan's and Mondale's performances were judged virtually equally in the control condition, but Reagan tended to be rated more negatively than Mondale in the other two conditions—particularly in the condition in which the soundbite remained but the audience reaction was deleted.

The ratings of each candidate's performance were subjected to separate 3 (soundbite manipulation) \times 2 (obtaining measures twice versus only after the debate) ANOVAs. The ANOVA revealed a significant main effect for the soundbite manipulation on ratings of Reagan, $F(2, 101) = 3.12, p < .05$. More specifically, Reagan received significantly higher ratings in the control condition than in the reaction-deleted condition ($M_s = 59.32$ & 48.09, respectively; $p < .05$ via Newman-Keuls test; $M = 52.81$ in the soundbite-deleted group). In contrast to Experiment 1, however, ratings of Mondale's performance were not significantly affected by the manipulation, $F < 1$.

Candidates' Qualities

As in Experiment 1, our manipulation influenced ratings of Reagan's character. On the same composite measure used in our prior studies (combined ratings of

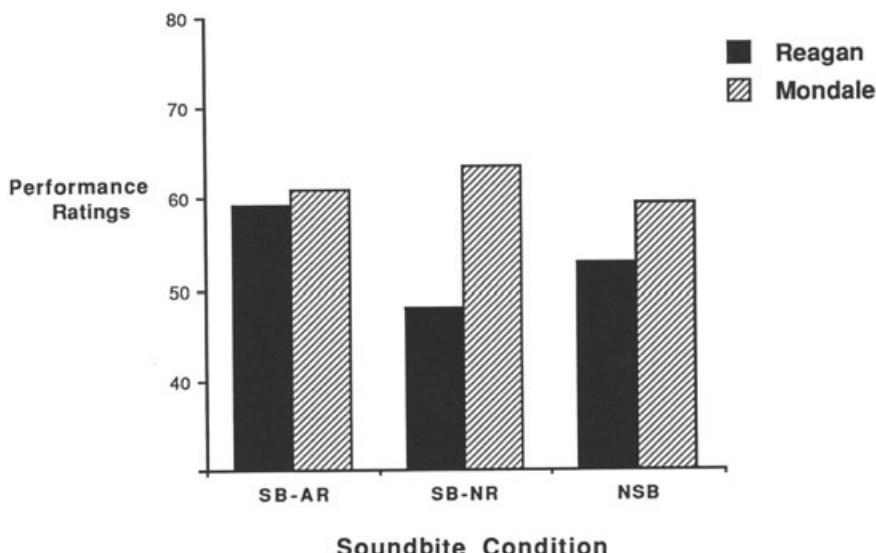


Figure 2. Postdebate performance ratings of Reagan and Mondale in Experiment 2 as a function of whether the clips contained the soundbites and the audience reaction (SB-AR), the soundbites with no audience reaction (SB-NR), or no soundbites (NSB).

intelligence, sense of humor, likeability, leadership, strength, competence, sincerity), there was a significant main effect, $F(2, 101) = 3.21, p < .05$, as Reagan was viewed more favorably in the control condition than in the reaction-deleted condition ($M_s = 65.94$ and 55.85 , respectively; $p < .05$ via Newman-Keuls test; $M = 62.03$ in the soundbite-deleted condition). This pattern was particularly strong in ratings of Reagan's intelligence, sense of humor, and likeability. Once again in contrast to the results of Experiment 1, the effect of the soundbites on evaluations of Mondale's character was not statistically significant, $F(2, 101) = 2.79, p < .10$.

Were Soundbites Perceived as Significant?

Finally, participants were asked on the postdebate questionnaire to describe two or three highlights from the debate. These free-response data were coded for whether or not the participants included a soundbite on the list. The result was informative. Demonstrating the profound importance of audience reactions, 78% of all participants in the control condition included a soundbite among the highlights, compared to only 15% of those in the reaction-deleted condition, $X^2(2, N = 74) = 29.17, p < .001$. In short, participants were impressed by the soundbites only when the remarks were followed by the strong reactions from the audience.

Summary

Supporting the results of the first experiment, participants in this study tended to offer very different relative ratings of the two candidates as a function of the presence or absence of the soundbites and the reaction to them. Although the results are more muted than in the first study, perhaps due to the extra 50 minutes of material—much of it quite dry—diluting the impact of these critical moments, it was still the case that participants’ relative judgments were affected significantly. Whereas participants in the control condition perceived a very close debate, those in the condition in which the reactions to Reagan’s two soundbites were deleted—representing less than a minute of time in a 90-minute debate about some of the most important issues facing the country—saw Mondale clearly outperforming Reagan. Moreover, absent the applause, laughter, and general approval of Reagan’s one-liners, these responses were not seen as particularly noteworthy by the participants. With these audience reactions intact, however, the majority of the participants saw them as a highlight of the debate.

Experiment 3

Although pivotal soundbites stand out remarkably well in debate lore, they are rare, and often debates are won or lost in their absence. In the past several election cycles it is hard to locate defining moments in debates, even in debates with clear winners. The first Bush-Kerry debate of 2004, for example, was judged a clear Kerry win despite its lack of “zingers.” Kerry’s successes were seen holistically (as were his opponent’s foibles). The same could be said of Kennedy’s performance in the first debate of 1960. Therefore, conformity should also be examined in the—arguably more common—context of these victories by attrition.

Experiment 3 used segments of the 1984 debate in that light using a 10-minute tape that excludes the critical soundbites from the previous two experiments. Unlike in Experiments 1 and 2, this study did not use studio audience behavior as the independent variable. Rather, participants received false feedback concerning the alleged reactions of their peers in the room with them as they watched the debate together in a group.

The feedback was administered using a technology first displayed during the televised debates of 1992. For those events, the cable television news network, CNN, showed viewers continuous, real-time responses of focus-group members who were watching the debate in an auditorium and using hand-held response dials to record their changing impressions of the event. This computer-based technology thus provided TV viewers with real-time, public-opinion poll data. We used this technology to consistently insert audience feedback into each participant group, with participants being told that the displayed “data” were their own. This procedure enabled us to assess whether judgments of the candidates would be influenced by exposure to the group norm.

Method

Participants and Design

Ninety-four introductory psychology students participated in this study for course credit. Participants were run in groups of approximately 15–20 students. The groups were randomly assigned to either a pro-Reagan or pro-Mondale false feedback condition.

Procedure

The stimulus tape used in this study was a 10-minute segment consisting of excerpts of the second 1984 presidential debate between Reagan and Mondale. Upon their arrival at the laboratory, participants completed a brief predebate questionnaire and were then each given a wireless hand-held dial. Each dial was equipped with a digital display that indicates the current numerical setting—in this case, it was set to range from 0 (with the dial hand pointed to the far left) to 100 (with the dial hand pointed to the far right), and having a midpoint of 50 (with the hand pointed up at a 90-degree angle). Participants were told that they would see videotaped excerpts of a presidential debate and were instructed to set the dial at 50 and move it up or down to varying degrees to indicate their changing opinions during the debate. Before seeing the debate, participants got some practice using their dials, which also served to allay potential suspicions they may have had about the validity of these dials—as the experimenters demonstrated how they could tell what each participant's dial was set to at any given moment.

Participants in both groups were informed that they would receive continuous real-time feedback about their group's average opinion in the form of a line graph superimposed over the debate videotape. In fact, the feedback they received was false, pre-programmed to indicate that Reagan or Mondale was gaining in support over the other. In both groups, participants saw the same line graph begin at the neutral midpoint of 50 and move gradually, in a fluctuating pattern, to a final value of 85. For those in the pro-Reagan group, the 0–100 point response scale was defined in such a way that 50 meant that neither candidate was outperforming the other, that numbers higher than 50 meant that Reagan was outperforming Mondale, and that numbers lower than 50 meant that Mondale was outperforming Reagan; for those in the pro-Mondale condition, the scale was reversed. To guard against the possibility that a participant would turn his or her dial down to 0 or up to 100 and become suspicious of the false feedback manipulation when it appeared not to affect the group's average, participants were told that the most extreme high and low scores would always be excluded from the computation.⁴

⁴ It perhaps should be noted that participants found these dials to be easy to use and did not find them, or the superimposed graph, to be particularly distracting. The first author has used this technology in

After watching the tape, participants completed the same postdebate questionnaire as that used in the first experiment.

Results and Discussion

Overall Performance

Supporting the results of the first two studies, the results of Experiment 3 indicated that others' reactions can have a significant effect on individuals' judgments of the candidates' overall performance and personal qualities. As can be seen in Figure 3, participants who saw a graph suggesting that their peers saw Reagan as winning the debate rated Reagan's performance more than 15 points better than Mondale's, whereas they rated Reagan's performance more than 20 points *worse* than Mondale's if they saw a graph suggesting that their peers saw Mondale perform better, resulting in a net difference of about 36 points, $F(1, 92) = 51.74, p < .0001$.

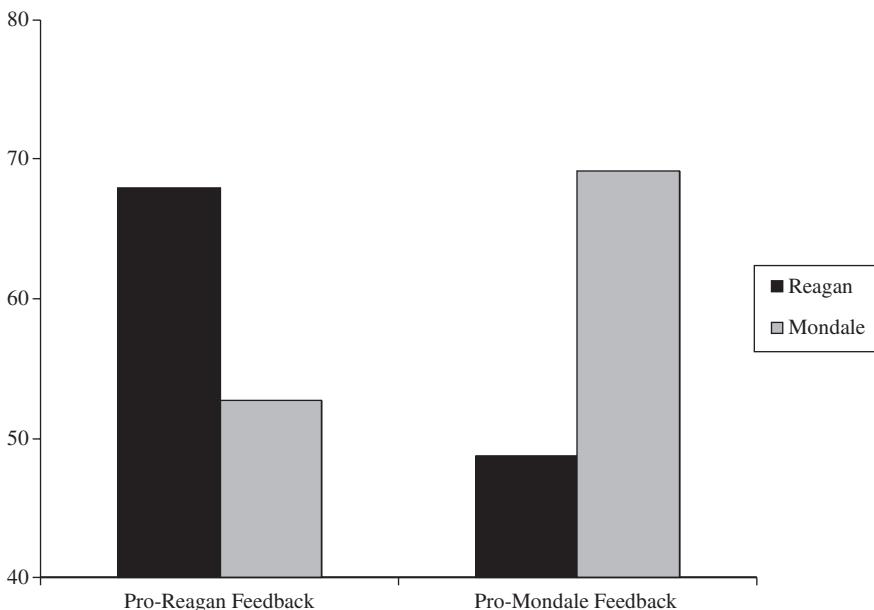


Figure 3. Postdebate performance ratings of Reagan and Mondale in Experiment 3 as a function of whether participants were exposed to pro-Reagan or pro-Mondale feedback.

several studies in subsequent years (e.g., Fein, Hoshino-Browne, Davies, & Spencer, 2003), and it is extremely rare for any participants to be suspicious of or very distracted by it.

On ratings of each candidate's debate performance, the results again supported the hypothesis that participants would be influenced by the false normative feedback concerning the opinions of others in the room. Reagan's performance ratings were significantly higher in the pro-Reagan feedback condition than in the pro-Mondale condition ($M_s = 67.98$ versus 48.81), $F (1, 92) = 31.67, p < .001$. In contrast, Mondale's ratings were significantly higher in the pro-Mondale condition than in the pro-Reagan condition ($M_s = 69.17$ versus 52.69), $F (1, 92) = 30.03, p < .0001$.

A chi-square analysis on participants' responses to the questions of who won the debate (Reagan, Mondale, or neither) yielded additional support for our hypothesis. Participants in the pro-Reagan condition were significantly more likely to judge Reagan to be the winner of the debate (67.31% versus 7.69% for Mondale and 25% for neither) than were participants in the pro-Mondale condition (11.91% versus 71.43% for Mondale and 16.67% for neither), $\chi^2 (2, N = 94) = 43.61, p < .0001$.

Further indicating the strength of the audience reaction effect was that there was also a significant and sizable difference on the key question, "who do you think you would have voted for?" (Reagan, Mondale, or neither), $\chi^2 (2, N = 94) = 9.88, p < .008$. Participants in the pro-Reagan condition were significantly more likely to indicate they would have voted for Reagan (46.15% versus 34.62% for Mondale and 19.23% for neither) than were participants in the pro-Mondale condition (16.67% versus 61.91% for Mondale and 21.43% for neither).

Candidates' Qualities

As in the first two studies, audience reaction also had a marked effect on perceptions of each candidate's personal characteristics. On a composite measure that combined mean ratings of intelligence, likeability, leadership, sense of humor, strength, competence, and sincerity ($\alpha = .91$), participants evaluated Reagan more favorably in the pro-Reagan feedback group than in the pro-Mondale group ($M_s = 67.02$ versus 58.57), $F (1, 92) = 7.18, p < .009$. Similarly, Mondale ($\alpha = .86$) was rated more favorably in the pro-Mondale condition than in the pro-Reagan group ($M_s = 65.50$ versus 56.58), $F (1, 92) = 14.52, p < .0003$. Indicating the broad range of effects produced by audience reaction, individual analyses revealed that the two groups differed significantly in their ratings of Reagan's likeability, leadership, and sense of humor, and in their ratings of Mondale's likeability, leadership, intelligence, competence, strength, and sincerity (all at $p < .05$). On estimates of how well the two candidates fared among voters in general, audience reaction also affected mean estimates of Reagan's popularity ($M_s = 74.23$ and 57.98 , respectively, in the pro-Reagan and pro-Mondale groups), $F (1, 92) = 35.58, p < .0001$, as well as Mondale's ($M_s = 61.55$ and 44.33 , respectively, in the pro-Mondale and pro-Reagan groups), $F (1, 92) = 37.56, p < .0001$.

Self-Reported Effects of Feedback

Finally, participants rated the extent to which their evaluations of the candidates were influenced by the online normative feedback (i.e., on a 0–100 point scale, whether it led them to favor Reagan or Mondale, where 50 = no effect). Interestingly, the pro-Reagan and pro-Mondale conditions did not differ on this measure, $F(1, 92) = 1.47, ns$. With a combined mean rating of 51.20, it appears that participants in general believed that they were not influenced by the feedback. This self-reported lack of influence thus contrasts sharply with all of the above results.

Summary

The results of Experiment 3 indicated that evaluations of the 1984 presidential candidates were significantly influenced by fabricated real-time feedback in the form of a line graph allegedly indicating the opinions of their peers in the room with the participants. For example, seeing the graph of their peers' alleged reactions changed the percentage of participants who thought Mondale had, in their own minds, won the election from fewer than 8% to more than 71%. Overall, these results, which were striking in their breadth and magnitude, suggest that the practice of presenting TV viewers with continuous focus-group data (or, as can be expected in upcoming elections, the practice of seeing others' real-time reactions on the internet as people watch the debate) constitutes a powerful source of social influence information.

Experiment 4

The previous experiments all showed significant conformity effects. However, the 1984 presidential election took place years before these studies were run. The participants knew the outcome of the election, and their concern about the accuracy of their judgments might not have been particularly high. One might presume that when citizens are watching an actual debate live, and sizing up the candidates in order to decide how to vote, they may be much less influenced by the apparent evaluations of others. Furthermore, there are other factors that might reduce conformity effects in this situation—viewers may have well-formed impressions of candidates, been subjected to months of predebate spin, been influenced more strongly by personal biases, etc. In short, there may be something fundamentally different about a debate preceding an election in which one will actually vote. Experiment 4 was intended to address this concern by having research participants evaluate a presidential debate on the night it actually took place, during the height of the campaign. If strong conformity effects were found under these circumstances, a host of potential concerns relating to external validity could be dismissed.

Participants and Design

Sixty-one introductory psychology students participated in exchange for course credit. Participants were randomly scheduled to appear in one of three classrooms to watch the third and final presidential debate of that year's campaign. By random assignment, one room was designated as pro-Bush, a second as pro-Clinton, and a third as neutral.⁵

Procedure

The stimulus tape used in this study was the third and final 1992 presidential debate involving George Bush, Bill Clinton, and Ross Perot. Upon recruitment, all participants were asked to refrain from watching the debate live that evening, as they would see it just minutes after its conclusion as part of the study (indeed, they were scheduled to come to the lab several minutes before the debate would conclude). This was a mere 15 days prior to the election. As soon as they arrived at their assigned room, all participants completed a brief predebate questionnaire. The entire 90-minute debate tape was then played simultaneously in the three rooms. Afterward, all participants were administered a postdebate questionnaire, fully debriefed, and thanked for their participation.

In this study, social influence was manipulated via audience reactions. In a *pro-Bush* group, 21 participants watched the debate in the company of 13 student confederates who had been instructed to quietly but audibly applaud statements by Bush and disapprove (hiss, jeer) statements by Clinton. In a *pro-Clinton* group, 20 participants were joined by 12 confederates who were told similarly to applaud Clinton and disdain Bush. All confederates were rehearsed to react in ways that seemed natural. Their reactions increased gradually during the course of the debate, but they were cautioned not to interfere with the participants' ability to hear the tape. In a third *control group*, there were 20 participants and no confederates.

The dependent measures were similar to those used in the previous experiments, with the notable exception that in this study, at the end of the postdebate questionnaire participants were asked also to estimate the mean 0–100 point rating that each candidate would receive (1) from the other participants in the room, and (2) from viewers all over the country. The participants were asked to indicate which of the candidates, if any, they planned to vote for on Election Day.

⁵ Although he also participated in the debate, third-party candidate Ross Perot was not considered to have a reasonable chance of even coming close to winning the 1992 election, and so we focused our manipulations on the two major candidates.

Results and Discussion

Overall Performance

As shown in Figure 4, the effect of audience reaction on the relative performance ratings of Clinton and Bush was striking. In the control group, Clinton's rating was 23 points higher than Bush's (and, indeed, Clinton was seen as the convincing winner of the debate by analysts and opinion polls). Yet the difference was up to 51 points in the pro-Clinton group and down to only 6 points in the pro-Bush group, $F(2, 58) = 20.59, p < .0001$.

On ratings of each candidate's debate performance, the results indicated that the effects of audience reaction were significant. Specifically, Clinton's ratings were higher among participants in the pro-Clinton audience ($M = 82.75$) than in the pro-Bush ($M = 67.14$) and control groups ($M = 69.50$), $F(2, 58) = 12.99, p < .0001$. In contrast, Bush's ratings were higher in the pro-Bush audience ($M = 60.95$) than in the pro-Clinton ($M = 31.50$) and control groups ($M = 46.25$), $F(2, 58) = 9.94, p < .0002$.

Participants' postdebate voting preferences were relatively consistent with these results, but the chi-square did not approach significance, $X^2 (8, N = 61)$

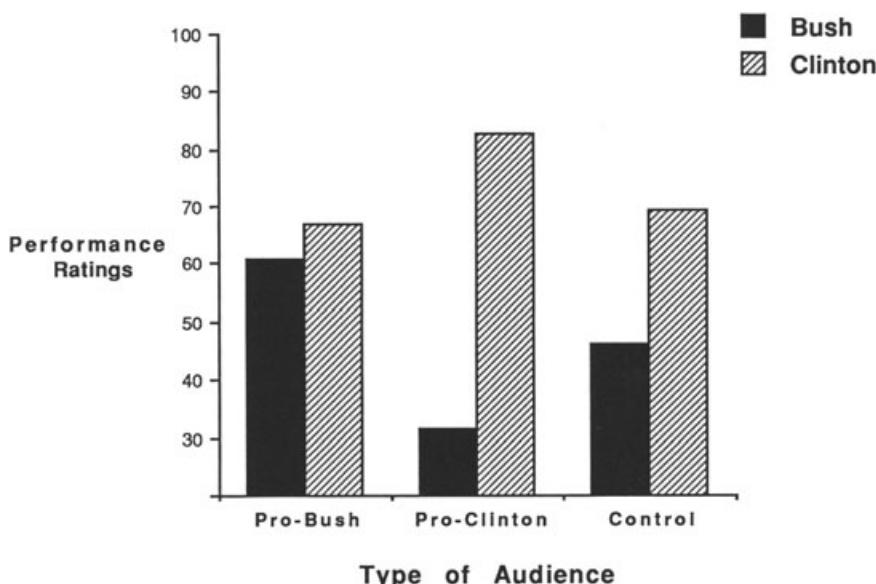


Figure 4. Post-debate performance ratings of Bush and Clinton in Experiment 4 as a function of whether confederates were exposed to pro-Bush, pro-Clinton, or no (Control) confederates.

= 6.95, *ns*, perhaps in part because there were five rather than only two options for participants to choose—for Clinton, for Bush, for Perot, did not know, or did not plan to vote.

There were two additional and interesting results on these overall evaluations of the candidates. First, questionnaires were administered to the 13 pro-Bush and 12 pro-Clinton confederates, and an analysis of their responses indicated that they too were significantly influenced by audience reaction in their ratings of both Bush, $F(1, 23) = 10.91, p < .01$, and Clinton, $F(1, 23) = 6.93, p < .02$. Despite knowing that audience reaction was systematically manipulated—indeed, despite the fact that they themselves had to vary their behavior according to instruction—the confederates exhibited the same effect. Second, although we manipulated audience reactions only to Bush and Clinton, ratings of Perot's performance were also affected somewhat by the manipulation, $F(2, 58) = 2.63, p < .09$. Reflecting the unanticipated fact that Perot had aligned himself more with Clinton in this debate than with Bush, ratings of Perot were higher in the pro-Clinton group ($M = 76.75$) than in the pro-Bush or control conditions ($M_s = 65.00$ and 66.10 , respectively), $F(1, 58) = 5.21, p < .02$.

Candidates' Qualities

On the more specific evaluation measures, the results followed a similar pattern as the overall performance ratings. Combining each candidate's performance ratings on the economy, foreign affairs, education, health care, and crime, results showed that Bush's mean ratings ($\alpha = .83$) were higher in the pro-Bush group than in the other conditions, $F(2, 58) = 3.70, p < .05$, although no comparable difference was found in ratings of Clinton on this same constellation of issues ($F < 1$). It is interesting that although audience reactions had no effect on the Clinton composite measure ($\alpha = .73$), participants did rate Clinton more favorably on economic issues in the pro-Clinton group than in the other conditions, $F(2, 58) = 5.17, p < .01$. In fact, the effect of audience reaction on ratings of Bush was also strongest on his economy performance, $F(2, 58) = 9.41, p < .001$. Although this pattern was not predicted, the likely reason for it is that the debate focused primarily on economic issues.

Overall, audience reaction effects were even more pronounced on ratings of each candidate's personal characteristics. Participants in the pro-Bush condition, relative to those in the pro-Clinton and control groups, saw Bush as more trustworthy ($F(2, 58) = 3.60, p < .05$), more intelligent ($F(2, 58) = 9.64, p < .001$), more caring ($F(2, 58) = 3.51, p < .05$), and as having a better sense of humor ($F(2, 58) = 3.22, p < .05$). Similarly, participants in the pro-Clinton condition relative to those in other groups saw Clinton as more trustworthy ($F(2, 58) = 5.64, p < .005$), more intelligent ($F(2, 58) = 9.59, p < .001$), more likeable ($F(2, 58) = 4.91, p < .01$), and more mature ($F(2, 58) = 2.98, p < .05$).

Estimates of Others' Perceptions

Finally, participants were asked to estimate the mean ratings that each candidate would receive from the other participants in the room, and from all viewers nationwide. On the first of these measures, the effect was striking in its magnitude, as participants estimated the highest Bush rating in the pro-Bush group and the lowest in the pro-Clinton group ($M_s = 59.76$ and 11.50 , respectively; $M = 41.00$ in the control condition), $F (2, 58) = 51.50, p < .0001$. The data were similar for ratings of Clinton, as participants estimated his highest rating in the pro-Clinton group and his lowest in the pro-Bush group ($M_s = 89.75$ and 62.38 , respectively; $M = 73.75$ in the control condition), $F (2, 58) = 33.20, p < .0001$. Essentially, these results served as a check on the audience reaction manipulation. However, there was also a highly significant effect on estimates of how the candidates fared more generally among viewers across the country. On this measure as well, those in the pro-Bush group projected the rest of the country to be much more positive about Bush's performance ($M = 60.38$) than did those in the control ($M = 53.75$) or pro-Clinton conditions ($M = 40.00$), $F (2, 58) = 7.85, p < .001$. Similarly, those in the pro-Clinton group projected the rest of the country to be much more positive about Clinton's performance ($M = 78.25$) than did those in the control ($M = 69.50$) or pro-Bush conditions ($M = 63.33$), $F (2, 58) = 11.53, p < .0001$. Thus, even though all participants saw the same debate (which included an actual live audience), they inferred that the reactions of the confederates in their sessions were diagnostic of national public opinion.

Summary

The results from Experiment 4 revealed that manipulated peer audience reactions to a presidential debate strongly influenced participants' judgments of each candidate's overall performance, performance on the focal substantive issue, and personal characteristics, even under the most realistic circumstances possible. The magnitude of this effect was astonishing (it produced a 45-point shift in the relative ratings of Bush and Clinton), and was obtained not only with naive participants who were low in political involvement, but with high-involvement participants—and even confederates who knew that the cheers and jeers were experimentally manipulated. This effect occurred even though the stimulus event was a heated 90-minute presidential debate and despite the presence of a live, reactive (i.e., at the debate) audience. Indeed, our manipulation led participants not only to alter their own judgments of the candidates, but to infer that these judgments were diagnostic of national public opinion, as was also seen in Experiment 3.

General Discussion

The results of the four studies reported here were highly consistent. In each of them participants' judgments about debate performance were strongly influenced by their mere exposure to other people's reactions. The magnitude of these effects is striking. Our participants made *very* different judgments about debate performance as a function of the information we manipulated concerning the reactions of others in the audience. In Experiment 1, participants watched Mondale and Reagan. In one condition the audience reaction to Reagan's winning "one-liners" was removed. Compared to the control condition where the audience reaction was presented intact, Reagan's ratings relative to Mondale's declined by 31 points. The results of Experiment 2 were also significant despite the fact that the manipulation was embedded within the full 90-minute debate. In Experiments 3 and 4, participants saw what they believed to be the reactions of their peers who were watching the debate with them. In Experiment 3, as a function of this false group feedback, Reagan's ratings relative to Mondale's varied by 30 points. In Experiment 4, participants' judgments about overall performances in a 1992 Bush-Clinton-Perot debate, on the night of the live debate itself, were strongly influenced by whether they watched the debate with pro-Clinton or pro-Bush confederates who audibly vocalized their opinions. On a 101-point scale, Clinton's advantage over Bush increased by 45 points across conditions. Taken together, these studies all demonstrated that mere exposure to the judgments of others strongly influenced our participants.

One striking feature of the results is that they were highly similar across a number of dimensions. First, information about others' judgments was presented in a variety of ways: the recorded laughter and cheers of audiences attending actual debates, the visual fabricated feedback of other group members, or the vocal reactions of live confederates. Second, audience reactions significantly affected perceptions of candidate performance across a broad range of participants, including Democrats and Republicans, those who defined themselves as politically involved or uninvolved, and those who were high or low in need for cognition. In each case there were similar effects for the different kinds of participants.⁶ Third, effects were obtained across a range of dependent measures, including overall performance ratings as well as perceptions of specific qualities such as intelligence and likeability. Fourth, participants' judgments about a presidential debate in a current campaign cycle were influenced as much as, or even slightly more than,

⁶ It should be noted that the participants in our studies all were college undergraduates, and therefore we cannot know whether older samples of people would respond similarly. We have conducted one study in a different line of work concerning social context and presidential debates in which our participants came from a large sample of public-school teachers, and the results were similar to those we found with undergraduate samples. It is clear, though, that including older and broader samples of participants should be an important goal for future research.

their judgments about debates that took place nearly a decade earlier. The consistency of the influence effects is striking.

Our participants seemed to exhibit informational social influence and internalized conformity, reflecting, in Insko, Drenan, Solomon, Smith, and Wade's (1983) terms, a desire to be right rather than a desire to be liked. Their judgments were given privately and anonymously. In Experiments 1 and 2 the source of influence was simply a videotaped audience from a debate which had taken place years earlier. In those cases, participants had no reason to be concerned about pressure from others or others' opinions about them. Importantly, then, the influence we obtained seems to be informational social influence affecting privately held judgments. This is precisely the kind of influence one would expect in a situation where physical reality is ambiguous but social reality is clear. Our participants probably believed that there was in fact a right answer to questions about how the candidates performed—and that the reactions of others signaled that clear answer.

What is a likely account of the process by which informational social influence was produced in these studies? It may be that the identification and interpretation of specific behaviors in debates were influenced by others' reaction. When Ronald Reagan responded to a very serious question about his energy and intellectual capacity by joking that he would not exploit his opponent's youth and inexperience, there was room for considerable interpretation. Observers could identify his reply as a feeble or even disgraceful attempt to dodge the issues or as a deft and masterful turning of a hostile question to his advantage. The audience reaction of strong approval did not *recognize* Reagan's quip as a knockout punch so much as it *made* it one. Throughout a 90-minute debate there are numerous segments of behavior (Newton, 1976), each of which can be framed and defined by the reactions of others. Furthermore, once a particular expectation is adopted about how each candidate is performing in a debate, individual behaviors can be interpreted within that framework (Fiske & Taylor, 1991; Kelley, 1950; Kugler & Goethals, 2005; Norton & Goethals, 2004). Positive or negative moments can therefore define entire debates, or even campaigns.

It is perhaps worth underlining the fact that the audiences who produced the influence in these studies were most likely perceived as having no intent to influence. This is most clearly the case in Experiments 1 and 2 where the audience's reactions were televised years before. Earlier studies of overheard communications suggest that they are perceived as not being designed to persuade and for this reason often have more credibility and influence than communications delivered with intent to persuade (Brock & Becker, 1965). They also generate less psychological reactance (Brehm, 1966).

What are the policy issues that are suggested by this research? Perhaps the most fundamental is that those who stage, sponsor, and broadcast political debates should take steps to minimize social influence, or psychologists should take steps to publicize that influence, or both. Within the current debate format there are at

least three sources of influence that are under virtually no control. The first is the reactions of live audiences. Since 1984 live audience reaction has been part of the mix. In the vice-presidential debate in 1988 Lloyd Bentsen partisans laughed and cheered at his “You’re no Jack Kennedy” put-down of Dan Quayle, while the latter’s supporters competed with a loud chorus of boos. In several of the recent debates, the moderator has strongly admonished the audience to remain quiet. Since audience reaction has such a strong effect on viewers, this warning is well justified.

A second possible source of influence is the behavior of the commentators who moderate debates and ask questions. In 1984 journalist Henry Trewhitt commented on Reagan’s age quip by saying “I’ll try to run to the fence to catch that one before it goes out,” just in case anyone missed the fact that Reagan had hit a home run. In 1988, Jim Lehrer implicitly supported Bush’s attempts to make light of his persistent errors by quoting and endorsing the vice-president’s “nobody’s perfect” response to Michael Dukakis’ derisive response to Bush’s misstatements. Also in 1988, NBC newsman Tom Brokaw’s headshakes and nods influenced Dan Quayle’s response as he walked into Lloyd Bentsen’s “You’re no Jack Kennedy” put-down. Broadcasters should exercise caution and restraint over the subtle ways in which they guide, diminish, or enhance a debater’s remarks.

Third, networks now report on the results of instant polls and focus groups that may have a significant effect on people’s judgments about how candidates performed and who won. Some networks also give high-ranking campaign officials free reign to apply their “spin” to the evening’s debate performances. All this information is consumed by viewers before they have had a chance to consolidate, deliberate, and make up their own minds about the large amount of content they have just seen. The potential for influence is therefore all the stronger. In this regard, the news media might be more thoughtful about their role in channeling social realities to viewers. While the news media may have no conscious intent to influence, our research suggests that they may have great influence nonetheless. Their good intentions aside, the power of the networks needs to be considered.

We can expect in future debates that communication via the internet will play an increasingly important role. It is easy to imagine that very large numbers of citizens, particularly younger voters, will provide spontaneous audience reactions, along the lines of what we manipulated in Experiments 3 and 4, by posting to blogs, electronic forums and boards, and so on. The normative information that will be provided in this manner has tremendous potential to influence the perceptions and decisions of many individuals. When information such as this, or such as the immediate postdebate spins presented by analysts and campaign officials, bombards individuals before they have had the chance to reach their own interpretations and conclusions independently, individual voters may be unaware that the privacy of their voting booth has been compromised by the long reach of the opinions of a great many friends and strangers.

In conclusion, we have shown that people's judgments of presidential debates can be strongly influenced by their knowledge of other people's reactions. These debates appear to be highly ambiguous events. For people to develop their own independent opinions, great care must be taken on the part of organizers and sponsors to limit the impact of irrelevant cues like the ones discussed herein. One potentially useful direction for future research is to examine the factors that would make people less vulnerable to such effects. Would being educated about the results of studies such as the present ones reduce individuals' susceptibility to the kind of informational social influence we have documented? Of course, other people can and often do provide important, useful, and even correct information. In the theater of political campaigns, however, with spin doctors, hand-picked audiences, and highly charged emotions galore, this may be less likely to be the case than is usual. The often-adaptive practice of checking in with social reality can be a trap in these orchestrated contexts. What is especially important, therefore, is for individuals to look at the social context around these debates with vigilance and caution.

ACKNOWLEDGMENTS

We would like to thank Amelia Cottrell, Jessica Cross, Leigh Frost, C.J. Gillig, and Lauren Parkhill for their help in collecting the data presented in this paper. We would also like to thank Columbia Information Services for their assistance with the Perception Analyzer technology used in Experiment 3. Correspondence concerning this article should be addressed to Steven Fein, Department of Psychology, Williams College, Williamstown, MA 01267, or Al Goethals, Jepson School of Leadership Studies, University of Richmond, Richmond, VA 23173. E-mail: sfein@williams.edu and ggoethal@richmond.edu.

REFERENCES

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership, and men*. Pittsburgh: Carnegie Press.

Axsom, D., Yates, S., & Chaiken, S. (1987). Audience response as a heuristic cue in persuasion. *Journal of Personality and Social Psychology*, 53, 30-40.

Baron, R. S., Vandello, J. A., & Brunsman, B. (1996). The forgotten variable in conformity research: Impact of task importance on social influence. *Journal of Personality & Social Psychology*, 71, 915-927.

Brehm, J. W. (1966) *A theory of psychological reactance*. New York: Academic.

Brock, T. M., & Becker, L. A. (1965). Ineffectiveness of "overheard" counterpropaganda. *Journal of Personality and Social Psychology*, 2, 654-660.

Cacioppo, J. T. & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42, 116-131.

Cantril, H. (1940). *The invasion from Mars*. Princeton: Princeton University Press.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal & Social Psychology*, 51, 629–636.

Fein, S., Frost, L. A., Goethals, G. R., & Kassin, S. M. (1994). *The effects of expectations on perceptions of presidential debate performance*. Presented at the 102nd meeting of the American Psychological Association, Los Angeles.

Fein, S., Hoshino-Browne, E., Davies, P. G., & Spencer, S. J. (2003). Self-image maintenance goals and sociocultural norms in motivated social perception. In S. J. Spencer, S. Fein, M. Zanna, & J. M. Olson (Eds.), *Motivated social perception: The Ontario symposium* (vol. 9, pp. 21–44). Mahwah, NJ: Erlbaum.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill.

Germond, J. W., & Witcover, J. (1989). *Whose broad stripes and bright stars? The trivial pursuit of the presidency 1988*. New York: Warner.

Insko, C. A., Drenan, S., Solomon, M. R., Smith, R., & Wade, T. J. (1983). Conformity as a function of the consistency of positive self-evaluation with being liked and being right. *Journal of Experimental Social Psychology*, 19, 341–358.

Jamieson, K. H., & Waldman, P. (2002). *The press effect: Politicians, journalists, and the stories that shape the political world*. Cambridge: Oxford University Press.

Kaid, L. L., & Bystrom, D. G. (1999). *The electronic election: Perspectives on the 1996 campaign communication*. Mahwah, NJ: Erlbaum.

Kelley, H. H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, 18, 431–439.

Kiesler, C. A., & Kiesler, S. B. (1969). *Conformity*. Reading, MA: Addison-Wesley.

Kraus, S. (1962). *The great debates: Background, perspectives, effects*. Bloomington: Indiana University Press.

Kugler, M. B., & Goethals, G. R. (2005). *Trait-focused spin in presidential debates: Surviving the kisses of death*. Unpublished manuscript, Williams College.

McKinnon, L. M., & Tedesco, J. C. (1999). The influence of medium and media commentary on presidential debate effects. In L. L. Kaid & D. G. Bystrom (Eds.), *The electronic election: Perspectives on the 1996 campaign communication* (pp. 191–206). Mahwah, NJ: Erlbaum.

McKinnon, L. M., Tedesco, J. C., & Kaid, L. L. (1993). The third 1992 presidential debate: Channel and commentary effects. *Augmentation and Advocacy*, 30, 106–118.

Newtonson, D. (1976). Foundations of attribution: The perception of ongoing behavior. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (pp. 223–247). New York: John Wiley and Sons.

Norton, M. I., & Goethals, G. R. (2004). Spin (and pitch) doctors: Campaign strategies in televised political debates. *Political Behavior*, 26, 227–248.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Schroeder, A. (2000). *Presidential debates: Forty years of high-risk TV*. New York: Columbia University Press.

Sears, D. O., & Chaffee, S. H. (1979). Uses and effects of the 1976 debates: An overview of the empirical studies. In S. Kraus (Ed.), *The great debates: Carter versus Ford, 1976*. Bloomington: Indiana University Press.

Sherif, M. (1936). *The psychology of social norms*. New York: Harper.

Sigelman, L. (1992). There you go again: The media and the debasement of American politics. *Communication Monographs*, 59, 407–410.

Sigelman, L., & Sigelman, C. K. (1984). Judgments of the Carter-Reagan debate: The eyes of the beholders. *Public Opinion Quarterly*, 48, 624–628.

Steeper, F. T. (1978). Public response to Gerald Ford's statements on Eastern Europe in the second debate. In G. F. Bishop, R. G. Meadow, & M. Jackson-Beeck (Eds.), *The presidential debates: Media, electoral, and policy perspectives* (pp. 81–101). New York: Praeger.

Zarevsky, D. (1992). Spectator politics and the revival of public argument. *Communication Monographs*, 59, 411–414.

Copyright of Political Psychology is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.