My Friend's Enemy: How Split-Screen Debate Coverage Influences Evaluation of Presidential Debates
Dietram A. Scheufele, Eunkyung Kim and Dominique Brossard
Communication Research 2007 34: 3
DOI: 10.1177/0093650206296079

The online version of this article can be found at:
http://crx.sagepub.com/content/34/1/3

Published by:
http://www.sagepublications.com

Additional services and information for Communication Research can be found at:

Email Alerts: http://crx.sagepub.com/cgi/alerts

Subscriptions: http://crx.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://crx.sagepub.com/content/34/1/3.refs.html
My Friend’s Enemy
How Split-Screen Debate Coverage Influences Evaluation of Presidential Debates
Dietram A. Scheufele
Eunkyung Kim
Dominique Brossard
University of Wisconsin–Madison

Leading up to the 2004 presidential debates, there was considerable discussion about the mode of presentation on television, that is, debate coverage with simultaneous reaction shots of the opponent while a candidate was speaking or coverage with isolated shorts of each candidate only. In fact, many commentators argued that split-screen coverage of Bush’s reactions to Kerry’s statements hurt the president during the first debate. This study analyzes the impact of split-versus single-screen debate coverage in the 2004 campaign using a large-scale experimental design with about 700 participants, conducted at a large midwestern university. Consistent with explanatory models from political science and social psychology, findings show that split-screen coverage led Bush supporters to become more extreme in their positive and negative judgments about Bush and Kerry, respectively. Kerry supporters however, had strong views about Bush from the beginning and changed little based on the mode of coverage.

Keywords: negativity; presidential debates; split-screen coverage; turnout

Body language and nonverbal cues have always been an important part of presidential debates in the United States. Nixon’s visible perspiration during his debate with Kennedy on September 26, 1960, was widely interpreted as a sign of nervousness and is often cited as one of the key reasons for Kennedy’s narrow victory in the general election. A similarly prominent example of nonverbals were George H. W. Bush’s quick glances at his wristwatch during the second, town hall–type presidential debate in Richmond, Virginia, during the 1992 race. The most recent example before this past election cycle was Al Gore’s performance in Boston, Massachusetts, during the first presidential debate of the 2000 race on October 3, 2000, with George W.

Authors’ Note: All correspondence regarding this article should be addressed to Dietram A. Scheufele, Department of Life Sciences Communication, University of Wisconsin–Madison, 440 Henry Mall, Madison, WI 53706; e-mail: scheufele@wisc.edu.
Bush. Gore’s clearly audible and exacerbated sighs made him the target of much post-debate commentary and criticism, and he adjusted his display of nonverbals in the second debate. Nonetheless, he lost the electoral college vote in the general election.

In part as a result of these examples, the two major presidential candidates in 2004 agreed on a 32-page “Memorandum of Understanding” before the debates that outlined clear guidelines of behavior for each of the candidates and also for television networks (Jurkowitz, 2004). Specifically, the memorandum suggested that broadcast and cable networks were only allowed to show isolated shots of the candidates while they were speaking, and no reaction shots or split-screen shots. Many broadcast and cable networks, however, did not follow the memorandum, arguing that there was no evidence that “anyone [makes] a judgment based on the shape of the split screen” (Phil Griffin, vice president for prime-time programming at MSNBC, as cited in Anderson & Collins, 2004, p. 18).

In spite of the networks’ argument that the mode of coverage did not make a difference, the debate was followed by days of postdebate coverage, discussing the potential damaging impact of George W. Bush’s supposed expressions of anger, frustration, and arrogance (e.g., Milbank, 2004). In fact, the Democratic National Committee quickly compiled a 51-sec video titled “Faces of Frustration” that was posted on their Web site. Experts overall seemed to agree that the split-screen shots had hurt Bush (Milbank, 2004).

Of course, this is an empirical question. Does the mode of presentation influence how people evaluate candidates and their debate performance? And it is an empirical question that matters beyond the realm of presidential debates and has direct relevance for our understanding of how people form judgments based on media portrayals. In this article, we examine this question empirically using a large-scale experimental design. Specifically, we use theoretical models from communication, political science, and social psychology to explore how the nature of debate coverage has influenced viewers’ perceptions in the 2004 debate. Before going into the specifics of the current study further, however, it is necessary to locate this particular question in the larger context of current research on debate effects.

**Debate Effects—An Overview**

Previous research on debate effects has focused on two broad groups of outcome variables: learning about candidates’ issue positions and impression formation about candidates and its potential impact on vote choice.

*Political learning from debates.* First, a number of researchers have investigated the effects of debates on voters’ learning about the issue positions of candidates. The discussion of cognitive debate effects is mostly focused on the issue of whether voters who are exposed to debates also change their policy preference based on new information and thereby possibly their vote choice (Chaffee, 1978; Lanoue, 1992; Lemert, 1993; Miller & MacKuen, 1979; Zhu, Milavsky, & Biswas, 1994). Zhu
et al. (1994), for instance, found that televised debates improve content-specific issue knowledge. They concluded that debate viewing has an effect on issue knowledge only for issues proposed by the candidates themselves in the debate. Chaffee (1978) found that the effect of debate viewing on issue voting is greatest for regular viewers of debates who are likely more interested in politics than other groups.

Impression formation. Debate watching has also been found to significantly affect overall debate evaluation by priming certain types of candidates’ personality traits and character (Druckman, 2003; Lang & Lang, 1962; Mondak, 1995; Tannenbaum, Greenberg, & Silverman, 1962). Specifically, exposure to debates can influence perceptions of the candidates’ character and voter preferences. Druckman (2003), for example, conducted an experimental study using stimulus material from the first Kennedy-Nixon debate. He concluded that televised debates prime participants to rely more on personality perceptions when evaluating candidates, which, in turn, can affect overall evaluations of debate performance.

Communication modality as a key independent variable. Examining the impact of the nature of debate coverage more carefully, some researchers have also focused on the effect of communication modality (channel difference) on voters’ issue knowledge, character evaluation, and vote choice. The advent of television, in particular, had a strong influence on the study of debate effects (Perloff, 1998). Given the unique characteristics of televised debates, such as visual imaginary and nonverbal cues, several researchers have argued that televised debates may highlight a candidate’s personal qualities more than listening to the debates on the radio would, for instance (Hellweg, Pfau, & Brydon, 1992; Keeter, 1987). Examining political coverage more generally, other researchers have concluded that the choice of camera angle can influence viewers’ perception of a politician, particularly among his or her supporters (Kepplinger & Donsbach, 1987).

Furthermore, with rapid technological developments, the presentation of debates on television has also changed. One very prominent example in the 2004 presidential campaign was the argument between the Bush and Kerry camps over split-versus single-screen presentation of the debates. As outlined earlier, the contested question was if television stations were allowed to show reaction shots of one candidate while the other candidate was speaking. Especially the Bush campaign was concerned about potential detrimental effects of reaction shots on their candidate.

Unfortunately, there is little empirical research in the effects of the mode of presentation on debate outcomes. This is especially surprising given that “reaction shots are one of the most commonly used editing devices used to capture and manipulate non-verbal cues in film and television” (Davis, 1999, p. 477; see also Messaris, 1994). Visual persuasion research in advertising suggests, however, that juxtaposing two visual images to imply a contrast is equivalent to posing a complex puzzle for viewers and prompts cognitive processes. The solution to the puzzle will depend on the viewer’s predispositions toward politics and the media (Messaris, 1997). Following
this logic, split-screen coverage is likely to make people more aware of the nonverbal reactions of their candidate’s opponent, and—as a result—viewers may be given higher character and competence ratings to the candidate they support if his or her opponent shows strong nonverbal disagreement (Seiter, 1999; Seiter, Abraham, & Nakagama, 1998). In addition to nonverbal cues about character or expertise, split-screen coverage may also increase perceptions of incivility during debates. Mutz and her colleagues (Mutz & Holbrook, 2003; Mutz & Reeves, 2005), for example, examined the level of incivility during televised debates and its impact on evaluations of candidates and on the political process. They argued that televised incivility during televised political discourse (not necessarily presidential debates) influences respondents’ emotional responses and levels of trust toward the government and politicians.

Explicating Effects of Split- Versus Single-Screen Debate Coverage

In the current study, we examine voters’ postdebate candidate evaluations and the influences that predebate partisan predispositions and mode of debate coverage have on them. Specifically, we argue that predispositional factors (e.g., prior partisan preferences) play an important role in how people make sense of different modalities of debate coverage. In fact, past research suggests that predispositional factors, more generally (Benoit & Hansen, 2004; Benoit, McKinney, & Holbert, 2001), and partisan considerations (e.g., party affiliation), in particular (Berelson, Lazarsfeld, & McPhee, 1954; Campbell, Converse, Miller, & Stokes, 1960), are very important tools for citizens to interpret incoming information Sears and Chaffee (1979), for example, found that “information flowing from the debates appears to have promoted higher levels of consistency among voters’ party identifications, candidate evaluations, issue positions, and vote intentions” (p. 244). In the same vein, Sears, Freedman, and O’Connor (1964) found that one effect of debates is to reinforce partisan voters’ commitments to a candidate whom they initially support. Their explanation of this phenomenon is based on selective perception. Strong partisans are also more likely to support their preferred candidate’s presentation. Partisan leanings “boomerang” against the opposite candidate, however, whose performance is interpreted more negatively, based on preexisting partisan leanings.

In sum, we argue that the effect of televised presidential debates on candidate performance evaluation (emotional response) is a function of interaction of participant’s initial candidate preference (based on party identification) and communication modality. To examine these links empirically, however, it is necessary to outline two explanatory theoretical models in greater detail.

The First Model: An RAS Explanation for Split-Screen Effects

The first potential explanation for differential effects of split- versus single-screen debate coverage is based on the memory-based model of information processing. This
model assumes that people form attitudes based on the considerations that are most salient, that is, most easily accessible, when they make the decision (Hastie & Park, 1986). In other words, judgments and attitude formation are directly correlated with “the ease in which instances or associations could be brought to mind” (Tversky & Kahneman, 1973, p. 208; see also Hodges & Wilson, 1993; Houston & Fazio, 1989). Instead of using every relevant piece of information from their memory when forming a judgment on an issue, participants therefore only use information that happens to be “conveniently ‘located’ or accessible” (Iyengar, 1990, p. 2; see also Abelson, 1986).

The RAS model. Zaller’s (1992) receive-accept-sample (RAS) model develops the idea of memory-based attitude formation further. Similar to the basic notion of memory-based information processing, the RAS model assumes that people do not possess long-term attitudes on political issues but rather “base their attitude reports on the considerations that are immediately salient to them” (Zaller, 1992, p. 62).

The RAS model is based on four assumptions or axioms, as Zaller (1992) calls them. First, higher levels of political engagement\(^1\) make it more likely for a participant to be exposed to and be able to comprehend a political message he or she is exposed to. Political awareness, for Zaller, is operationally defined as a “general measure of political knowledge . . . , i.e., a summary score across a series of neutral, factual tests of public affairs knowledge” (p. 43). Second, resistance to messages or arguments depends on the perceived congruity between message and personal predispositions. Messages are rejected if they are incongruent with a respondent’s personal predispositions. Third, the more recently a consideration or argument has been brought to mind, the less time it takes a participant to retrieve that argument or related considerations from memory. Fourth, when making attitudinal judgments, participants tend to use considerations most easily available or accessible to them (Jervis, 1993).

Types of resistance. The first two axioms of the RAS model refer to being exposed to and resisting political messages. Specifically, Zaller’s RAS model specifies three types of resistance that—cumulatively—allow us to make specific predictions about the effects of split- versus single-screen debate coverage on partisans.

The first type of resistance, partisan resistance, assumes that individuals with strong partisan predispositions “may refuse to internalize new dominant messages that they recognize as inconsistent with their predispositions” (Zaller, 1992, p. 121).

Although partisan resistance is based on the assumption that people selectively expose themselves to persuasive messages and do not even receive messages (the \(R\) in RAS) when they are inconsistent with their personal predispositions, the models of countervalent and inertial resistance describe resistance processes during or after accepting a message.

Countervalent resistance occurs if respondents are more likely to accept (the \(A\) in RAS) attitude-consistent or countervailing messages. They will, of course, also receive
and internalize some inconsistent arguments or considerations. These inconsistent considerations, however, are rejected because their effects are counteracted by large numbers of incoming consistent messages.

The third type of resistance, inertial resistance, assumes that individuals possess certain amounts of preexisting considerations on political issues. Even if contradicting information or messages are received and accepted, their effects will be sampled out (the $S$ in RAS) or “swamped by the effects of previously formed considerations” (Zaller, 1992, p. 121).

**RAS and single-versus split-screen debate coverage.** Based on the idea of partisan resistance, partisans are less likely to expose themselves (or attend to) attitude inconsistent messages. It can be expected that partisans are therefore also less likely to be influenced by counterarguments made by the opposing candidate than nonpartisans. This difference between partisans and nonpartisans, however, can be expected to be largely consistent across models of presentation and is therefore less relevant for this current study. The other two types of resistance, however, have important implications for the potential effects of split-versus single-screen debate coverage.

Countervalent and inertial resistance are based on the idea that partisans evaluate each argument or consideration they receive based on how consistent their preexisting attitudes. For the candidate whom a person supports, this is relatively easy. In other words, all arguments this candidate makes are probably considered attitude consistent. It may be harder for people, however, to evaluate arguments of the candidate they initially disagreed with. This is in part a function of what Price and Tewksbury (1997) call “applicability effects.” Applicability effects assume that audiences process information based on preexisting cognitive schemas. Only if a message is applicable to a preexisting interpretive schema will it have an effect on audience members. Of course, many interpretive schemas are directly tied to partisanship. Abortion discourse is a good illustration. It is heavily dominated by “pro-choice” and “pro-life” frames that are the predominant heuristics used by partisans on the left and the right to justify their respective issue stances. Based on the applicability model, it is therefore easier for a Democratic viewer to correctly identify an abortion message framed in pro-choice terms as attitude consistent than it is for a Republican to correctly identify the same message as attitude inconsistent.

This is where the difference between split- and single-screen coverage is likely to show the strongest effect. Watching single-screen coverage of a debate is similar to being exposed to separate messages by the candidates with little relationship between the two. When watching split-screen coverage, however, partisan viewers see the candidate they oppose speaking while also watching their own candidate and his or her nonverbal reactions (frowns, sighs, etc.). These nonverbals by their own candidate make partisan viewers constantly aware of the fact that he or she disagrees with and disapproves of the statements made by his or her opponent. Making this disagreement salient through the split-screen coverage, therefore, also increases the
chances that partisans will label the opponent’s messages as inconsistent with their own attitude and reject them. As a result, we put forth the following hypothesis:

**Hypothesis 1**: Split-screen coverage will be linked to more negative evaluations by partisans of the candidate they oppose than single-screen coverage.

**Alternative Explanations for Split-Screen Effects: Using Cognitive Consistency Theories and Social Judgment Theory**

Differential effects of split- versus single-screen debate coverage can also be explained based on cognitive consistency models in social psychology (Festinger, 1957; Heider, 1958; Newcomb, 1953; Osgood & Tannenbaum, 1955, Tannenbaum, 1968). Congruency theory (Osgood & Tannenbaum, 1955), in particular, provides a useful framework for theorizing about the differential effects of split- versus single-screen coverage on viewers’ attitudes toward presidential candidates. Congruency theory takes into account three elements: an individual (p), dealing with an assertion about an object or a concept (x) about which p feels a certain way, and an assertion made by a source (s) toward whom p also has some preconceived attitude. According to cognitive congruency theory, when p’s attitude toward s and x are positive and s’s assertion is equally strong and of the same valence, p has no motivation to change his or her attitude toward the object or toward the source because the attitudinal triad has a congruous structure. However, when p’s attitude toward s is positive, and p has a positive attitude toward x that s later negatively evaluates, an incongruous structure is established that p will make it congruent by either changing his or her attitude toward s, or changing his or her attitude toward x. Unlike other consistency theories that take into account only the valence of attitudes, congruency theory also takes into account the strength of the attitudes. In other words, incongruency can be created not only when the valences of the attitudes are not congruent but also when the strengths of the attitudes in the triad are not congruent. The change in attitudes, however, will be less drastic in the case of weaker relationships (Osgood & Tannenbaum, 1955).

We argue that exposure to the split-screen condition can trigger an incongruent situation when the viewer has a neutral, or moderately negative, attitude toward the political opponent of his or her favorite candidate. If a viewer (p) has, for example, a preference for President Bush and a strong negative attitude against Kerry (Kerry being x in this case), seeing President Bush’s negative nonverbal responses will create a congruent triad in a viewer’s mind in the split-screen condition. However, if a viewer has a preference for President Bush and a weakly negative or neutral attitude toward John Kerry, split-screen coverage will expose the viewer to Bush’s nonverbal reactions toward Kerry while Kerry is speaking. This strongly negative message implicit in Bush’s negative nonverbal reactions, visible in the split-screen condition, does not correspond with the viewer’s own moderate views toward Kerry. Based on
congruency theory, one would therefore expect viewers to try to establish more congruency by making his attitude toward Kerry more negative.

Single-screen coverage, on the other hand, will not create such incongruencies. It leads the viewer to make assessments about each candidate in a set of dyadic relationships that never create incongruency in his or her mind. The split-screen condition, on the other hand, creates a need for incongruency reduction that fosters attitudinal change. As far as outcomes are concerned, cognitive congruency theory therefore seems to be consistent with the RAS model when it comes to predicting the effects of split-screen coverage on negative evaluations by partisans. However, it does offer an additional prediction when put into the context of a participant viewing strong nonverbal reactions to his or her preferred candidate by his or her opponent. In this case, cognitive incongruency might lead viewers to adjust their attitudes toward a more positive view of their own candidate.

These predictions are further confirmed when using social judgment theory as a theoretical framework. According to social judgment theory (Sherif & Sherif, 1967), individuals tend to categorize persuasive information as belonging to one of three groups: (a) the expressed position is judged to be within the range of positions the participant accepts (i.e., within his or her latitude of acceptance); (b) the expressed position is judged to be within a range of positions that the participant neither accepts nor rejects (i.e., within his or her latitude of non commitment); and (c) the expressed position is judged to be within the range of positions that the participant clearly rejects (i.e., within his or her latitude of rejection). According to social judgment theory, individuals also hold an “anchor” attitude. If the incoming persuasive information falls within the latitude of acceptance and is close to the anchor position, then the participant assimilates the new position by making it seem to be even more acceptable than it really is. However, if the incoming persuasive information falls outside of the participant’s latitude of acceptance, he or she will tend to make it seem worse than it really is, through a “contrast” effect. In the case of split-screen coverage of a debate, the visual cues presented by one’s own candidate are likely to fall within one’s latitude of acceptance and to be close to the participant’s anchor position. In this case, this will lead to assimilation effects, and the viewer’s initial negative attitude toward the other candidate will be reinforced. On the other hand, if the participant is exposed to strong nonverbal cues from the opponent, this might fall outside the participant’s latitude of acceptance and lead to contrast effects. In this case, the participant’s initial positive attitude toward his or her own candidate will be made even stronger.

Congruency theory and social judgment theory are consistent with the RAS model when it comes to predicting the effects of split-screen coverage on negative evaluations by partisans. We will therefore not restate Hypothesis 1 here and will add a second hypothesis:

**Hypothesis 2:** Split-screen coverage will be linked to more positive evaluations by partisans of the candidate they support than single-screen coverage.
Method

The data in the current study came from a Web-based experiment with participants who enrolled in communication courses at a large Midwestern university. Their instructors offered extra credit for participating in this research experience. All potential participants were contacted by e-mail and given the Web site of the online survey. The study was conducted during the 2 weeks prior to the 2004 presidential election. A total of 698 students participated in the experiment.2

Design

The study was based on a $3 \times 3$ experimental design embedded within an online survey. In addition to a battery of pre- and posttest questions, participants were exposed to two experimentally manipulated factors: a modified, 5-min segment of the first presidential debate (a split-screen, a single-screen, and an audio-only condition) and a 400-word fictional postdebate news analysis. Participants were randomly assigned to the manipulations

Manipulations

The debate clip manipulation took one of two modes: split-screen or single-screen screen.3 In the split-screen mode, participants saw the debate segment as it was aired on most television networks, that is, both candidates were on screen at all times, and the reactions and movements of one candidate were therefore visible while the other one spoke. In the single-screen condition, footage alternated between candidates as they spoke, such that participants were shown only footage of the candidate speaking. No reaction shots of the other candidate were presented.

The debate segments used in the experiments were selected from the first presidential debate, which took place September 30, 2004, in Coral Gables, Florida, with moderator Jim Lehrer, and were chosen for the topic covered and the way the candidates presented themselves. Pretest questions controlled for whether participants had seen any of the debates and how much attention they paid to them.4 In keeping with the first presidential debate’s theme of foreign policy, the candidates were asked by Lehrer to address the issue of nuclear weapons and diplomacy with North Korea and Iran. The candidates spoke at length about their views on how best to deal with the escalating nuclear threat presented by both countries, with Bush supporting multilateral talks and Kerry favoring bilateral talks.5

Measures

Criterion variables. After watching the televised presidential debate clip, participants answered questions focusing on the criterion variable: a feeling thermometer
measure. We asked participants how favorable they felt toward Kerry and Bush, respectively. Participants rated the candidates on a 100-point scale, ranging from unfavorably (e.g., ratings between 0 to 50) to favorably (e.g., ratings between 50 to 100; $M = 38.04$, $SD = 31.57$, for Bush; $M = 60.86$, $SD = 26.71$, for Kerry).

**Blocking variables.** Even in an experimental design with random assignment to conditions, controls help ensure that an experimental treatment is not confounded with other factors in accounting for criterion variables (Keppel, 1991; Kirk, 1995). As outlined earlier, our analyses did not include the effects of the postdebate analysis manipulation (e.g., policy vs. performance) that was part of our overall design. Previous research found that exposure to televised debate watching and postdebate analysis can influence respondents’ performance ratings of candidates (Lemert, Elliott, Bernstein, Rosenberg, & Nestvold, 1991). In contrast to the study of Lemert et al. (1991), however, our postdebate analysis manipulation was not designed to lean toward either Kerry or Bush but rather to portray a neutral position. Given that this second manipulation was not directly relevant to the theoretical relationships tested here and given its lack of one-sided partisan content, its effects on our outcome variables should be minimal. To rule even small potentially confounding influences, however, we treated this manipulation (postdebate analysis, policy vs. performance) as covariates or “blocking variables” in our analyses. It is important to note, however, that additional analyses not reported here showed no significant effects of postdebate analysis on our criterion variables, such as thermometer ratings or candidate evaluations.

Given the design of our experiment, we were also faced with prior debate viewing as a potential confounding factor. Specifically, it is possible that the effect of our manipulation is attenuated by prior debate viewing. We therefore controlled prior debate viewing as a blocking variable in all tests related to our hypotheses.

**Predictor variables.** In the current study, two main predictor variables were used: the experimental factor (e.g., single- vs. split-screen) and prior candidate preference. Prior candidate preference was measured by asking respondents in the pretest which major party candidate they would vote for if the presidential election were held today. Respondents were categorized into a low and high group by splitting the sample at the median.6

**Results**

To test hypothesized relationship, the effects of different modality and predispositions on candidate evaluations were analyzed by employing two-way ANCOVAs, where two main effects of modality (e.g., single-screen [single shot] vs. split-screen
[reaction shot]) and prior candidate preference, and the interaction effect between them were assessed. All other experimental factors were controlled as covariates in the analysis, functioning as blocking variables.

**The Effects of Modality (Split vs. Single) and Prior Candidate Preference on Candidate Evaluations**

Supporting our hypothesis, the effect of split-screen versus single-screen debate coverage interacted with prior candidate preference in their effects on debate evaluation. The effect, however, was only significant for evaluations of Kerry and not of Bush. As Table 1 shows, split-screen coverage leads to significantly more negative evaluations of Kerry than single-screen coverage. The effect, however, was dependent on prior candidate preference, that is, it was significantly stronger for Bush supporters than for Kerry supporters, $F(1, 432) = 12.439, p < .001$. As each of our two independent variables has only two categories, additional post hoc analyses are not necessary to establish that there are significant differences between these values. We did not find the same patterns for evaluations of Bush, however (see Table 2). In fact, there was no significant interaction between prior candidate preference and modality of coverage for Bush, $F(1, 432) = .441$, ns.

Figure 1 illustrates these relationships in a more intuitive fashion. As this figure shows, the split-screen condition polarized evaluations of Kerry compared to the single-screen condition, depending on prior candidate preference. Of course, prior candidate preference showed a significant main effect on evaluations of debate performance in both models. The modality of coverage (split- vs. single-screen), however, did not show main effects in either model.

**Potential Priming Effects on Perceptions of Debate Performance**

Although Zaller’s (1992) RAS model and various types of cognitive consistency models provide a consistent predictive framework for our hypothesized relationships, there is one possible alternative explanation for the phenomena we outline in our hypotheses. Based on the memory-based model outlined earlier (e.g., Hastie & Park, 1986), it is possible that split-screen debate coverage produces a stronger effect on candidate evaluations than single-screen coverage because it primes character traits in addition to exposing viewers to different debate messages (Mondak, 1995). In other words, seeing two candidates on the screen at the same time may focus viewers’ attention on the characteristics of the candidate rather than the content of their exchange. As a result, these character evaluations may be more salient and influence subsequent evaluations of the candidate (Iyengar & Kinder, 1987).

If this assumption is correct, of course, we should observe an effect of split-screen versus single-screen coverage for both candidates. As we outlined earlier, we did not.
Nonetheless we controlled for potential character-priming effects by examining our hypothesized relationships for three types of character evaluations as outcome variables: expertise, civility, and nervousness. In our experiment, we tapped these three dimensions using a battery of questions about the participants’ evaluation of each candidate’s debate performance. Specifically, participants were asked how they felt about the debate performance of each candidate in this particular video clip. For expertise, participants were asked to assess the debate performance of each candidate using three semantic differential-type scales (competent/incompetent, credible/not believable, and informed/uninformed; Cronbach’s $\alpha = .92$, $M = 16.04$, $SD = 8.13$ for Bush; Cronbach’s $\alpha = .84$, $M = 22.11$, $SD = 5.29$ for Kerry). For civility, we used the following three items: hostile/friendly, aggressive/passive, and agitated/calm (Cronbach’s $\alpha = .62$, $M = 19.28$, $SD = 5.21$ for Bush; Cronbach’s $\alpha = .63$, $M = 15.25$, $SD = 5.10$ for Kerry). Finally, perceptions of nervousness were measured

### Table 1

<table>
<thead>
<tr>
<th>Variance Source</th>
<th>df</th>
<th>M</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance (Dummy)</td>
<td>1</td>
<td>810.522</td>
<td>2.784</td>
<td>.096</td>
</tr>
<tr>
<td>Policy (Dummy)</td>
<td>1</td>
<td>249.917</td>
<td>0.858</td>
<td>.355</td>
</tr>
<tr>
<td>Modality Manipulation</td>
<td>1</td>
<td>487.060</td>
<td>1.673</td>
<td>.197</td>
</tr>
<tr>
<td>Prior Candidate Preference</td>
<td>1</td>
<td>184,721.715</td>
<td>634.445</td>
<td>.000</td>
</tr>
<tr>
<td>Modality $\times$ Prior Candidate Preference</td>
<td>1</td>
<td>3,621.790</td>
<td>12.439</td>
<td>.000</td>
</tr>
<tr>
<td>Within-cell errors</td>
<td>432</td>
<td>291.155</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Variance Source</th>
<th>df</th>
<th>M</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance (Dummy)</td>
<td>1</td>
<td>63.264</td>
<td>0.176</td>
<td>.675</td>
</tr>
<tr>
<td>Policy (Dummy)</td>
<td>1</td>
<td>396.453</td>
<td>1.102</td>
<td>.294</td>
</tr>
<tr>
<td>Modality Manipulation</td>
<td>1</td>
<td>0.395</td>
<td>0.001</td>
<td>.974</td>
</tr>
<tr>
<td>Prior Candidate Preference</td>
<td>1</td>
<td>286,010.861</td>
<td>795.140</td>
<td>.000</td>
</tr>
<tr>
<td>Modality $\times$ Prior Candidate Preference</td>
<td>1</td>
<td>158.558</td>
<td>0.441</td>
<td>.507</td>
</tr>
<tr>
<td>Within-cell errors</td>
<td>432</td>
<td>359.699</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
using respectful/disrespectful, composed/frustrated, and confident/nervous (Cronbach’s $\alpha = .82$, $M = 14.75$, $SD = 7.09$ for Bush; Cronbach’s $\alpha = .76$, $M = 7.35$, $SD = 4.70$ for Kerry).

We ran three two-way ANCOVAs for each candidate’s traits to examine our initial finding for the three dimensions of candidate evaluation. As Table 3 shows, however, we only found significant results for perceptions of expertise for Kerry and not for other five ANCOVAs. Specifically, the one significant finding shows effects of split-screen debate exposure on perceptions of Kerry’s expertise for people who support Bush, $F(1, 433) = 5.636, p < .05$.

**Potential Moderating Effects of Prior Debate Viewing and Political Knowledge**

As outlined earlier, Zaller’s (1992) RAS model is based on the assumption that citizens reject political messages based on the consistency between message origin and slant and their own partisan leanings. Zaller also assumed, however, that only if

![Figure 1](http://example.com/figure1.png)

**Figure 1**
**Effects of Modality and Prior-Candidate Preference on Kerry’s Feeling Thermometer**

<table>
<thead>
<tr>
<th>Candidate Preference</th>
<th>Thermometer Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-Screen</td>
<td>80</td>
</tr>
<tr>
<td>Split-Screen</td>
<td>70</td>
</tr>
</tbody>
</table>

Downloaded from crx.sagepub.com at Ebsco Electronic Journals Service (EJS) on January 19, 2011
a participant is politically aware enough to in fact perceive the incongruity between
message and own predispositions will he or she reject the message.

At the same time, our manipulations were based on footage from the original
campaign. In other words, the effect of the manipulation may be minimal for partic-
ipants who watched most of the original debates on television already. We therefore
factored in two potential covariates: prior factual political knowledge and prior
debate viewing.

Political knowledge measure was an additive index constructed from eight questions
concerning the issues of political party and presidential candidates. Correct answers
were scored one, all other answers zero (Cronbach’s $\alpha = .88, M = 6.19, SD = 2.64$).7

Prior debate viewing was measured by asking participants how many of the three
presidential debates they had watched ($M = 2.00, SD = .98$).

To account for potential moderating effects, we tested additional three-way inter-
actions for each of these variables using the feeling thermometers as outcome vari-
ables. Neither model showed significant support for a potential moderating role
of prior knowledge or prior debate viewing.8 For example, prior knowledge did
not interact with the experimental condition in predicting feelings toward Bush,
$F(1, 432) = 1.660, ns$; and Kerry, $F(1, 432) = .054, ns$. In addition, introducing prior
debate viewing as an additional factor in the analyses did not substantially alter our
findings, $F(1, 432) = 1.116, ns$, for feelings toward Bush; and $F(1, 432) = .461, ns$,
for Kerry. In sum, our data suggest that political knowledge and prior debate view-
ing did not significantly alter the effects of different modalities (single- vs. split-
screen) and predisposition (prior candidate preference) on candidate evaluations.

Discussion

The current study was concerned with the effects of split- versus single-screen
coverage on evaluations of presidential candidates. As outlined earlier, this area of
debate research is surprisingly understudied, especially given the controversy surrounding the 32-page “Memorandum of Understanding” issued by the Bush and Kerry campaigns before the 2004 debates, limiting broadcasts to single-screen shots of each candidate (e.g., Anderson & Collins, 2004). As outlined earlier, networks largely ignored the memorandum (Jurkowitz, 2004). There seemed to be a consensus among pundits and journalists, however, on the detrimental effects that Bush’s body language and facial expressions had on public perceptions of his performance (e.g., Anderson & Collins, 2004).

Our data suggest that these largely intuitive judgments were far too simplistic and in part even incorrect. Specifically, our analyses showed that the split-screen coverage had significant effects mostly for Bush supporters who developed a stronger dislike for Kerry as a result of split-screen coverage. Attitudes of Kerry supporters remained relatively stable, regardless of the mode of coverage. Before discussing these findings further, it is necessary, however, to examine some of the specific characteristics of our data and the implications for the conclusions that can be drawn from our results.

Our analyses are based on a large-scale experimental study of debate effects. When designing the current study, we faced a number of methodological decisions that are worth discussing further. Most important, we decided to use a student sample with almost 700 student respondents. College students, of course, tend to be different from the general population with respect to their political preferences, their likelihood of voting, and other attitudinal and behavioral variables. For our current study, however, there is one particular issue that needs to be addressed.

This issue relates to the problem of within-sample variation. For some of the variables in our analyses, for instance, college students differ significantly from the general population in terms of variable means and variation. This includes evaluations of candidates, partisanship, and other variables. This is especially a concern for variables where the current study sample shows less variation than the general population. In these cases, some of the relationships we report here may in fact be attenuated, compared to what could be observed in the general population. For two reasons, however, this is not a major concern in the current study. First, the distribution of key variables in the current study was similar to the distributions observed in the 2004 American National Election Study (ANES), suggesting that constrained variances or artificial floor and ceiling effects because of our sample probably did not play a role in confounding our findings. Second, the design used in the current study may in fact lead to an overestimation of relationships. As outlined earlier, we exposed our experimental groups to a 5-min segment from the first presidential debate as part of the overall study design. This, of course, could be interpreted as a classic case of what Hovland (1959) would call a “captive audience.” In other words, we asked college students to pay close attention to a political debate that many of them would otherwise never be exposed to, which may artificially increase the size of the effects we observed.
These concerns, however, were outweighed by the obvious advantages of the design we employed. First, the experimental nature of our study allowed us to make strong causal claims that would have been impossible to make using cross-sectional or even most longitudinal survey designs. Second, our stimulus was based on footage from the original debate broadcast. To rule out effects due to perceptions of source credibility or source bias, we created the two versions of the debate clip (single and split screen) from the raw footage, that is, without station logo or other on-screen elements. As a result, we ended up with a very externally valid stimulus. The only possible drawback was the fact that people who had seen the first debate may potentially not show any effects based on our short 5-min stimulus from the same debate. This concern was obviously not warranted, given that we found significant effects in line with our original hypothesis.

In addition, we also need to address a potential concern related to the use of single versus multiple message designs in experimental studies. According to Reeves and Geiger (1994), messages are complex structures; and when theorizing about a particular attribute of messages, “it is quite difficult (if not impossible) to isolate this single feature to use as a stimulus” (p. 166). It is therefore necessary, Reeves and Geiger argued, to apply “message sampling” to create external valid experimental designs. In other words, the conclusions drawn from a single-message design are too unique to be able to generalize to other, similar messages, thus producing inconsistencies between studies.

Message sampling addresses this problem in two ways. First, a large sample of messages can minimize errors from confounding message characteristics. Second, a large sample of messages helps stabilize the responses in different treatment conditions. This reasoning has most recently been summarized by Leshner (2006), “The problem with single-message design is,” he argues, “that any conclusion that can be made about the effect of the manipulation must be constrained to that particular message” (p. 2).

Despite these very valid critiques of single-message designs, Bradac (1983) argued there are instances where a single-message design can be preferable to multiple messages. For instance, when “a particular message is important in and of itself, there is little need for message multiplicity” (p. 184). One example of this would be the study of the idiosyncratic effects of particular messages, such as a special newscast or political debate.

Bradac (1983) also argued that “the longer the message, the greater the number of irrelevancies and (potentially) the greater the number of representatives of the theoretically interesting construct distributed uniformly throughout the message” (p. 185). As outlined earlier, participants were exposed to a lengthy, 5-min debate clip. If we had manipulated negativity of messages, for instance, and created a stimulus using ten 30-sec campaign ads, our stimulus materials would have been of similar length but would have been considered a multi message design. Or to put it in Bradac’s words: “The long message can be viewed as many within-study replications comprising shorter messages” (p. 185).
It is important to note, of course, that we are not arguing against multimessage designs. In face, we believe that Reeves and Geiger (1994) and Leshner (2006) have identified a critical threat to the external validity of experimental findings. Following Bradac’s logic, however, we relied on a single, lengthy message in the current design. This also means, of course, that our results should be generalized with caution particularly in the context of presidential debates.10

With these considerations in mind, the current study offers a systematic test of the effects of split- versus single-screen debate coverage during the 2004 presidential campaign. Specifically, the findings support our predictions for evaluations of Kerry, but not for Bush. As outlined earlier, these predictions were based largely on Zaller’s (1992) RAS model, congruency theory, and social judgment theory. Two questions need to be answered in this context. First, why did the effect hold for Kerry but not for Bush? This is especially surprising, given the widespread assumption after the debate that his nonverbals had hurt Bush (Anderson & Collins, 2004). Second, was this effect a simple priming effect where the split-screen coverage made personality traits more salient, or did the effect really follow our theoretical rationale?

The answer to the first question lies in the level of dislike that Kerry supporters had for Bush in the first place. The average thermometer rating for Kerry among his supporters increased from 73.03 in the single-screen condition to 77.13 in the split-screen condition. In other words, seeing Bush’s reaction to Kerry’s answers and vice versa made Kerry supporters feel more favorable toward their candidate. The single-split distinction, however, made little difference with respect to their evaluations of Bush. In the single-screen condition, Kerry supporters showed a floor effect and rated Bush almost as low (22.26) as in the split-screen condition (20.97). Bush supporters, on the other hand, were steadfast in the support for their candidate. Being exposed to the split-screen condition increased their thermometer ratings of Bush to 80.19 from 78.77 in the single-screen condition. This combination of floor effects (for Kerry supporters’ rating of Bush) and ceiling effects (for Bush supporters’ ratings of Bush) most likely prevented a similar interaction between prior candidate preference and model of presentation for Bush as we observed for Kerry.

The second question relates to the appropriateness of our theoretical model. As outlined earlier, it is possible that the mode effects are simply a function of the split-screen manipulation making personality traits of the candidates more salient and therefore influencing thermometer ratings. We therefore measured perceptions of personal characteristics for each candidate during the debates. The mode effects were not significant for almost all of the personality characteristics we measured, suggesting that our hypothesized effects are not a function of making personality traits more salient in the split-screen condition. Rather, our results are consistent with outcomes that could be predicted on RAS models, congruency theory, and social judgment theory. In other words, based on the RAS model, it is reasonable to assume that people use their initial candidate preference as an interpretive schema for incoming information. In other words, they will perceive the opposing candidate...
as even more opposed to their own viewpoints if they see him opposing their favored candidate in a split-screen setup. As a result, they will also rank him or lower on the thermometer scales.

Congruency theory and social judgment theory theoretical models offer a similar explanation. In addition, however, they allow us to explain some of the specific results even better than the RAS model. Split-screen coverage creates a triad in people’s minds where they have to evaluate their relationships with the candidate they initially support and with his or her opponent. Seeing their own candidate’s reaction while his or her opponent is speaking in debates makes them aware of the strong negative link between the two and forces them to reassess a potential neutral or weak negative affect toward the opponent. This is very consistent with our findings for the thermometer ratings toward Kerry and Bush. Given their initial assessments of both candidates, seeing the split-screen coverage did not force Kerry supporters to readjust their attitude toward either candidate. As outlined earlier, they were already very negative on average in their assessments of Bush. Consistent with congruency theory, however, split-screen coverage led to an ever stronger positive assessment of Kerry by his supporters. This slight upward adjustment in their judgment of Kerry is not explained by the RAS model that is only concerned with the resistance toward attitude-inconsistent messages. It makes perfect sense in the context of congruency theory, however. Also consistent with congruency theory, split-screen coverage made Bush supporters reinforce the links within their perceptual triad toward an even more positive evaluation of Bush and even more negative evaluation of Kerry. In fact, Bush supporters ended up with a similar level of assessment toward their own candidate and his opponent as Kerry supporters had in the first place.

In sum, our findings show how the mode of presentation interacts with preexisting candidate preferences to influence people’s evaluations of candidates. The Los Angeles Times headlined after the first 2004 debate: “Split screen wasn’t kind to Bush” (Anderson & Collins, 2004). Our findings suggest that the predominant mode of coverage for this debate was much kinder to Bush than it was to Kerry.

Notes

1. Zaller’s (1990, 1992) concept of political awareness is controversial to say the least and has been criticized from conceptual and operational perspectives. He conceptualized awareness as a latent construct, represented by a number of variables like interest, knowledge, and behavioral participation (Zaller, 1990). Criticism, especially based on empirical grounds, was raised by Krosnick and Milburn (1990) and Krosnick (1990).

2. Approximately 74% of students were female and 26% were male. The audio-only manipulation was not relevant for our theoretical model and therefore excluded from the analyses. Our total sample size therefore was 234 in the single-screen condition and 235 in the split-screen condition.

3. As outlined earlier, the audio-only condition was created to test a different set of research questions and was therefore not included in these analyses. All Ns reported refer to the two conditions using television clips. Participants were also exposed to written postdebate news analyses. The stories were focused
on one of two themes: policy or performance. A control group was presented with no postdebate analysis story. The second manipulation was also created to test a different set of research questions and was therefore not included in the analyses presented here.

4. We selected this specific clip after careful screening of the complete debate. We chose this clip for two reasons. First, it was a segment of the debate that was devoted to a single policy issue without either of the candidates getting off on tangents. Second, we wanted to make sure to capture a segment of the debate that was balanced in terms of the time each candidate spent on an issue and also in terms of the nonverbal reaction of each candidate to the other candidate’s answer.

5. In the second experimental manipulation, participants were presented with one of two postdebate news analyses. A control group was presented with no postdebate analysis story. The stories were focused on one of two themes: policy or performance. This second manipulation was also not part of the current study and was therefore excluded from the results presented here. Even though the random assignment to conditions should eliminate any potential contamination due to the second manipulation, all analyses presented here also include the second manipulation as a covariate to rule out this possibility. We coded the debate clip with two independent coders. In particular, the coders examined the debate clip including reaction shots in 10-sec segments. For each segment, they coded nonverbal reactions as positive (nods, affirmative smiles, etc., coded as 1), negative (shaking head, raising eyebrows, etc., coded as −1), or neutral (no or neutral nonverbal reactions, coded as 0). The average Cohen’s kappa was .901. The clip included Bush and Kerry’s responses, timed by the moderator, and a follow-up question by the moderator with both candidates providing short answers. There were no significant differences between Bush and Kerry in the number of nonverbals or in the average valence of nonverbals per segment. Specifically, there were 17 nonverbals coded for Kerry and 20 for Bush, which was largely a function of Kerry giving a slightly longer initial response to the question. Kerry’s mean nonverbal reaction per segment was −.24, and Bush’s was −.25.

6. This means, our analyses are based on a $2 \times 2$ cell design (single-screen vs. split screen, and prior candidate preference Bush vs. prior candidate preference Kerry).

7. In the process of constructing political knowledge we noticed that some respondents spent excessive amounts of time answering the knowledge questions. Because time spent also correlated highly with the number of correct responses, we assumed that some participants had used a Web browser to search for the correct answer while filling out the survey. We therefore examined the log-file data for each participant to see how much time he or she had spent on the knowledge battery. After plotting the distribution for all respondents, we treated all respondents who spent more than 200 secs responding to the knowledge battery as outliers.

8. It could be argued, of course, that our control for the number of debates that a respondent has seen may not be the most important control here. In fact, having watched the first debate may be the one factor that could confound the effectiveness of our stimulus. We therefore ran the three-way interactions for prior debate viewing in general and prior viewing of the first debate only. The results were virtually identical.

9. To check for potential sampling biases in our student sample, we compared our data with the 2004 American National Election Study (ANES). Our results indicate that the distribution of our dependent variable (i.e., the feeling thermometer), broken down by respondents’ willingness to vote for a particular candidate (which is comparable to our measure of prior candidate preference), does not seem to be skewed or constrained in terms of its variance. For instance, in our data Bush’s average feeling thermometer score among his supporters is 78.72 ($SD = 16.24$). For Kerry supporters, Bush’s feeling thermometer average is 20.78 ($SD = 19.56$). The average feeling thermometer score for Kerry among Bush supporters is 28.66 ($SD = 19.96$). Among Kerry supporters, Kerry’s feeling thermometer average is 74.65 ($SD = 16.24$). These numbers are similar to the distributions we found in the 2004 ANES. Bush’s average thermometer scores among his supporters in the ANES dataset are 83.60 ($SD = 15.87$). Among Kerry supporters Bush’s average feeling thermometer ratings are 27.81 ($SD = 23.74$). Kerry’s feeling thermometer score among Bush supporters had a mean of 33.16 ($SD = 21.69$). Among Kerry supporters, Kerry’s average feeling thermometer score is 71.81 ($SD = 15.95$).
10. This was also based on a number of additional considerations: (a) The first reason relates to the fact that we could not use materials from different debates. The heavy coverage of nonverbal reactions after the first debate shaped the expressive behaviors of both candidates in subsequent debates. Similarly, the “surprise element” of some networks choosing to broadcast the debates in split-screen versus single-screen format was gone and the candidates knew that their reactions were potentially visible to the audience. This rendered material from the second and third debate largely useless for our purposes. (b) Only one clip was used because we designed the current study to be an online experiment. This enabled us to collect data from a large number of participants while simultaneously minimizing history effects during the ongoing campaign. Using multiple lengthy clips would have attenuated participation rates in our experiments and dramatically reduced attention levels among participants. This, in turn, would have undermined the external validity of the current study, which is what a multimessage design is designed to protect.

References


**Dietram A. Scheufele** (PhD, University of Wisconsin) is professor and director of Graduate Studies in the Department of Life Sciences Communication and professor in the School of Journalism & Mass Communication at the University of Wisconsin. His current research focuses on the impact that new and traditional media have on public opinion and political behavior. He is particularly interested in the intersection of politics and scientific controversies.

**Eunkyung Kim** (MA, University of Wisconsin) is a doctoral student in the School of Journalism & Mass Communication at the University of Wisconsin. Her research interests include the impact of interpersonal discussion and media use on information processing and political behavior, the processes underlying opinion formation about science and politics, and quantitative research methods in media and market research.

**Dominique Brossard** (PhD, Cornell University) is an assistant professor in the School of Journalism & Mass Communication at the University of Wisconsin. Her current research focuses on the dynamics underlying public opinion processes related to controversial scientific issues.