Revisiting the Role of Information Format in Candidate Evaluation: An 'Update' Model of Evaluation*

Scott H. Huffmon
Winthrop University

Over a decade ago, Lodge, Stroh, and Wahlke (1990) wrote that all major models of vote choice and candidate evaluation were "black-box" models because they were "silent about the processes that drive their explanations" (p. 13). In an attempt to peer inside that black-box, Rahn, Aldrich, and Borgida (1994) highlight the importance on candidate evaluation of sophistication and the format in which information is received; they conclude that individuals use a memory-based, as opposed to an impression driven, model of evaluation. This article offers a re-examination of Rahn, Aldrich, and Borgida's model. I conclude that their use of fictitious candidates and the construction of some of their variables may have clouded some of their findings, but several of their theoretical assumptions are sound. I also posit an "update" model of evaluation that demonstrates individuals tend to use a mixed strategy of both on-line and memory-based approaches in creating evaluations and highlights the importance of the individual's policy proximity assessment, sophistication, and information format on the evaluative process as individuals update their evaluations in light of new information.

* This research is funded by the National Science Foundation (SES-9809223). I would like to thank the Social Science Research Laboratory at the University of Mississippi for the use of their facilities and John M. Bruce, Robert D. Brown, and Charles E. Smith, Jr. for their helpful comments.

THE JOURNAL OF POLITICAL SCIENCE
VOLUME 31 2003 PAGES 35-69

Over a decade ago, Lodge, Stroh, and Wahlke (1990) wrote that all major models of vote choice and candidate evaluation were "black-box" models because they were "silent about the processes that drive their explanations" (p. 13). While many scholars have tried to illuminate the dark recesses of that "black-box," we must often settle for more prediction than explanation from our models.

Examples of research that probe the black-box can be found in Lodge, McGraw, and Stroh (1989) and Rahn, Aldrich, and Borgida (1994). The former posits an impression-driven model of candidate evaluation, while the latter tests the conclusions of Lodge and his colleagues against a memory-based model, concluding that individuals employ a memory-based approach to candidate evaluation. This research offers a reexamination of Rahn, Aldrich, and Borgida's (1994) model and builds on the foundation they laid.

The crux of the difference between impression-driven and memory-based models lies in the role of information in the creation of an evaluation. Impression-driven models assume that evaluations consist of a running tally that is updated with the application of new information, the details of which are discarded or buried deep within the memory once the evaluation has been updated. Therefore, when asked to produce an evaluation, voters need only consult their tally without retrieving the actual information that served as its basis. Memory-based models, on the other hand, assume no preexisting evaluation, or tally, residing within the memory. If there is no previous evaluation to rely upon, voters must recall detailed information from memory in order to form an evaluation. The question of how individuals process political information is further complicated, however, when we introduce the concepts of motivation, sophistication, and information context.

Srull and Wyer (1986) demonstrated that when individuals pursue the specific purpose of forming an evaluation, an impression-driven, or on-line, approach is taken. One might assume that the campaign season would orient the voter to the goal of forming an evaluation leading up to election day. However, depending on the nature of the election, character of the candidates, and complexity of the issues, the impact of motivation may be eclipsed or mitigated by political sophistication. Bargh and Thein (1985) and McGraw, Lodge, and Stroh (1990) found sophistication to have significant impact on the use of an on-line strategy. Lodge, Steenbergen, and Brau (1995) highlight the impact of time as memory fades (the "forgetting curve") and the importance of campaign information and messages. Further, Rahn, Aldrich, and Borigda (1994) found that the format in which information was received had a significant interactive impact on the individual's ability to create candidate evaluations; according to their findings, information received in a debate-style format contributed to the use of a memory-based approach to candidate evaluation.

Unfortunately, the researcher is confronted with a difficult dilemma. In order to test the impact new information on evaluation, the researcher must be certain that the evaluation, once given, is really the product of the stimulus information and not merely a reflection of a previous evaluation, one held before the research began. The traditional way to overcome this problem has been to ask the research subject to evaluate fictional candidates. Such was the strategy in the evaluation research done by Lodge, McGraw, and Stroh (1989) and Rahn, Aldrich, and Borgida (1994). As Rahn, Aldrich, and Borgida (1994) note, "[i]n a real campaign, it would be virtually impossible to isolate voters before they had any exposure to the candidates" (p. 195).

Using fictitious candidates allows researchers the cleanest type of experimental design. Since subjects could not possibly

have an opinion prior to the experiment, the evaluation *must* be the result of information presented in the course of the study. However, the use of ficticous candidates has an unfortunate side effect: conclusions drawn from results based on the use of completely unknown candidates are not applicable to most real world electoral scenarios. If, in real campaigns, individuals never knew anything about candidates until they were given specific campaign information, experimental design using fictitious candidates would be directly applicable to the "real world." Rarely, however, are voters introduced to new candidates for the first time in elections. Although it may be the case more often in local politics that one or more prospective candidates are relative or complete unknowns, such occurrences are infrequent in elections that extend beyond the county line. Given the situation of two completely unknown candidates competing for the same office, a model that explains how individuals process information and how they integrate it into opinions being simultaneously formed for the first time is appropriate, but for most elections of note, it is simply not the case. It is difficult to demonstrate either impression-driven or memory-based candidate evaluation strategies when there are no long held opinions to be updated using one of these models.

My analysis replicates Rahn, Aldrich, and Borgida's (1994) test of impression-driven versus memory-based evaluation and the impact of information context on evaluation, but I use actual candidates to create a more realistic evaluative environment. Additionally, naturally created variables are used that more accurately reflect respondents' thinking and attitudes, including a latent measure of attitudinal strength that thwarts second guessing and artificial attitude justification in respondents' evaluative reports. Finally, an "update" model of candidate evaluation is introduced. The model controls for the impact of prior opinion

while still measuring the impact sophistication, memory, and information structure.

Attitudinal Strength

Models of comparative candidate evaluation hinge on the accurate measure of the strength of positive or negative opinions that individuals hold concerning the candidates. Although easy enough to ask respondents whether they approve or disapprove of a candidate, it is much more difficult to measure the strength of that opinion in an accurate way without forcing respondents to conduct an unnatural over-analysis of their opinions. The dilemma leads researchers unintentionally to change what they want to measure because of the obtrusiveness of the measurement itself. Traditionally, the best way to measure attitudinal strength with minimal intrusion has been to ask respondents to rate their opinions on a scale—a feeling thermometer or a straight-forward like-dislike scale, for example. Although these methods keep the artificial noise of over-analysis to a minimum, they still create an evaluative environment that is not likely to be reproduced in a real electoral setting.

Attitude strength is of paramount importance in the understanding of how individuals process information. One is more likely to use strong attitudes when forming evaluations than weak ones (Fazio 1993; Krosnick and Petty 1995). What is needed is a way to measure latent attitudinal strength without forcing respondents to construct an opinion report artificially that is more subject to measurement error. Studies in the field of psychology have developed such a measure.

Researchers in cognitive psychology, led by Russell Fazio and others, have used measures of response latency, or reaction time, to gauge the accessibility and activation of attitudes (Fazio et al. 1982; Powell and Fazio 1984; Fazio et al. 1986; Houston and Fazio 1989; Bargh et al. 1992; Doll and Ajzen 1992; Down-

ing, Judd, and Brauer 1992; Roskos-Ewoldsen and Fazio 1992; and others). Additional research has shown that important attitudes are more accessible than those attitudes of less importance (Szalay and Deece 1978). In a meta-analysis of nine previous experiments, Fazio (1993) confirmed that stronger attitudes lead to greater accessibility, as measured by reaction time. Similar conclusions were reached by Bargh *et al*. (1992) who found that subjects demonstrated faster reaction times for their strongest attitudes and slower reaction times for their weakest. Fazio and Williams (1986) found that evaluations of Republican candidates in presidential and vice presidential debates were more favorable among Reagan supporters and that this relationship was stronger among individuals whose attitudes about Reagan were more accessible. In short, extensive research has confirmed that reaction times, or response latencies, serve as accurate measures of how strongly an individual holds an attitude. Drawing on this research, I use response latencies of expressions of approval or disapproval of candidates as a measure of the strength of respondents' evaluative attitudes. The response latency measure has additional advantages: it prevents experiment participants from being "tipped off" as to the real nature of the attitudinal inquiry.

Research Design

Computer surveys were administered in the Social Science Research Laboratory at the University of Mississippi. Subjects were recruited from introductory political science courses at the university; 107 students participated in the experiment. Respondents ranged in age from 18 to 40 years old. Use of the standard National Election Study (NES) questions revealed that the majority, 82.17%, considered themselves either "strong" or "not so strong" partisans while 17.83% considered themselves to be political "independents." More than half of the independents were

partisan leaners, leaving only 7.92% who considered them selves to be pure independents.

Although some may question the validity of studies that use college students as subjects, David Sears (1986) notes that the consensus in the field of psychology is that "reliance on college student subjects does not have major negative consequences" for experiments. Sears (1986) did, highlight, however, several possible pitfalls of the use of college students. In fact, the major findings of Sears' research were that college students are likely to have less crystallized attitudes and stronger cognitive skills. Both of these conditions would work systematically against the models in this research, *suppressing* the expected relationships. One might infer that any significant findings would be magnified in the public.

Upon arrival at the computer lab, subjects were told that they would be participating in a political opinion survey. Subjects were randomly assigned to computer stations. The station assignment would determine which type of stimulus, the information context, the subject would see (see Appendix A). For the stimulus, subjects were exposed to information about Al Gore and George W. Bush in one of two formats. In the candidate-centered format, information was presented separately for each candidate. Subjects saw all of one candidate's positions on the issues followed by the positions of the other candidate on those issues. In the dimension-centered format (also commonly referred to as "debate-style" format) information was presented separately for each issue. Presented with each issue separately, subjects saw one candidate's position on that issue followed by the other candidate's position on that issue before being presented with the next issue.

Subjects first answered a series of twelve questions on whether they approved or disapproved of several items. The first three questions were non-political and designed to familiarize the

subject with the questionnaire format. The remaining questions probed approval or disapproval of George W. Bush, Al Gore, politicians in general, the two major political parties, parties in general, and the institutions of government. Unbeknownst to participants, their reaction time to each question was recorded as they responded. The next section of the survey asked subjects to identify the party affiliation of several notable political actors and to match policy positions to the correct political party as well as express their own policy preference. Subjects were then asked to report their partisan identification and ideology. Ideology was measured on a scale ranging from extremely liberal to extremely conservative. Subjects were then asked to place George W. Bush, Al Gore, and the two major political parties on the ideology scale. At this point, participants were presented with the stimulus that provided policy information for George W. Bush and Al Gore.

The information was presented in either a *candidate-centered* format, where subjects saw all of one candidate's views followed by all of the other candidates' views, or a *dimension-centered* format, in which an issue was presented followed by both candidate's views on that issue (see above). The information was taken from the candidates' official internet websites. The information presented in both stimuli was identical. Immediately after the stimuli, subjects were asked again to place George W. Bush and Al Gore on ideology scales. To serve as a distractor task, the participants were surveyed on a number of non-political items before being presented with a quiz on the information that had been presented in the stimuli. Finally, subjects were again asked the series of approve-disapprove questions. As before, reaction times were taken for the approval questions.

Creation of Variables

Structure (format). Respondents viewed information of Al Gore's and George W. Bush's policy positions on four issues (see Appendix A). The policy position information was presented in either in a candidate-centered format (coded 0) or in a dimension-centered format (coded 1).

Sophistication. Rahn, Aldrich, and Borgida (RAB) based their sophistication measure on the established composite method. An index was created by standardizing and summing four variables: self-reported interest in politics, exposure to television news, behavioral participation in political activities, and the number of correct answers to a five-question political information test. These four variables were collapsed into a dummy variable along the median score: 0 = "non-sophisticates" and 1 = "political sophisticates." The sophistication measure used in the test of RAB's model presented here is also a composite score, although different components and summing techniques were used.

For the research presented here, pre-test interviews with non-participants revealed that individuals who reported similar exposure levels to television news were often referring to different types and qualities of newscasts. For example, some who reported high levels of exposure to television news were referring to watching a few minutes of CNN's Headline News several times a day, while others meant never missing the local news, while yet others referred to being "CNN junkies." Consequently, I made a judgment that the inclusion of a question addressing exposure to television news in the composite sophistication measure might lead to intrusive "noise" in the measure. Moreover, the mediating nature of political efficacy for the relationship between sophistication and participation in political activities led me also not to use the media exposure question. In the final wash, it was decided that sophistication would best be

tapped by using base measures of political knowledge coupled with self-reported interest in politics.

The measure of political sophistication used in the alternate model was created by averaging the scores from two political knowledge tests and weighting this average with the subjects' political interest. The first test asked respondents to identify the partisan affiliation of several prominent political actors, while the second asked subjects to match policy positions to the correct party. These scores were collapsed into a dummy variable along the median score: 0 = "non-sophisticates" and 1 = "political sophisticates."

One might argue that by collapsing the political sophistication variable, valuable information inherent in its original variance is lost. The question appears to be whether the trade-off between subtlety and parsimony is valid. In fact, the question is a bit more complicated. A defense of the use of a collapsed measure of sophistication along a median split begins with the observation that a collapsed measure is "a standard measure of political sophistication" (Lodge, Steenbergen, and Brau 1995: 314). A collapsed measure of sophistication is the norm across seminal works that serve as a standard for this research (see Lodge, McGraw, and Stroh 1989; McGraw, Lodge, and Stroh 1990; Zaller 1990; Rahn, Aldrich, and Borgida 1994; Lodge, Steenbergen, and Brau 1995). Moreover, while it is possible to gain greater nuance with more subtle gradations, the collapsed measure presents a direct comparison with the model of evaluation created by Rahn, Aldrich, and Borgida (1994). Use of a collapse measure allows for more direct comparisons with their work.

Party Identification. While both this research and that by RAB use the traditional seven point partisan scale, there is a key difference in how the measures were constructed. This research used the traditional NES format and battery of questions. A score of "1" represents a strong Democrat and a score of "7" repre-

sents a strong Republican. RAB used one question showing respondents all seven points on the partisanship scale; the result was an astounding 75% of RAB's respondents claiming political independence. The use of the multi-question NES format resulted in only 17.83% of the subjects of this research asserting that they were political independents. The implications of the difference are discussed further in the examination of the results.

Policy Similarity. RAB asked subjects to place themselves on five seven-point policy scales. All those responding more conservatively, on the "Republican" side of the scale, were coded "1," those in the center, "0," and those on the "Democratic" side of the scale were coded "-1." These scores were summed to create a +5 to -5 scale representing proximity to Republican or Democratic policy positions, respectively. The score served as a proxy for proximity to the fictitious Republican and Democratic candidates of the RAB study.

For the alternate model presented here, subjects were asked to place themselves on an ideology scale immediately after answering questions about their policy preferences. Subjects were also asked to place George W. Bush and Al Gore on ideology scales immediately after reading the stimuli containing their policy positions. A relative proximity score was created by subtracting the absolute value of the difference between the respondent's ideology and his or her perception of Gore's ideology from the absolute value of the difference between the respondent's ideology and his or her perception of Bush's ideology. More than just similarity, this calculus creates a relative, or comparative similarity score. Higher scores reflect a greater proximity to Gore. In the formula, I_B represents respondents' placement of Bush on the ideology scale, I_G represents their placement of Gore, and I_R represents their own self-placement:

$$\left|I_B - I_R\right| - \left|I_G - I_R\right|$$

Comparative Memory. The creation of this variable marks one of the more notable differences between this model and that of RAB. The alternate model includes a memory variable. The memory variable reflects respondents' ability to remember more facts about one candidate than about the other candidate, weighted by their approval of each candidate. After completing a distractor task, respondents were quizzed on the information they had seen in the stimuli. The comparative memory variable was created by subtracting the number of correct Bush/policy matches weighted by Bush's approval rating from the number of correct Gore/policy matches weighted by Gore's approval rating. Approval was coded "1" while disapproval was coded "-1." This score was weighted by the overall percentage correct on the post-stimulus quiz.

RAB, on the other hand, created a memory score that was "positively [or] negatively valenced" in a fundamentally different manner. Respondents were asked to list everything that they could remember about the candidates. They were then instructed to note whether they considered these items positive, negative, or neutral. A "net" memory score was created for both candidates by subtracting the number of recalled items considered negative from those considered positive. The net score for one candidate was subtracted from the other.

It is a standard practice to incorporate participants' affective responses to what they recall, but the method used in the alternate model is much more latent than that employed by RAB, making it more accurate and less prone to *post hoc* justification by the respondents. In the alternate model, information recalled is valenced by the underlying level of approval for the candidate who espouses the recalled position.

Comparative Candidate Evaluation. The biggest difference between RAB and the alternate model is in the dependent variable. First, and most importantly, the alternative model uses the

evaluation of real candidates in lieu of fictitious ones. Participants were asked to evaluate George W. Bush and Al Gore. The experiment was conducted in November 1999, at a time when each man was considered the front-runner for his party's nomination, but no national debates had taken place.

RAB asked respondents to rate fictitious candidates on a seven point like-dislike scale. The comparative candidate evaluation was constructed by subtracting one candidate's score from the other. The alternate model uses a measure of attitudinal strength based on attitude accessibility as discussed above. After reading their respective policy positions in the stimuli, respondents were asked whether they approved or disapproved of Bush and Gore. Their response latencies to these questions were measured in order to gauge the accessibility of their opinions. Since shorter reaction times equate to greater accessibility, all of the latencies were subtracted from the highest score, that is, the longest reaction time. The procedure reversed the accessibility measure so that higher scores equated to greater accessibility.

Approval was coded "1" for "approve," "-1" for "disapprove," and "0" for "don't know." The approval score was then multiplied by accessibility to create a measure of evaluative strength for each candidate. Bush's evaluation was subtracted from Gore's to create a comparative candidate evaluation, standardized by dividing by the average accessibility for both candidates. In the formula, G_{AP} and G_{AC} represent approval and accessibility for Gore, respectively, while B_{AP} and B_{AC} represent the same measures for Bush:

$$\frac{\left(G_{AP} * G_{AC}\right) - \left(B_{AP} * B_{AC}\right)}{\frac{G_{AC} + B_{AC}}{2}}$$

Allowing "don't know" as a response, and coding it as "0," makes an important distinction in the creation of the variable.

Not forcing respondents to espouse an opinion that they may not really hold reduces measurement error. Further, these non-attitudes are, indeed, "zeroed out" of the attitudinal strength calculus by coding them as "0"; therefore, it is impossible to have either strong or weak non-attitudes, a fact that controls the introduction of measurement error by non-attitudes as noted by Converse (1970).

Questions about Rahn, Aldrich, and Borgida's (1994) Model

It is likely that several of RAB's findings are the result of how they constructed their model. After being introduced to the "candidates" for the first time in the video stimulus, subjects were asked to rate each candidate on a seven point like-dislike scale. Their dependent variable, comparative candidate evaluation, was constructed by subtracting one candidate's score from the score of the other candidate. The primary bases for liking or disliking a candidate would have been partisanship and an opinion formed from the policy positions they just heard each candidate espouse. There were no other available tools for evaluation, except, perhaps, visual affect, which the model is not designed to measure.

The key independent variables were (a) an interaction between the format (structure in which the information was presented) and comparative candidate memory and (b) an interaction between these two variables and the subjects' political sophistication. However, there may be a problem in the construction of the comparative candidate memory variable, the key variable in both of these interaction terms. In RAB, comparative candidate memory was constructed by having the subjects list everything they could remember about what they had just heard or seen. Subjects then went back and noted whether information they had listed about the candidates was positive, negative, or neutral. The number of negative items for each candidate was

subtracted from the number of positive items. Finally, one candidate's "net" memory score was subtracted from the other candidate's "net" memory score.

The dependent variable measures whether subjects *like* or *dislike* the candidates based on the policy positions they just heard. A primary component of the comparative candidate memory score is based on whether subjects *like* or *dislike* what they just heard: the policy positions of the candidates. Since the bases of both of these variables are so intricately interconnected, it seems unlikely that any significant relationship between them, or any interactive term of which they are components, could be attributed accurately to anything other than their common roots. The broader implication of this fact is that the model, although ingeniously conceived, may not offer a fair test of memory-based versus impression-driven evaluation because of potential bias in the variable construction. Further, if there is a question about the construction of key variables, the test of the influence of information structure, an important contribution to the literature on evaluation, may be flawed as well. Simply put, the use of fictitious candidates, for whom it is *impossible* to test a "running tally," the construction of the memory variable that mimics that of the dependent variable, and the non-traditional construction of the partisanship variable, which, as noted above, netted 75% of participants claiming to be independents, limits the ability of the RAB model to test the original hypotheses.

Expectations and Results of Model Comparison

This analysis offers a more realistic test of the RAB model, using variables that are more naturally constructed, yet controlled by experimental design. Based on research from the field of cognitive psychology, the dependent variable used here reflects the strength of respondents' evaluations of Gore relative to the strength of their evaluation of Bush. The alternatively created

comparative memory score is more latently valenced, and, therefore, offers a direct test of the impact of information recall on candidate evaluation without the confounding effect of introducing an evaluation of the information provided by the candidate.

The results of the model comparison may be found in Table 1. In the alternate model, only partisanship and policy similarity drive comparative candidate evaluation, explaining more than half of the model variance. The notable increase in amount of variance explained is likely due to the use of the traditional three part NES question to determine partisanship as opposed to showing subjects all seven points on the partisanship scale and allowing them to position themselves on the scale. Note that because the alternative model uses Gore as the baseline candidate, and higher numbers on the partisanship scale represent Republicans, the relationship between evaluation and partisanship should be negative; the directional difference with RAB reflects their use of the Republican as the baseline candidate. The increase in the size and significance of the policy similarity variable is likely the result of a direct comparison to subjects' perceptions of Gore's and Bush's policy positions as opposed to a broader similarity to the general policies of the candidates' parties.

The use of a variable measuring the quality and content of memory, without a *conscious* evaluation of its contents, shows that recall memory does not appear to be a driving factor in candidate evaluation. However, RAB's findings should not be readily dismissed. Their assumptions were based on solid theory. Perhaps the failure of this variable in the replication is related to a problem with the dependent variable in the alternate model.

TABLE 1
Regression on Comparative Candidate Evaluation on Memory, Information Structure, and Political Sophistication

Variable	RAB Model*	Alternate Model
Partinship	.28+ (.16)	-.29§ (.06)
Policy Similarity	.18+ (.09)	.17§ (.05)
Comparative memory	.02 (.11)	.00003 (.00002)
Comparative Memory • Sophistication	.16 (.14)	-.00001 (.00002)
Comparative Memory • Structure	.32↕ (.12)	-.00004 (.00003)
Comparative Memory • Sophistication • Structure	-.47↕ (.17)	.00005 (.00003)
Sophistication	-.68 (.56)	.52 (.29)
Structure	-.18 (.54)	.03 (.30)
Sophistication • Structure	-.37 (.81)	-.11 (.43)
Adjusted R^2	.33	.54
N	104	101

*Reproduced from Table 1, p.198 in Rahn, Wendy M., John H. Aldrich, and Eugene Borgida. 1994. "Individual and Contextual Variations in Political Candidate Appraisal." *American Political Science Review* 88:193-99.

Note: Cell entries are unstandardized coefficients; standard errors are in parentheses. +p < .05 ↕p < .01 §p < .001

An "Update" Model of Evaluation

The alternate model presented above is open to the criticism that all of the relationships are confounded by the influence of prior opinion, the reason RAB used fictitious candidates. For example, an individual's opinion of Gore or Bush may very well be memory based, but drawn from the wealth of information received prior to the introduction of the stimulus and, therefore, would show no relationship to a memory score based solely on recollection of the stimulus.

The following formula controls for the influence of prior opinion on candidate evaluation. Comparative candidate evaluation is measured before any new information is introduced to the subjects, as well as after the introduction of the controlled information. The dependent variable in this model represents change in comparative candidate evaluation from "before" (T_1) to "after" (T_2) the presentation of new information, the introduction of which is controlled by experimental design. Both pre- and post-stimulus comparative candidate evaluations are standardized to allow comparability and control for decreased response time based solely on familiarity with experimental procedure. By design, any change in comparative candidate evaluation is the result of the introduction of this information. In the formula, G_{AP1} and G_{AC1}, and B_{AP1} and B_{AC1}, represent approval and accessibility for Gore and Bush prior to the presentation of the stimuli while G_{AP2} and G_{AC2}, and B_{AP2} and B_{AC2}, indicate the same measures after the presentation of the stimuli:

$$\left(\frac{\left(G_{AP2} * G_{AC2}\right) - \left(B_{AP2} * B_{AC2}\right)}{\frac{G_{AC2} + B_{AC2}}{2}} \right) - \left(\frac{\left(G_{AP1} * G_{AC1}\right) - \left(B_{AP1} * B_{AC1}\right)}{\frac{G_{AC1} + B_{AC1}}{2}} \right)$$

The formula posits that in a more realistic setting, individuals integrate new information by checking this information against prior "impressions," that is, individuals' previously created proximity assessments. Individuals' initial assessment of the candidate (subconsciously logged impressions of the candidate's ideological and policy positions) relative to individuals' established ideological preferences must be consulted in light of new information. As they update the assessment of their proximity to the candidates based on new information, they will update their evaluations.

The ability to update impressions is likely to be mediated by sophistication, dependent on information structure, and individuals' relative recall ability. Measuring these phenomena requires the use of a mixed evaluative strategy. When asked to create an evaluation, individuals compare their existing, or "on-line," assessment to new information in their memories. The updated model draws on research by Lodge, Steenbergen, and Brau (1995) to posit the mixed strategy. Lodge, Steenbergen, and Brau (1995) highlight the importance of campaign message, but note that the utility of memory decays as people forget specific items. A mixed strategy is adopted in the updated model because it is believed that individuals will draw from direct memories of campaign messages if they have any, but will also rely on the relation of new memories to previously created impressions; impressions formed, in part, from forgotten memories of previous messages.

To test the updated theory, a variable measuring the proximity reassessment function occurring before the evaluative update is needed. A variable was created reflecting the policy similarity correction. The variable reflects the change in individuals' assessments of policy proximity from before and after the introduction of new information. A measure of similarity (described above) is taken before (denoted with a subscript "1") the intro-

duction of the stimulus, representing the current assessment, or "tally," and another is taken after the introduction of the stimulus (denoted with a subscript "2"). The difference between the before and after measures is the policy similarity correction score. It represents the impact of the new information on the assessment of similarity.

In the formula, I_{G1} and I_{B1} represent the respondent's placement of Gore and Bush prior to the introduction of the stimuli, while I_{G2} and I_{B2} reflect respondent's placement of Gore and Bush after viewing the stimuli. Respondents' self-placement is represented by $I_{R.}$. Measures of partisanship, sophistication, structure, and comparative memory are the same as above:

$$\left(\left|I_{B2}-I_R\right|-\left|I_{G2}-I_R\right|\right)-\left(\left|I_{B1}-I_R\right|-\left|I_{G1}-I_R\right|\right)$$

The full model posits

> Change in Comparative Candidate Evaluation = a + b_1(Partisanship) + b_2(Comparative Memory) + b_3(Structure) + b_4 (Sophistication) + b_5(Policy Similarity Correction [PSC]) + b_6(PSC * Sophistication) + b_7(Comparative Memory * Sophistication) + b_8(PSC * Structure) + b_9 (PSC * Comparative Memory).

Partisanship is still expected to play a significant role; again, because of the coding scheme, the coefficient should be negative. Comparative memory is expected to be significant due to the importance of campaign message recall in evaluation. However, because of the construction of the dependent variable, the sign should be negative. If respondents can produce an evaluation as quickly (creating no significant impact) or more quickly (inducing a positive sign) then it would seem that respondents are not integrating new information in their memory. However, if they

are drawing from their memory, they should have to pause as new information is integrated into candidate evaluation. Policy similarity correction should have a negative impact as well. If a notable reassessment of one's relative proximity to the candidates must occur, then it seems logical that one must pause to do this. Whether the process is significant likely hinges on the importance of the interaction of proximity variable with sophistication and with structure. On their own, structure and sophistication are not expected to be significant since their influence is likely to be mediated by other variables in the model, hence the interactive terms.

The interaction between policy similarity correction and sophistication should be significant and negative. Respondents who have a greater discrepancy between pre- and post-stimulus policy assessment should pause as they re-evaluate the candidates; however only the sophisticated have the ability to reevaluate, meaning that the "pause" should occur only among the sophisticated. A deeper understanding of the process might be seen by comparing it to findings by McGraw, Lodge, and Stroh (1990) regarding information overload and non-sophisticates. In their model, those who are overloaded by a change in policy similarity correction discount the change allowing the post-stimulus comparative evaluation to be made as quickly, or perhaps even more quickly because of other factors. Another way to put it is that the unsophisticated will fail to recognize that their previous assessment was in error; therefore, no change should be seen for them.

The interaction between comparative memory and sophistication should be positive and significant. The politically sophisticated should be better equipped to apply the newly acquired information to their evaluation.

The interaction between policy similarity correction and structure should be positive and significant if information format truly makes a difference. The argument is that the debate-style

format makes it easier to catalog information in order to determine whether a correction of the respondent's understanding of policy similarity is necessary. The debate format allows for the easier creation of a "checklist" of policy position proximities. Because of the "forgetting curve" (see Lodge, Steenbergen, and Brau 1995), the previous proximity assessment is likely the result of an impression driven evaluation. If that evaluation is to be corrected in light of the memory of recently acquired information, then this process should be facilitated by a format allowing for maximum distinction between the candidates on each policy issue. The result would be the use of a mixed strategy for evaluation facilitated by the format of the information. Previous impressions are being updated with new information from recent memory and the debate-style format presents the information in the most accessible manner.

The interaction between policy similarity correction and comparative memory should be significant and negative. Higher scores on the component variables of the interaction indicate that respondents having more information in recent memory to consult must do more correction to their previous proximity assessment. Comparisons between earlier and later information should take longer, resulting in a negative relationship with the dependent variable. However, because a mixed strategy is theorized, a variable representing both memory-based information and impression-derived information should be significant.

Findings

The results of the "update" model may be found in Table 2. As expected, partisanship still plays a significant role in evaluation. On its own, sophistication did not reach statistical significance: this was expected because of the influence of sophistication in the interactions. The same is true for the structure variable. Also as expected comparative memory was signifi-

cant and in the expected direction, lending some support to the memory based approach suggested in RAB. Policy similarity correction, however, did not reach statistical significance, indicating that if an on-line ("impression-driven") strategy is used, it is mediated by some other variable and/or is part of a mixed strategy.

Lending evidence to the possibility of the use of a mixed strategy, the interaction between policy similarity correction and sophistication worked as expected, meaning that sophisticated individuals are more likely to *pause* as they update their comparative candidate evaluation in order to integrate their new assessment of policy similarity. However, the pause effect only occurs when individuals must adjust, or *correct*, their assessment of the candidates' ideologies after being given information on their policy positions. Further, the pause only happens if individuals are sophisticated enough to realize that their original assessment was in error. We are almost seeing a snapshot of individuals' thought process as they reevaluate candidates in light of new information.

Politically sophisticated respondents do seem more likely to pause and ponder the implications of erroneous previous impression driven (or "on-line") evaluations as they update their evaluations. Finding evidence for the on-line strategy here as well as the memory based strategy demonstrated above seems to confirm the hypothesis of the use of a mixed strategy.

The interaction between comparative memory and sophistication was significant and in the expected direction. Applying recently acquired information to an evaluation is easier for those who are more politically sophisticated. As evidenced by this and the previous variable, the ability to use efficiently a mixed strategy seems to be related to political sophistication.

The interaction between policy similarity correction and structure was significant and in the expected direction, as well. If

respondents need to reassess their relative policy proximity to candidates, the debate-style format, highlighting the comparative policy positions of the candidates, makes it easier to do so.

Finally, the interaction between policy similarity correction and comparative memory (variables representing an impression-driven evaluative strategy and a memory-based strategy, respectively) is significant, and in the predicted direction. The results produce further evidence of the use of a mixed strategy. If individuals are applying information from recent memory to an

TABLE 2
An "Update" Model of Candidate Evaluation
Regression of Change in Candidate Evaluation on Policy Correction, Information Structure, Memory, & Political Sophistication

Independent Variables	Unstandardized Coefficients
Partisanship	-.158↕ (.002)
Comparative Memory	-.00003↕ (.00001)
Structure	.002 (.210)
Sophistication	.282 (.225)
Policy Similarity Correction	-.02 (.09)
Policy Similarity Correction • Sophistication	-.247+ (.109)
Comparative Memory • Sophistication	.00004+ (.00002)
Policy Similarity Correction • Structure	.240+ (.1003)
Policy Similarity Correction • Comparative Memory	-.00001↕ (.000004)
Adjusted R2	.24
N	101

Note: Cell entries are unstandardized coefficients; standard errors are in parentheses. +p < .05 ↕p < .01

evaluation that is significantly different from earlier impressions of the candidates, then the comparison takes more time, causing pauses as candidate evaluations are updated.

Conclusions

The research presented here accomplishes two things. First, it has replicated the work by Rahn, Aldrich, and Borgida (1994) using real candidates and alternatively created variables. Second, it builds upon RAB and other research to offer a model of candidate evaluation that maps the processes used by individuals to update their comparative evaluations. In so doing, the "update" model offers support for a mixed strategy of candidate evaluation.

The reexamination of RAB's model calls into question the construction of several of their variables and several of their findings. However, much of their theory and many of their assumptions appear to be correct. The common roots of the construction of comparative candidate evaluation and comparative candidate memory cast doubt upon the accuracy of their findings regarding the relationship between these two variables and the relationships between comparative candidate evaluation and all interactive independent variables of which comparative candidate memory is a component. However, the difference is with their measures, not their theory. Abandoning the standard NES battery of questions measuring partisanship in favor of showing respondents all seven points of the partisanship scale at once, raises concerns about the validity of their partisanship variable. The minimal role of partisanship in their model is probably because of variable measurement. Moreover, the use of fictitious candidates biases the model against the possibility of finding evidence for an impression-driven approach to candidate evaluation: respondents cannot have an impression evaluating candidates of whom they have never heard. Therefore, even a failure

to reject the null hypothesis that comparative candidate memory does not drive the comparative evaluation of fictitious candidates in RAB's model cannot be interpreted as supporting the theory of impression-driven evaluation. Finally, their memory score is likely to show a significant relationship to their dependent variable because of the construction of both variables

Although this article questions several about their measures, it does not cast doubt on Rahn, Aldrich, and Borgida's theoretical contribution to the literature. Their model may not include a full test of memory-based versus impression-driven evaluation, but it does highlight the important role that information context, or format, plays in the evaluative process. The "update" model of evaluation presented here builds upon and expands that contribution.

All models of evaluation that use real political figures as their objects of appraisal are open to the criticism that the researcher cannot determine if the reported evaluation is the product of his or her independent variables, or simply reflects a long held opinion that preceded the measures of the model. The update model presented here controls for this problem by measuring comparative candidate evaluation both before and after the introduction of the stimulus information. By subtracting the measure taken prior to the introduction of new information from the post-stimulus measure, one is left with a variable that measures the effect of current information on candidate evaluations. Individuals *do* draw on recent memory to make evaluations, but there is evidence that they consult previous impressions as well.

Realizing that the information presented in the study might differ from what subjects "thought they knew" about the candidates, this analysis measures the change in individuals' perceived proximity to candidates. The policy similarity correction interacts with sophistication, structure (information format), and comparative memory. The interaction between perceived prox-

imity to a candidate and political sophistication demonstrates that receiving information incongruous with previous assessments produces a discrepancy more readily recognized by those sophisticated enough to pick up on the inconsistency.

It appears that individuals assess new information to update their candidate evaluations; however, the ability to do so is mediated by sophistication and facilitated by the "pre-sorted," checklist nature of information provided in a dimension-centered, or debate, format. Information received this way is more readily applied to proximity assessments that are then used to update evaluations, or the information is examined for discrepancies, depending on individual sophistication. The update model lends support to both the memory-based and on-line evaluation debate, suggesting that individuals tend to use a mixed strategy of evaluation. Although the function of recall memory is important on its own, it also serves as a tool to update the previous on-line tally. Both the on-line impression and memory recall play a role in the creation of an evaluation. It is possible that as the salience of the new information diminishes, its accessibility in the memory fades eventually leaving only an impression of the assessment created because of it. If true, then the use of an on-line, memory-based, or mixed strategy may well depend, in part, on a time and salience factor. Hopefully, future research will be able to answer this question.

Not addressed in the update model is how long the specifics of the new information remain in the memory of individuals and at what point the new information becomes part of the overall "impression" of candidates while the details of the information are lost. The loss of memory detail is referred to as "decay" or the "forgetting curve" (Lodge, Steenbergen, and Brau 1995). However, the finding of a mixed strategy of evaluation does add to the memory-based versus impression-driven debate (see Hastie and Park 1986; Lodge, McGraw, and Stroh 1989; Lodge,

Steenbergen, and Brau 1995; McGraw, Lodge, and Stroh 1990; Rahn, Aldrich, and Borgida 1994; Bargh and Thein 1985). Naturally, since evaluation is such an important part of vote choice this has direct bearing on models of vote choice.

Another finding of this research is that the format in which information is received makes a difference in individuals' ability to process information and recall it to update their existing candidate evaluations. The important implications for this finding relate to the role of the media and campaigns in the calculus of individual voters. It has already been established that the structural features of the presentation of information in television impact viewer attentiveness involuntarily (Lang 1990; Thorson, Reeves, and Schleuder 1985). There is also ample evidence that campaign messages have a substantial impact on evaluations (see Lazarsfeld, Berelson, and Gaudet 1948; Bartels 1993; Marcus and MacKuen 1993; Lodge, Steenbergen, and Brau 1995). This impact is independent of recall. In fact, if only recall played a role in evaluations, campaigns would matter little as time robbed them of their effect on decision-making with regard to long-term (non-election specific) evaluations (Lodge, Steenbergen, and Brau 1995).

Given the role of campaigns, however, these findings imply that new information garnered in campaigns is checked against existing impressions for comparability and used to update one's evaluation and this process is facilitated when the information is received in a dimension-centered format. Campaign information is relevant as it contributes to the impression used to create the evaluation and more effective when it is presented in the most digestible form. However, political sophisticates are more likely to give pause while creating an updated evaluation when the information received runs counter to their previously established impression. Sophisticates may be "screening" information for congruity, but their ability to do so may be related to the format

of the information. This greatly increases the importance of not only campaigns, in general, but also specific *types* of campaigns in the formation of political evaluations and, through evaluation, voter choice.

My findings have implications beyond those directly tested in the models. For example, one of the broader implications regarding information context, or structure, might speak to the way the media cover political campaigns and candidates. If information is received in a dimension-centered format, it appears to aid information processing as individuals reassess candidates and seek to comparatively evaluate them, is there an obligation on the media—or the candidates themselves—to provide information in the most digestible form? Given the fact that sophistication mediates the impact of policy reassessment on evaluation, what are the obligations to those with less political savvy? Such normative questions are left to future research.

APPENDIX A

Note: All information is taken from official candidate websites.

Candidate-Centered Stimulus

George W. Bush. George W. Bush has issued several statements concerning his position on some of the most important issues of the presidential campaign of 2000. Bush has stated that there should be no reduction in Social Security benefits for retirees or near-retirees. He also opposes any tax increase for Social Security and is against the investment of Social Security funds in the stock market by the government. With regard to Medicare, Bush supports providing more choice and more private sector alternatives for the elderly. He has also spoken in favor of medical savings accounts and offering prescription drug benefits to Medicare recipients. On gun control, Bush opposes the government mandated registration of all guns owned by law abiding citizens and supports their constitutional right to own guns. He has spoken in favor of the stronger enforcement of existing gun laws instead of the creation of additional ones. Bush is in favor of instant background

checks at gun shows. On the topic of education, Bush supports a "school voucher" type program that would give parents federal funds to be used for public and private school choice and innovative education programs. He further supports implementation of state accountability systems in which students are tested every year in grades 3-8 in reading and math in order to gauge the effectiveness of school systems.

Al Gore. Al Gore has addressed his issue positions on several of the most important issues of the presidential campaign of 2000. In order to keep Social Security solvent, Gore supports the investment of part of the budget surplus in the stock market and scaling back some benefit levels. He would also like to eliminate the limit on what older Americans on Social Security can earn. On Medicare, Gore has spoken in favor of a prescription drug option for all Medicare beneficiaries. He has also suggested eliminating the deductible and all co-payments on preventive benefits covered by Medicare. With regard to gun control, Gore supports the passage of new laws designed to keep guns off the streets. In the past, Gore worked to help enact a ban on assault weapons. He also has supported the Brady Law which requires background checks for gun purchases and imposes a five day waiting period for those states without "instant background checks." On the topic of education, Gore opposes "school voucher" programs, instead supporting additional funding made available to states for the purpose of reducing drop-out rates and increasing student achievement in failing school districts. Gore also wants to use technology to increase educational productivity and increase the number of computers in public schools.

Dimension-Centered Stimulus

Please read the following information about the policy positions of the two leading candidates for the Presidential Election of 2000.

Social Security. George W. Bush: has stated that there should be no reduction in Social Security benefits for retirees or near-retirees. He also opposes any tax increase for Social Security and is against the investment of Social Security funds in the stock market by the government.

Al Gore: In order to keep social security solvent, Gore supports the investment of part of the budget surplus in the stock market and scaling back some benefit levels. He would also like to eliminate the limit on what older Americans on Social Security can earn.

Medicare. With regard to Medicare, Bush supports providing more choice and more private sector alternatives for the elderly. He has also spoken in favor of medical savings accounts and offering prescription drug benefits to Medicare recipients.

Al Gore: On Medicare, Gore has spoken in favor of a prescription drug option for all Medicare beneficiaries. He has also suggested eliminating the deductible and all co-payments on preventive benefits covered by Medicare.

Gun Control. George W. Bush: On gun control, Bush opposes the government mandated registration of all guns owned by law abiding citizens and supports their constitutional right to own guns. He has spoken in favor of the stronger enforcement of existing gun laws instead of the creation of additional ones. Bush is in favor of instant background checks at gun shows.

Al Gore: With regard to gun control, Gore supports the passage of new laws designed to keep guns off the streets. In the past, Gore worked to help enact a ban on assault weapons. He also has supported the Brady Law which requires background checks for gun purchases and imposes a five day waiting period for those states without "instant background checks."

Education. George W. Bush: On the topic of education, Bush supports a "school voucher" type program that would give parents federal funds to be used for public and private school choice and innovative education programs. He further supports implementation of state accountability systems in which students are tested every year in grades 3-8 in reading and math in order to gauge the effectiveness of school systems.

Al Gore: On the topic of education, Gore opposes "school voucher" programs, instead supporting additional funding made available to states for the purpose of reducing drop-out rates and increasing student achievement in failing school districts. Gore also wants to use technology to increase educational productivity and increase the number of computers in public schools.

REFERENCES

Bargh, J.A., S. Chaiken, R. Govender, and F. Pratto. 1992. "The Generality of the Automatic Activation Effect." *Journal of Personality and Social Psychology*. 62:893-912.

Bargh, J.A., and R.D. Thein. 1985. "Individual Construct Accessibility, Person Memory, and the Recall-Judgment Link: The Case of Information Overload." *Journal of Personality and Social Psychology*. 49:1129-46.

Bartels, Larry M. 1993. "Messages Received: The Political Impact of Media Exposure." *American Political Science Review*. 87:267-285.

Doll, J. and I. Ajzen. 1992. "Accessibility and stability of Predictors in the Theory of Planned Behavior." *Journal of Personality and Social Psychology*. 63:754-65.

Downing, J.W., C.M. Judd, and M. Brauer. 1992. "Effects of Repeated Expressions on Attitude Extremity." *Journal of Personality and Social Psychology*. 63:17-29.

Fazio, R.H. 1993. Variability in the Likelihood of Automatic Attitude Activation: Data Re-Analysis and Commentary on Bargh, Chaiken, Govender, and Pratto (1992). *Journal of Personality and Social Psychology*. 64:753-58, 764-65.

Fazio, R.H., J. Chen, E.C. McDonel, and S.J. Sherman. 1982. "Attitude Accessibility, Attitude-Behavior Consistency, and the Strength of the Object-evaluation Association." *Journal of Experimental Social Psychology*. 18:339-57

Fazio, R.H, and C.J. Williams. 1986. "Attitude Accessibility as a Fazio, R.H., D.M. Sanbonmatsu, M.C. Powell, F.R. Kardes. 1986. "On the Automatic Activation of Attitudes." *Journal of Personality and Social Psychology*. 50:229-38.

Moderator of the Attitude-Perception and Attitude-Behavior Relations: An Investigation of the 1984 Presidential Election." *Journal of Personality and Social Psychology*. 51: 504-14.

Hastie, Reid and Bernadette Park. 1986. "The Relationship Between Memory and Judgement Depends on Whether the Task is Memory-Based or On-Line." *Psychological Review*. 93:258-268.

Houston, D.A., and R.H. Fazio. 1989. Biased Processing as a Function of Attitude Accessibility: Making Objective Judgements Subjectively." *Social Cognition*. 7:51-66.

Krosnick, Jon A., and Richard E. Petty. 1995. "Attitude Strength: An Overview." In *Attitude Strength: Antecedents and Consequences*, eds. Richard E. Petty and Jon A. Krosnick. Mahwah, NJ: Lawrence Erlbaum Associates.

Lang, Annie. 1990. "Involuntary Attention and Physiological Arousal Evoked by Structural Features and Emotional Content in TV Commercials." *Communication Research*. 17:275-299.

Lazarsfield, Paul, Bernhard Berelson, and Hazel Gaudet. 1948. The People's *Choice: How the Voter Makes up His Mind in a Presidential Campaign.* New York:

Lodge, Milton, Kathleen M. McGraw, and Patrick Stroh. 1989. "An Impression-Driven Model of Candidate Evaluation." *American Political Science Review* 83:399-419.

Lodge, Milton and Marco R. Steenbergen with Sawn Brau. 1995. "The Responsive Voter: Campaign Information and the Dynamics of Candidate Evaluation." *American Political Science Review.* 89:309-326.

Lodge, Milton, Patrick Stroh, and John Wahlke. 1990. "Black-Box Models of Candidate Evaluation." *Political Behavior* 12:5-18.

Marcus, George, and Michael B. MacKuen. 1993. "Anxiety, Enthusiasm, and the Vote: The Emotional Underpinnings of Learning and Involvement During Presidential Campaigns." *American Political Science Review*. 87:672-685.

McGraw, Kathleen M., Milton Lodge, and Patrick Stroh. 1990. On-Line Processing in Candidate Evaluation: The Effects of Issue Order, Issue Importance, and Sophistication." *Political Behavior* 12:41-58.

Powell, M.C., and R.H. Fazio. 1984. "Attitude Accessibility as a Function of Repeated Attitudinal Expression." *Personality and Social Psychology Bulletin.* 10:139-48.

Rahn, Wendy M., John H. Aldrich, and Eugene Borgida. 1994. "Individual and Contextual Variations in Political Candidate Appraisal." *American Political Science Review* 88:193-99.

Roskos-Ewoldson, D.R. and R.H. Fazio. 1992. "The Accessibility of Source Likability as a Determinant of Persuasion." *Personality and Social Psychology Bulletin*. 18:19-25.

Sears, David O. 1986. "College Sophomores in the Laboratory: Influences of a Narrow Data Base on Psychology's View of Human Nature." *Journal of Personality and Social Psychology*. 51: 515-530.

Srull, Thomas K. and Robert S. Wyer, Jr. 1986. The Role of chronic and Temporary Goals in Social Information Processing." In *Handbook of Motivation and Cognition*, ed. Richard M. Sorrentino and E. Tory Higgins. New York: Guilford.

Szalay, L.B. and J. Deece. 1978. *Subjective Meaning and Culture: An Assessment Through Word Associations*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Thorson, Esther, Byron Reeves, and Joan Schleuder. 1985. "Message complexity and Attention to Television." *Communication Research*. 12:427-454.