



ACADEMIC  
PRESS

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Consciousness and Cognition 12 (2003) 257–278

Consciousness  
and  
Cognition

[www.elsevier.com/locate/concog](http://www.elsevier.com/locate/concog)

## Halo and devil effects demonstrate valenced-based influences on source-monitoring decisions

Gabriel I. Cook,<sup>a</sup> Richard L. Marsh,<sup>a,\*</sup> and Jason L. Hicks<sup>b</sup>

<sup>a</sup> *Department of Psychology, University of Georgia, Athens, GA 30602-3013, USA*

<sup>b</sup> *Department of Psychology, Louisiana State University, Baton Rouge, LA 70853-5501, USA*

Received 22 April 2002

---

### Abstract

Source attributions based on positive versus negative valence were examined in four experiments. The two sources were individuals who were depicted positively or negatively, and the content of their statements was similarly valenced. When valenced information about the sources was provided after learning the statements, test biases to attribute positive statements to the positive source and negative statements to the negative source were strongly present. Providing the same information prior to learning improved memory, but did not entirely eliminate test biases based on valence. Signal detection analysis suggests that these “halo effect” biases are criterion-based and not memory-based. Therefore, the results are more consistent with descriptions of source-monitoring processes that can benefit from familiarity-based partial information as opposed to descriptions in which source monitoring is primarily recollection-based.

© 2002 Elsevier Science (USA). All rights reserved.

---

### 1. Introduction

People are often influenced by their own subjective theories concerning other people's personalities. When all that is known about a person is one positive trait

---

\* Corresponding author. Fax: 1-706-542-3275.

E-mail address: [rmarsh@uga.edu](mailto:rmarsh@uga.edu) (R.L. Marsh).

(e.g., high intelligence), a strong tendency exists to extrapolate that positive characteristic to other personality characteristics about which nothing is objectively known (e.g., neat, honest, friendly, socially adept, etc.). Thorndike (1920) coined the term “halo effect” to describe this erroneous extension of positive beliefs. He also coined the opposite phenomenon the “devil effect” in which observing one bad quality tends to create the belief structure that the person must have other negative qualities as well. As was shown by Nisbett and Wilson (1977) in their citation classic study, the basic halo effect can be readily obtained in the laboratory. In two different versions of a video tape, the same Belgian professor was initially depicted as friendly or arrogant. In the former case, participants rated traits like his physical appearance and mannerisms to be much more positive than in the latter case, despite being objectively the same. The specific question that we address in the current study is whether people’s source-monitoring attributions about who spoke statements would be influenced by the halo and devil effects.

In standard depictions of source monitoring, people identify the origin of information by inspecting qualitative characteristics that are stored in memory. These characteristics include perceptual information, emotional details, spatio-temporal data, semantic information, and records of internal (cognitive) operations of processing the information originally. When certain of these details are more plentiful than others, the original source of a memory can be specified relatively easily. For example, if the remembered details of information are imbued with lots of internal cognitive operations and very little perceptual detail, one’s self can be identified as the source as opposed to some external agent. The same sorts of characteristics can be inspected as a means of reducing the number of options among candidate sources. In this case, if information can be retrieved that a statement was uttered on a soccer field, the candidate sources are likely to be different from a statement remembered being spoken at a religious gathering.

In contrast to consulting qualitative characteristics, another basis by which a memory’s origin can be determined is by consulting schematic knowledge (e.g., Bayen, Nakamura, Dupuis, & Yang, 2000; Mather, Johnson, & De Leonardis, 1999). In these cases, statements can be attributed to people based on a schema, such as whether they are a doctor versus a lawyer, or a democrat versus a republican (also see Hicks & Cockman, *in press*). In this approach, a memory is attributed to a source based on its semantic content. For example, a statement such as “I have to be in court at 9 AM” is more consistent with one’s schema for an attorney although it is hypothetically possible for a doctor to have uttered it as well. Source attributions based on preexisting knowledge of gender, like occupations, also represents a semantic influence on origin identification (Marsh, Cook, & Durley, 2002).

In the present study, we wanted to determine whether there would be an analogous effect of emotional or affective (rather than semantic) consistency on source monitoring. Therefore, we presented positively and negatively valenced items from two hypothetical men (i.e., sources). One source was depicted as employed, married, and a college professor (i.e., positive) whereas the other source was depicted as unemployed, divorced, and a blue-collar worker (i.e., negative). All of the statements were slightly valenced events that befell the two sources (e.g., I got a flat tire driving home) or behaviors that they supposedly performed themselves (e.g., I gave the

homeless person a quarter). Given the existence of the halo and devil effects, we wanted to ascertain if source-monitoring attributions would be biased toward one source or another depending on the valenced content of the candidate memory being evaluated. If source-monitoring decisions are biased by the valence of the candidate memory being considered, then positively valenced items should be assigned more reliably to the source depicted in a positive light whereas negative items may gravitate toward the source depicted as less fortunate. Although the theoretical basis for the existence of valenced-based biases might not be fully identifiable with this one article, we do not know of any study that has investigated the potential bias to source attributions from evaluating the valenced content of a candidate memory.

The halo effect has mainly been studied in the context of attributions about personality characteristics. Therefore, biased attributions might be expected with personal behaviors that are consistent or inconsistent with personality information depicted for a given source. In these cases, more positive behaviors might be attributed to the successful source and more negative behaviors to the unsuccessful one. By contrast, because most everyone has experienced a flat tire or found a \$5 bill on the ground, there is no a priori reason to expect that positive versus negative external events should be biased in favor of one source versus another. For this reason, we began our investigation with behaviors that were associated with two different men (e.g., I cut to the front of the line; I helped the old woman across the street). We then tested whether any biases found with behaviors would extend to external (and uncontrollable) events that occurred to the two sources (e.g., My raffle ticket was the winning one; I tore my favorite shirt). We then removed the experimenter-provided descriptions of the two male sources as valenced and presented different proportions of positive versus negative behaviors and actions that befell the two sources. Our goal was to determine whether people would learn the correlations and then bias their source attributions based on what they had learned on their own (see Heit, 1992). In this fashion, we gradually reduced those conditions that might be deemed favorable to finding source attributions based on the halo and devil characterizations of the two sources. Finally, rather than depicting the valenced behaviors as associated with the two valenced sources, we depicted them as associated with a third party. Consequently, we predicted that the size of any biases that might result from the valenced content of memories would decrease across the experiments.

Theoretically, Hicks and Cockman (in press) demonstrated that when schemas were provided after learning and before the source test (for doctor versus lawyer sources) they replicated Bayen et al.'s (2000) effect of biased source attributions. In this case, the schema-consistent content of a candidate memory biased attributions in favor of a particular source. However, they also found that if the schema was presented before encoding, those biases disappeared. In that case, participants noticed the inconsistency of a doctor speaking a lawyer related item and their source memory was greatly improved. This effect is reminiscent of earlier work on schemas demonstrating elevated memory for schema-inconsistent events in a story (e.g., Graesser, Gordon, & Sawyer, 1979; Graesser & Nakamura, 1982). Therefore, manipulating whether the person descriptions were given prior to versus after learning will establish whether valenced content (i.e., emotional or affective) information operates in similar ways (or not) to established schemas.

## 2. Experiment 1

Equal numbers of positive and negative behaviors were presented from each of two hypothetical male sources during learning. After learning but before the source test, one source was depicted as successful and the other as down on his luck. If the halo and devil effects extend themselves to source-monitoring attributions, then positive behaviors associated with the negatively valenced source might be erroneously attributed to the positive source (and vice versa for the negative behaviors). By the same token, the bias to attribute studied items may also extend to new items such that when they are erroneously claimed to be old they will be attributed to the source with the same valence. The same biased source monitoring may not be observed if participants are provided with the person descriptions before they learn the behaviors (Hicks & Cockman, in press). If the biases do disappear, then the valence effects might be theoretically linked to the small literature on source monitoring and schemas, albeit in the present case with affective or emotional information.

### 2.1. Method

#### 2.1.1. Participants

Sixty-three undergraduates from the University of Georgia volunteered in exchange for partial credit toward a course requirement. Each participant was tested individually in sessions that lasted approximately 30 min. As described shortly, 32 participants were assigned haphazardly to the *before-study* condition and 31 were tested in the *after-study* condition, where the condition names reflect when the person/source descriptions were provided to participants.

#### 2.1.2. Materials and procedure

A total of 72 behaviors depicting either slightly positive or slightly negative conduct were developed (36 of each valence). Some of these were identical to those used by Sherman and Bessenoff (1999) and others were created in house.<sup>1</sup> The names Mark and Geoff were used to represent the two fictitious men. During learning, participants read 12 positive and 12 negative behaviors purportedly associated with each of the Mark and Geoff sources. Therefore, they read 48 behaviors that were randomly presented from the two sources. In this fashion, both sources were associated with equal numbers of each valenced type of behavior. The software randomly selected statements anew for each participant tested. A given study trial contained a male name in the center of the computer monitor with the behavior printed below it. Participants had 5 s to read the behavior before the computer chimed and the next study trial began automatically.

Descriptions of Mark and Geoff were presented either before studying or after studying the behaviors (and we have mnemonically labeled the conditions as such). Geoff was depicted as a college professor, happy, married, employed, and enjoyed drinking wine with friends. Mark was depicted as a blue-collar employee, unhappy,

---

<sup>1</sup> We appreciate Jeffrey Sherman sharing his stimuli with us.

divorced, temporarily out of work, and enjoyed drinking beer at a bar. Because the names Mark and Geoff seemed neutral enough, we did not counter-balance names of the positive and negative sources. All participants were asked to study the statement below the name on each trial for a later unspecified memory test. In the before-study condition participants were asked to read the two short statements about each individual before the study sequence ensued. In the after-study condition they were asked to do so after the study sequence but prior to the source-monitoring test. In all other procedural respects, the between-subjects conditions of before- versus after-study were identical. Both groups were asked to work on a distractor task for 5 min before receiving instructions for the source-monitoring task.

For the source test, the software randomly intermingled anew for each participant the 48 studied statements with the 24 new statements. The 24 new items were comprised of 12 positive and 12 negative behaviors drawn from the original sets of 36 positive and 36 negative behaviors (i.e., they had not been presented during the learning phase). Participants performed a three-alternative forced-choice test in which they specified on every trial whether Mark was the source, Geoff was the source, or a behavior was brand new.

## 2.2. Results and discussion

Unless specified otherwise by a  $p$  value, statistical significance does not exceed chance of a Type I error by the conventional 5% throughout this article. In order to distill the amount of data, conditionalized source monitoring (CSIM) scores are reported in Table 1. These scores represent correct source attributions for a given type of item and its source. For example, in the after-study condition negatively valenced behaviors originally spoken by the negatively depicted source (Mark) were correctly attributed to him at a rate of .68, incorrectly attributed to the positive source at .17, and called brand new .15 (N.B.,  $.68 + .17 + .15 = 1.0$ ). The CSIM score is calculated as the proportion of items declared old that were correctly attributed to their source, in this case, calculated as  $.68 / (.68 + .17) = .80$  in the first row of Table 1. None of the effects reported here would have changed if we had not used this data distillation technique, but rather, presented all 36 means from average raw claims.<sup>2</sup> Note that the first two columns of Table 1 represent correct claims for positive items and the second two columns represent correct claims about negatively valenced items. For ease of comprehension, we have labeled the subcolumn headings within each as negative (Mark) and positive (Geoff) to denote the source with whom the behavior was originally associated. In this way, the first column represents

---

<sup>2</sup> Currently, there is no universal agreement on how source-monitoring data should be reported (see Qin, Raye, Johnson, & Mitchell, 2001). Some argue that multinomial modeling or ACSIM scores (which merely average the CSIM scores we have reported) are to be preferred. Although we acknowledge this debate, we feel strongly that the effects reported in this article would be virtually identical no matter what data analytic technique was applied. Consequently, we have chosen the method that is most tractable to the widest possible audience. However, the full complement of data has been reported in the Appendix for those who wish to analyze the outcomes of our experiments with different data analytic procedures.

Table 1  
 Conditionalized source monitoring (CSIM) scores representing correct source attributions by type of item and source in Experiments 1–4

Experiment and condition	Item type and source			
	Positive stimuli		Negative stimuli	
	Negative	Positive	Negative	Positive
Experiment 1				
After-study	.64 (.04)	.71 (.04)	.80 (.03)	.61 (.04)
Before-study	.86 (.02)	.77 (.03)	.81 (.03)	.88 (.02)
Experiment 2				
After-study	.71 (.02)	.84 (.03)	.80 (.03)	.64 (.02)
Before-study	.86 (.03)	.84 (.03)	.82 (.03)	.87 (.02)
Experiment 3				
Full attention	.71 (.03)	.85 (.03)	.84 (.02)	.72 (.04)
Divided attention	.60 (.05)	.83 (.03)	.81 (.03)	.63 (.05)
Experiment 4	.34 (.05)	.71 (.05)	.68 (.05)	.48 (.06)

*Note.* Standard errors are in parentheses.

positive items correctly attributed to the male who was depicted negatively and the second column represents positive items correctly attributed to the male depicted positively.

We conducted a 2 (condition: after-study versus before-study) by 2 (item valence: positive versus negative) by 2 (source: Mark/negative versus Geoff/positive) mixed-model analysis of variance (ANOVA) with repeated measures on the last two factors. Of most relevance, the three-way interaction was statistically significant,  $F(1, 61) = 15.22$ . In the after-study condition, positive behaviors were correctly attributed to the positive source numerically more often than they were correctly attributed to the negative source, although this contrast failed to reach statistical significance,  $t(30) = 1.16$ , n.s. The reverse was true with negative behaviors where source accuracy was statistically higher for the negative source as compared with the positive one,  $t(30) = 4.17$ . These results suggest that participants relied on the person descriptions in determining the origin of the behaviors because biased attributions were obtained at test. By contrast, exactly opposite results were obtained in the before-study condition where the person descriptions were learned prior to studying the behaviors. In this case, positive behaviors were more often correctly attributed to the negatively depicted source and negative behaviors were more often correctly attributed to the positive source, smaller of the two  $t(31)$ 's = 2.28. We believe that having the schema

available during encoding is likely to have caused schema-inconsistent behaviors to attract more attention; and consequently, they were better remembered (cf., Graesser & Nakamura, 1982). Therefore, the bias that was present to attribute a behavior to a source based on its consistency with the valenced content of the candidate memory was entirely absent in the before-study condition. In fact, having the schema available vastly improved source memory in general,  $F(1, 61) = 25.34$ .

The bias and the lack of bias in the after- and before-study conditions, respectively, was also observed in the false alarm rate to new behaviors that were never studied. In the after-study condition more positively valenced new items were attributed to the positively depicted source than to the negatively depicted source (.11 and .03, respectively),  $t(30) = 2.75$ . The same bias was observed with brand new negative behaviors being attributed to the negative source as compared with the positive one (.13 and .03, respectively),  $t(30) = 2.66$ . In the before-study condition, the misattributions of new positive behaviors were basically equivalent to each of the two sources (.04 versus .05), as were the negative new behaviors (.05 versus .03), both  $t(31)$ 's  $< 1$ . Neither of these simple contrasts was statistically significant. That the valenced content of the behaviors affected performance on new items only in the after-study condition is very strong evidence for bias in that condition. Moreover, that bias in new items is consistent with other source-monitoring biases such as the it-had-to-be-you bias in which new items are more often attributed to an external source as opposed to one's self because self-generated information is usually much stronger in memory (e.g., Johnson & Raye, 1981; Johnson, Raye, Foley, & Foley, 1981; Marsh & Bower, 1993).

The results of this experiment neatly parallel those found by Hicks and Cockman (in press) who similarly manipulated whether a schema was present during encoding or available only at test. When available during learning, participants in the present experiment noted schema-inconsistent and schema-consistent behaviors which dramatically improved their source monitoring. When that information was provided later, the valenced content of the behaviors biased participants to claim that the behavior was associated with the person of the same positive or negative valence. Therefore, from a theoretical standpoint, the halo and devil influences on source monitoring appear to operate just like other sorts of schemas insofar as they will bias attributions in the absence of clear memory for the original source. However, one important difference between the current results and those found by Hicks and Cockman is that source performance for "schema inconsistent" items in the before-study condition was actually better than "schema consistent" items in the after-study condition. Hicks and Cockman only found that performance was better overall (i.e., in all conditions) when descriptions were provided before study. In this sense, the affective manipulation used here appears to be somewhat stronger than semantic schema information. We will return to this issue in discussing the results of Experiment 2.

### 3. Experiment 2

The halo effect on source judgments found in Experiment 1 may be the consequence of having used behaviors as stimuli (as did Sherman & Bessenoff, 1999).

Behaviors are usually consistent with personality types, and thus, the biases in source attributions may be a natural extension of a preexisting belief that “good” and “bad” people are associated with certain types of behaviors. In this next experiment, valenced events were used that could have befallen any individual. Some of these were negative such as getting a flat tire, and others were positive like winning a scratch off lottery ticket. On the one hand, participants may rely less on the schematic information with these stimuli because the events could have happened to anybody. On the other hand, in the absence of clear memory there may be no better basis on which to make source judgments in which case the effects found with behaviors in Experiment 1 may extend to common, albeit valenced events as well. For continuity, we tested again the after- and before-study conditions to ascertain whether the halo effect behaves like other schematic information.

### 3.1. Method

#### 3.1.1. Participants

Fifty-three undergraduates from the University of Georgia volunteered in exchange for partial credit toward a course requirement. Each participant was tested individually in sessions that lasted approximately 25 min. As described shortly, 26 participants were assigned haphazardly to the after-study condition and 27 were tested in the before-study condition.

#### 3.1.2. Materials and procedure

This experiment was conducted identically to Experiment 1 in all of its essential properties with the exception of the types of stimuli being learned and tested. A pool of 72 common (but valenced) events were developed. Half of these (36) were positive (e.g., finding money on the ground) and the other half were negative (e.g., accidentally tearing one’s favorite shirt). These were presented for learning and testing in an identical fashion as reported for Experiment 1. To briefly recapitulate, 12 positive and 12 negative events were randomly drawn anew from the pool of events to be presented as having happened to each of the two sources. This selection procedure left a random 12 each of the positive and negative events to serve as distractors on the source test. Participants in the before-study condition were asked to read the two short person descriptions prior to encoding whereas those assigned to the after-study condition read them after encoding the events. All participants were given a distractor activity for 5 min and then the source test was administered.

### 3.2. Results and discussion

The average CSIM scores representing accurate source monitoring by item types and sources are summarized in Table 1 using the same conventions as in Experiment 1. As before, the omnibus 2 (condition: before-study versus after-study) by 2 (item type: positive versus negative) by 2 (source: negative/Mark versus positive/Geoff) mixed-model ANOVA was tested. The three-way interaction was statistically significant,  $F(1, 51) = 14.54$ , but the nature of that interaction was quite different from the previous experiment. Replicating the numerical effect found in the after-study

condition of Experiment 1, more positive events were correctly attributed to the positive source, and fewer were correctly attributed to the negative source,  $t(25) = 3.22$ . Similarly, more negative items were correctly attributed to the negative source, and fewer were attributed to the positive source,  $t(25) = 3.47$ . In other words, participants were willing to use the valenced content of memory traces to assign items to sources even though these were everyday events that could have happened to anyone (and had occurred in equal numbers to each source). Therefore, similar to investigations of schematic influences on source monitoring, providing the valenced information about the sources just before the test biased source attributions.

Recall that providing the person descriptions prior to studying in Experiment 1 resulted in better source memory for “schema-inconsistent” behaviors (i.e., items which had the opposite valence to the source’s valence). The same was not generally true in this experiment. There was a 2% numerical difference for positive events associated with the negative source to be better remembered as compared to the positive source, but this difference was not statistically significant,  $t(26) < 1$ . Likewise, there was a 5% advantage for the positive source’s negative events as compared to the negative source’s negative events, but this contrast was also nonsignificant,  $t(26) = 1.61$ ,  $p > .10$ . In other words, with valenced events as stimuli, there was an extremely attenuated learning advantage (if any at all) for events that were inconsistent with a source’s depiction. One reason for the difference between Experiments 1 and 2 may be the degree to which participants view behaviors versus events as schema relevant. Behaviors (Experiment 1) are probably viewed as directly related to personality-driven behavior (i.e., they are motivated). In this respect, the source/person descriptions may have appeared more relevant with behaviors in the last experiment than they did in this experiment with events which are relatively unmotivated occurrences.

Recall that there was a valenced-based bias to attribute new items in the after-study condition in Experiment 1, but no bias for the same items in the before-study condition. Of interest, there were many fewer false alarms to new items in this experiment than in the previous one. When the person descriptions were provided before study, nominally more positive new items were attributed to the positive source (.03 and .02), and the same was true of the negative events being attributed to the negative source (.03 and .01), but neither contrast was statistically significant, both  $t(25)$ ’s  $< 1$ . When the descriptions were provided after study, similar attributions were made. More positive events were attributed to the positive source (.04 versus .02),  $t(25) = 2.06$ , and more negative events were attributed to the negative source than the positive one (.06 versus .03),  $t(25) = 2.09$ . Because only the last two of these four simple contrasts was statistically significant, very cautious interpretation of these data is warranted. The fewer false attributions of new items to sources in this experiment as compared with Experiment 1 suggests that the biases may indeed be stronger for behaviors than everyday events. This observation is consistent with our motivations for changing the stimuli in this experiment. Nonetheless, the overarching point demonstrated by this experiment is that attributions can be biased by matching the valence of the content of an everyday event to the valence of the candidate sources. However, the benefit of “schema-inconsistent” learning was absent in the before-study condition suggesting that emotional valence does not behave exactly like preexisting semantic knowledge.

#### 4. Experiment 3

Previous articles examining the influence of schematic information on source monitoring have provided information of the ilk supplied in the person descriptions used in Experiments 1 and 2. That sort of manipulation is consistent with the everyday cognitive processing that might be associated with learning who is a doctor versus a lawyer prior to going to a social event versus only finding out that information after the engagement is over (e.g., Hicks & Cockman, *in press*). No article has yet examined whether participants can deduce the schema-like information from the learning materials without being explicitly told anything about the two sources.<sup>3</sup> In this next experiment descriptions of the sources were eliminated and only the statements from the two sources were presented. Consequently, we combined all of the stimuli from the previous two experiments and associated 75% of the positive stimuli with one source and 75% of the negative stimuli with the other source. The remaining 25% of the statements from each valence was associated with the other source. If participants can learn these associations, and some research suggests that people are adept at learning very small implicit correlations (e.g., Heit, 1992), then source-monitoring biases might be observed at test. This prediction is bolstered by the fact that participants who learn to probability match in tests of recognition will produce a concomitant response bias (Ratcliff, Sheu, & Gronlund, 1992; see Erdfeelder & Bredenkamp, 1998, for a similar result on a source test).

Nonetheless, a self-learning manipulation could lead to performance similar to the before-study conditions of Experiments 1 and 2. After all, the information about the two sources is being presented during learning. From this perspective, perhaps no bias at all will be observed in this experiment. Recall that Sherman and Bessenoff (1999) found that younger adults only relied on priest versus skin-head schemas under divided attention at test. Similar to the present experiment, the manner in which their participants learned the materials was most consistent with the provision of the schema prior to learning. Consequently, we decided to introduce a manipulation of divided attention in this experiment. If a bias is not found under full attention at test in this experiment, then based on Sherman and Bessenoff's results we reasoned that one could emerge under conditions of divided attention. Thus, if we find biases only under divided attention during testing, this would suggest that Sherman and Bessenoff only found biased schema-based source attributions from the divided attention manipulation at test because their participants had the skin-head and priest schemas during learning. By contrast, to the extent that the bias is found under full attention at test in this experiment, that bias may be exacerbated by conditions of cognitive distraction during the test.

---

<sup>3</sup> Technically, Mather et al. (1999) attempted a similar learning manipulation with "secondary" schemas that were subordinate and orthogonal to their main schemas of republican versus democrat. However, they found no biases as a result of the secondary schema. Thus, the present experiment is a more direct test of the hypothesis concerning on-line learning.

#### 4.1. Method

##### 4.1.1. Participants

Fifty-nine University of Georgia undergraduates volunteered in exchange for partial credit toward a course research requirement. Each participant was tested individually in sessions that lasted approximately 30 min. Thirty volunteers were tested under conditions of full attention at test and 29 volunteers were tested under conditions of divided attention as described shortly.

##### 4.1.2. Materials and procedure

The procedural details of this experiment followed very closely those of the previous experiments. The behaviors used in Experiment 1 and the events used in Experiment 2 were folded into two large groups of 72 positively and 72 negatively valenced statements. Of the 72 of each type, 36 positive items and 12 negative items were randomly drawn and associated with the positive source (i.e., Geoff). Likewise, 36 negative items and 12 positive items were randomly drawn and associated with the negative source (i.e., Mark). Thus, each source had 48 statements associated with him but a preponderance (75%) were either from one valenced class of items or the other. The 48 remaining items were used as distractors on the source test. Obviously, half of them were positive and half were negative (i.e., 24 each). All randomization procedures occurred anew for each participant tested. Consistent with the goals of the experiment, no person descriptions were used.

The source test was the standard, three-alternative, forced-choice test used previously. However, it was conducted under conditions of full attention or divided attention. In the full attention condition, participants completed a puzzle distractor activity for 5 min and then completed the 144 item source test. In the divided attention condition, there was no distractor task. Rather, a baseline measurement was taken of the random number generation (RNG) task that was used to divide attention at test. The RNG task purportedly taps central executive resources (e.g., Hicks & Marsh, 2000) and is considered to be a rather difficult task. Participants had to generate a random number between 1 and 10 (inclusive) every time a metronome beat was heard played from a tape recorder. Beats of the metronome occurred every 1.5 s. The experimenter recorded performance on specially prepared sheets used to identify both missed beats and the numbers that were generated successfully. A minimum of 100 random numbers is needed to calculate a stable measure of RNG performance (Evans, 1978).

After the study phase in the divided attention condition, the RNG task was explained. Participants learned that in order to be random they should use the numbers 1–10 about equally often in the long run and avoid stereotypic sequences such as 1-2-3, 2-4-6, 9-8-7, etc. The baseline measurement was taken free of concurrent load and was explained to participants as a practice session. After obtaining the baseline measurement, the instructions for the source test were described and participants were informed that they were being asked to perform the RNG task and the source test concurrently. The RNG task was started approximately 20 s before the source test began. The experimenter started recording RNG performance about 1 min into the source test. After the experimenter obtained more than 100 stable digits that were

error free (not missed beats or errors out of the range of 1–10), she stopped recording RNG but the participant was unaware of this fact and continued the source test under the concurrent load until all 144 source judgments had been made.

#### 4.2. Results and discussion

The data are summarized in Table 1. A 2 (condition: full versus divided attention) by 2 (source: positive versus negative) by 2 (item type: positive versus negative valence) mixed-model ANOVA was conducted on the CSIM scores that reflect correct source attributions. The two-way interaction between item type and source was statistically significant,  $F(1, 51) = 37.70$ . Positive items were correctly attributed to the positive source at equal rates under both full and divided attention,  $t(51) < 1$ , n.s. The same was true of negative items being correctly attributed to the negative source,  $t(51) < 1$ , n.s. Thus, divided attention did not affect correct attributions of items whose valenced content was consistent with the source of the same valence. In the full attention condition, fewer positive items were correctly attributed to the negative source than the positive source, and fewer negative items were correctly attributed to the positive source than the negative source, smaller of the two  $t(29)$ 's = 2.76. Thus, based on the data from the full attention condition, participants were able to learn that the sources were associated with one or the other type of valence. In the divided attention condition, these differences in biased attributions were almost twice as large. Fewer positive items were correctly attributed to the negative source and fewer negative items were correctly attributed to the positive source under divided attention as compared to full attention, smaller of the two  $t(22)$ 's = 2.48.

Overall, correct source attributions were lower under divided attention than when full attention was available,  $F(1, 51) = 4.32$ , but as just mentioned, this result was mainly due to performance on the schema-inconsistent items. The performance decrement in the divided attention condition could be a consequence of sacrificing accuracy on the source task in order to perform well on the RNG task. To rule out a dual-task tradeoff interpretation of the results, we analyzed performance on the divided attention task itself (see Hicks & Marsh, 2000). Measures of RNG range from 0 to 1.0, with lower scores reflecting more random (i.e., better) performance. For reference, Evans (1978) found that RNG equaled approximately .30 when generating random numbers free of any other concurrent task. During baseline, participants had an average RNG of .29. When the RNG task was performed concurrently with the source task, RNG became less random and rose to .34. The difference between these scores was statistically significant,  $t(28) = 5.03$ . Therefore, randomness of generation declined as did source-monitoring performance. These results demonstrate that there was no dual-task tradeoffs because participants performed more poorly on both tasks when they were performed concurrently.

At the outset, we assumed that divided attention would make participants rely more on the self-learned schema (Sherman & Bessenoff, 1999), and we assumed further that divided attention would reduce recollective abilities (Jacoby, 1991; Jacoby & Dywan, 1990). Under these assumptions schema-consistent items should have been attributed correctly more often under divided attention than under full attention conditions (due to greater schema reliance), but this did not occur. Rather,

schema-inconsistent items were misattributed more often under divided attention because participants were relying on the schema. Therefore, divided attention may *only* be affecting schema-inconsistent items by reducing the recollective details that are needed to identify them as being schema inconsistent (e.g., associated cognitive operations). By this account, participants cannot recall the “tag” or other source information that the item was associated with the inconsistent source because attention was divided (e.g., Graesser et al., 1979). The overarching point is that participants did learn the affective correlations associated with each of the two sources, and they used this information both under full and divided attention conditions.

## 5. Experiment 4

In this last experiment, we briefly examine one reason for why the biases could have arisen in the after-study conditions of Experiments 1 and 2 (and in Experiment 3). In all three experiments, the behaviors and the events were depicted as being spoken by each male in the first person. Thus, the valence of the behavior or event reflected directly on the speaker. In this next experiment, the behavior statements used in Experiment 1 were rewritten in the third person so that the valence of the behavior reflected on somebody else other than the speakers (e.g., The old man gave the homeless person a quarter; The store clerk cut to the front of line). In this fashion, the speakers were explicitly providing valenced behavior information about other people and not themselves. On the one hand, behaviors of third parties may not be susceptible to the biases previously observed for valenced information about the two sources in an after-study condition. This outcome would be consistent with a person schema being irrelevant to what that person observes. On the other hand, the same test biases may be present in an after-study condition even though the behaviors being described are of a third party. If this outcome occurs, then it might arise because participants have some naive theory that negative people are somehow in a greater position to observe negative behaviors (and vice versa for positive people and positive behaviors).

### 5.1. Method

#### 5.1.1. Participants

Twenty undergraduates from the University of Georgia volunteered in exchange for partial credit toward a course research requirement. Each participant was tested individually in sessions that lasted approximately 30 min.

#### 5.1.2. Materials and procedure

The procedure for this experiment was identical in all respects to the after-study condition of Experiment 1 (i.e., full attention conditions). The only difference was that all of the behavioral statements were rewritten from the first person to the third person. We used a variety of occupations and innocuous references to third parties. To briefly recapitulate, participants learned 24 statements from each of two sources. After learning, they performed a distractor task for 5 min and then they read the two brief descriptions of the male sources. Finally, they took the source-monitoring test.

## 5.2. Results and discussion

The data are summarized at the bottom of Table 1. A 2 (source) by 2 (item type) ANOVA model was tested that revealed a significant interaction,  $F(1, 19) = 10.93$ . Positive behaviors of third parties were more often correctly attributed to the positively depicted speaker than they were correctly attributed to the negative source,  $t(19) = 4.52$ . Similarly, negative behaviors were more often correctly attributed to the negative source than to the positive source,  $t(19) = 2.02$ ,  $p = .05$ . Thus, the bias to attribute behaviors of third parties was identical to the bias observed when the behaviors were presented in the first person in Experiment 1. We do note that source performance was worse than in Experiment 1 to which the data are most comparable, and especially so for behaviors that were inconsistent with the valence of the speaker. That outcome suggests that rather than attenuate the valenced-based biases, they may have been slightly exacerbated by third party descriptors. One reason for the poorer learning in this experiment could be because participants were reluctant to relate the content of the third person statements to the speaker, as compared to the previous experiments. With poorer learning, the test biases became exaggerated.

The biases just discussed were reflected in new items as well. Positive behaviors were nominally more often associated with the positive (.19) rather than the negative source (.10), but this effect failed to reach conventional significance,  $t(19) = 1.73$ ,  $p = .10$ . In addition, negative behaviors were more likely to be attributed to the negative (.20) rather than positive (.11) source,  $t(19) = 1.97$ ,  $p = .06$ . As mentioned earlier, one way to account for these biases is that people may assume that a positive source has more opportunity to observe positive events whereas a negative source has the opposite opportunities. Although other explanations might work, the overarching point is that when people lack source specifying details in a memory trace they do appear to rely on the match of valenced information in the candidate memory and the potential sources being considered.

## 6. Familiarity versus recollection in Experiments 1–4

Thus far we have depicted the halo and devil influences on source monitoring as a bias effect. In doing so, we have been agnostic as to whether the bias is relatively automatic and therefore better classified as familiarity-based, or whether it is more consciously driven by recollective processes. Exploring this issue may be informative concerning what information is used, and how it is used, in making source attributions. After all, a debate currently exists whether source-monitoring processes can take advantage of partial memorial information of the sort giving rise to familiarity or whether these processes require retrieving recollective details. Many have argued that source monitoring is almost entirely recollection-based (e.g., Guttentag & Carroll, 1997; Perfect, Mayes, Downes, & Van Eijk, 1996; Yonelinas, 1999). On the other hand, we and others have argued that it can be familiarity-based (Banks, 2000; Dodson, Holland, & Shimamura, 1998; Gruppuso, Lindsay, & Kelley, 1997; Hicks, Marsh, & Ritschel, 2002). The results from the present study could be important insofar as they may help to resolve whether source-monitoring processes require

recollective details, or in some cases, can be based on “less perfect” information such as familiarity.<sup>4</sup>

Framed in this way, the fundamental question is whether the schematic influences of valence are best characterized as arising from memory versus bias. Although other analytic procedures are available (e.g., multinomial modeling), signal detection analyses of the data from Experiments 1–4 are one way to shed some light on this issue. A measure of discrimination such as  $d'$  would indicate participant's ability to distinguish the items belonging to each source. As such, effects of item valence that result in differences in  $d'$  can be classified as memory-based (or strength-based). For example, if the halo (or schematic) effects helped participants either to encode or to recall source information, then that influence would be recollective in nature. By contrast, effects on the criterion  $C$  are independent of discrimination  $d'$  and can be classified as criterial or familiarity-based. Based on the data in Table 1, we expected that providing person descriptions before study would translate into higher measures of  $d'$  as compared with providing them afterward. Likewise, we also expected that  $d'$  would be higher under conditions of full as compared with divided attention at test. The critical analysis determines whether  $d'$  measures are different for these conditions as a function of the valence of the items. In other words, if valence affects source monitoring through its influence on memory (i.e., strength) then an interaction between condition and item type should be found for  $d'$  in Experiments 1 and 2. By the same token, if the influence is criterial or familiarity-based, then the interaction between condition and item valence would be found for criterion  $C$  (and not discrimination  $d'$ ).

We conducted the signal detection analyses for Experiments 1–4 to assess these possibilities. In the computations, correct source attributions were considered “hits” and incorrect source attributions of old items were considered “false alarms.” Because it mattered little whether attributions to the positive (Geoff) versus the negative (Mark) source was analyzed, we chose the former. Of course, our analyses assume Gaussian distributions of equal variance despite the fact that lure distributions have only about 80% of the variance found in distributions of old items (Ratcliff et al., 1992). Table 2 summarizes the  $d'$  and  $C$  measures for each experiment and condition (i.e., rows) as a function of item valence (i.e., pairs of columns). For Experiment 1, a 2 (condition: after-study versus before-study) by 2 (item type: positive versus negative) mixed-model ANOVA was conducted on the  $d'$  measures. There was no significant interaction,  $F(1, 61) = 1.08$ . The absence of this interaction indicates that item valence did not affect discrimination differently in the after- versus before-study conditions. Rather, there was a main effect of condition indicating that having person descriptions before learning improved memory, and memorability of negatively valenced information was better, both  $F$ 's (1, 61) > 10.90. The identical ANOVA model tested on criterion  $C$  revealed only a significant interaction between condition and item valence,  $F(1, 61) = 18.79$ . Note that smaller values of  $C$  represent more liberally-placed criteria, and because the  $C$  values were calculated based on the positive source, there is a strong liberal bias to attribute positive items to the positive source in the after-study condition but simultaneously a strong reticence (conser-

---

<sup>4</sup> We appreciate the theoretical issues in this section of the article, and the suggested analyses that were conducted to address them, being raised in the review process.

Table 2  
Signal detection measures of discriminability and bias for each condition and item valence in Experiments 1–4

Experiment and condition	Item type and signal detection measure			
	Positive stimuli		Negative stimuli	
	<i>d'</i>	<i>C</i>	<i>d'</i>	<i>C</i>
Experiment 1				
After-study	.77	.25	.96	.44
Before-study	1.53	.42	1.91	.06
Experiment 2				
After-study	1.27	.06	1.07	.34
Before-study	1.87	.22	1.92	.06
Experiment 3				
Full attention	1.12	.19	1.60	.46
Divided attention	.79	.10	.97	.53
Experiment 4	.12	–.05	.40	.52

vative bias) existed to attribute negative items to the positive source. Thus, these signal detection measures strongly argue that participants are responding based on criterial considerations (i.e., familiarity).

The analogous results from Experiment 2 tell the identical story. The critical analysis of *d'* found no significant interaction between condition (before- versus after-study) and valence of the candidate items,  $F(1, 51) < 1.0$ . Although negative items were not remembered better than positive items in this experiment as they were in Experiment 1, the reader will recall that everyday events rather than behaviors were used. The absence of the interaction in *d'* is complemented by the statistically significant interaction in criterion *C*,  $F(1, 51) = 7.97$ . Neither main effect was significant for the criterion *C*, both  $F$ 's  $(1, 51) < 1.64$ . As such, the signal detection analysis of this experiment replicated in all of its essential properties the analysis conducted for Experiment 1.

Experiment 3 manipulated full versus divided attention, but in both conditions participants learned the correlation between the valenced content of items and the sources without experimenter-provided descriptions. Even a quick scan of Table 2 reveals that the signal detection measures in both conditions are entirely consistent with the after-study conditions of both Experiments 1 and 2. Therefore, significant main effects of condition and item valence are not surprising for discrimination *d'*, smaller  $F(1, 57) = 6.20$ . Only the effect of item type was significant for criterion *C*,  $F(1, 57) = 22.32$ . For this experiment, an interaction between item valence and condition would have indicated a stronger bias (*C*) under conditions of divided attention as compared with full attention, but obviously this did not occur. Because the signal detection analysis of Experiment 4 was entirely consistent with Experiments 1–3 in terms of interpretation and statistical significance, we leave it for the reader's visual inspection at the bottom of Table 2.

In summary, the overarching points that the reader should derive from these analyses are that (1) signal detection analysis of source data can differentiate between empirical effects due to memory ( $d'$ ) versus bias ( $C$ ) as effectively as, say, multinomial modeling (see Banks, 2000), (2) valence behaved like schematic information to change memory when source descriptions were provided before learning or when attention was divided at test, and most important, (3) source-monitoring decisions were found, in part, to be criterion-based which is most consistent with influences from familiarity or what has otherwise been labeled partial knowledge (e.g., Dodson et al., 1998; Hicks et al., 2002). Regarding this last point, the signal detection data strongly suggests that source-monitoring decision processes are not entirely recollection-based, and theories of source monitoring that strongly endorse this stance will need to be updated accordingly (e.g., Yonelinas, 1999).

## 7. General discussion

One purpose of this study was to ascertain whether source attributions would be influenced by the halo and devil effects. The answer to this question depends on when information about the candidate sources is learned. When this information is presented after encoding (but before testing), then halo-induced test biases are strongly observed in raw attributions (see Appendix), in the CSIM scores reflecting correct performance (see Table 1), and in criterion shifts analyzed with signal detection measures (see Table 2). However, similar to the effect that Hicks and Cockman (in press) found with standard semantic schemas (i.e., occupations), knowing valenced information about the candidate sources prior to learning improved source performance and more so for schema-inconsistent information (cf., Graesser et al., 1979). That effect was most strongly observed in Experiment 1 with behaviors, but also evidenced in Experiment 2 with valenced events. These outcomes suggest that valence can act similarly to schematic information about sources such as doctors and lawyers. However, whereas schema effects are based on semantic consistency (and inconsistency), the effects found in the present study are based on affective information about the positive and negative valence of candidate sources. Both effects are similar insofar as participants appear to be matching information in a candidate memory trace to discriminating information about candidate sources. However, in one case it is matching semantic information and in the present case it is matching affective information.

Together, the present results and other schema-based findings that have been reported previously may suggest a more general principle about source monitoring. According to the source-monitoring framework, memory traces are comprised of qualitative characteristics such as spatial and temporal details, emotional information, cognitive operations, perceptual details, semantic information, and so forth. We assert that when candidate sources can be differentiated on the basis of *any* of these qualitative characteristics, memories whose *content* shares features of this differentiation will be attributed based on the consistency of the content of the memory trace and the characteristic(s) supporting the differentiation. For example, suppose that one has an older acquaintance with a predilection to reminisce about past events, as well as a younger acquaintance who likes to engage in conversation

about current events. A statement about the Vietnam war might be equally likely to have been uttered by either source, but because the content of the memory has temporal content that is backward looking there may be a bias to attribute such a statement to the former acquaintance rather than the latter. Such a bias would not be based on semantic information, *per se*, but rather based on the referent age of the material in the candidate memory trace. Our point here is that what have appeared as schema-based effects previously in the literature (Bayen et al., 2000; Hicks & Cockman, *in press*; Mather et al., 1999) may just be a subset of a *more global* test bias to match any content-oriented information in a memory trace to any dimension that discriminates one source from another.

In the source-monitoring literature, contrasts within a modality have been studied (e.g., male versus female sources) and these contrasts generally yield worse source monitoring than contrasts between different qualitative characteristics (e.g., reality monitoring) due to the relatively high similarity between the sources (e.g., Lindsay, Johnson, & Kwon, 1991). The within-modality contrasts that have been studied previously share some similarity to the schematic sources that have been studied in the past and the valenced sources being studied here. The primary difference between the two types of studies is that schema-based (and the valenced-based) sources have candidate memories whose content overlaps with the distinguishing feature that separates one source from another. Previous studies using within-modality contrasts between sources have generally used study materials whose content is neutral with respect to the dimension that distinguishes one source from another. In that sense, those previous studies were able to avoid the test (perhaps guessing) biases that have been observed when source monitoring can be based on world knowledge, schematic information, and beliefs.

However, using neutral materials to avoid the test bias observed here is not necessarily a desirable quality in all studies of source monitoring. Neutral materials will yield studies of source monitoring that inform the inspection process for episodic details associated with a memory under “best case” scenarios when full attention can be devoted to source monitoring (e.g., Marsh, Landau, & Hicks, 1997). However, the inspection of such details appears to be augmented by a second process of attributing the origin when the content of memories maps on to the distinguishing feature(s) between candidate sources. In unpublished work from our laboratory we have found that seen versus heard and seen versus generated sources (using neutral materials) are not substantially affected by dividing attention with the identical RNG task used in Experiment 3. By contrast, we found significant performance decrements for valence-inconsistent items as a consequence of divided attention in Experiment 3. The contrast between these outcomes strongly suggests that examination of episodic qualitative characteristics is more resilient to cognitive distraction than source monitoring based on assessing the consistency of a memory’s content with the salient dimension distinguishing candidate sources.

This difference in resistance to manipulations of divided attention is likely to result from the type of decision criteria that are being used when inspecting qualitative characteristics versus more content-oriented general knowledge. The inspection of qualitative characteristics may involve fast, heuristic processing (Johnson, Hashtroudi, & Lindsay, 1993). In these cases, deciding whether one versus another

characteristic is more plentiful in a candidate memory may be relatively more automatic and therefore more immune from conditions of cognitive distraction. By contrast, sometimes source-monitoring decision criteria must be applied more effortfully and deliberately which is likely to require more attentional demands (see Chaiken, Lieberman, & Eagly, 1989). Under these cases, the mental agenda may engage systematic changes in decision criteria such as might arise from plausibility checks. In the present case, participants may be reflecting more deliberately on whether the valenced content of the memory is consistent or inconsistent with the valence of the candidate sources. This reflective process appears to be more susceptible to disruption from divided attention. If our assertion is correct that a test bias will arise whenever the dimension distinguishing between sources matches the content of memories, then this suggests that these cases of source monitoring are most likely to suffer when full attention is not available.

In summary, we have shown that the halo and devil effects influence source attributions. Those effects appear similar to semantic influences of schematic information, but the emotion associated with valenced information has its own unique characteristics as well. For example, inconsistent valence associated with everyday events at learning does not attract attention the way schema-inconsistent semantic information does (Experiment 2). In addition, valence associated with behaviors of third parties (and not the sources themselves) produces particularly strong test biases. Whether other qualitative characteristics that are confounded with the content of memories and the distinguishing attribute(s) of sources will show similar effects remains to be determined. Nonetheless, the results from this study suggest that the current working hypothesis is that they will, but perhaps with their own unique idiosyncrasies that are yet to be discovered.

### Acknowledgments

We thank Dana Schuler, Rachel Parrott, and Bonnie Woodson for their help in collecting the data. We also appreciate Jaime Durley's assistance with data analyses in earlier stages of the project.

### Appendix A

Experiment, condition, and claim	Item type and original source					
	Positive stimuli		Negative stimuli		New stimuli	
	Negative	Positive	Negative	Positive	Negative	Positive
Experiment 1						
After-study						
“Negative”	.49	.20	.68	.33	.13	.03
“Positive”	.27	.55	.17	.52	.03	.11
“New”	.24	.24	.15	.16	.84	.86

Before-study						
“Negative”	.70	.17	.71	.11	.05	.04
“Positive”	.11	.63	.15	.81	.03	.05
“New”	.19	.20	.14	.09	.91	.91
Experiment 2						
After-study						
“Negative”	.62	.13	.72	.33	.06	.02
“Positive”	.25	.70	.19	.57	.03	.04
“New”	.13	.17	.09	.10	.91	.94
Before-study						
“Negative”	.79	.14	.74	.10	.03	.03
“Positive”	.12	.75	.16	.81	.02	.02
“New”	.10	.11	.10	.09	.95	.95
Experiment 3						
Full attention						
“Negative”	.57	.10	.68	.21	.05	.02
“Positive”	.22	.64	.12	.62	.02	.04
“New”	.21	.27	.20	.16	.94	.94
Divided attention						
“Negative”	.46	.13	.65	.28	.11	.06
“Positive”	.30	.62	.16	.51	.07	.12
“New”	.25	.24	.20	.21	.82	.81
Experiment 4						
Third person						
“Negative”	.26	.23	.54	.43	.20	.10
“Positive”	.50	.54	.25	.38	.11	.19
“New”	.24	.23	.22	.18	.69	.71

*Note.* New stimuli do not have an original source, and thus, the labels negative and positive refer to their valence.

## References

- Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science*, *11*, 267–273.
- Bayen, U. J., Nakamura, G. V., Dupuis, S. E., & Yang, C. (2000). The use of schematic knowledge about sources in source monitoring. *Memory & Cognition*, *28*, 480–500.
- Chaiken, S., Lieberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212–252). New York: Guilford Press.
- Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1121–1136.
- Erdfelder, E., & Bredenkamp, J. (1998). Recognition of script-typical versus script-atypical information: Effects of cognitive elaboration. *Memory & Cognition*, *26*, 922–938.

- Evans, F. J. (1978). Monitoring attention deployment by random number generation: An index to measure subjective randomness. *Bulletin of the Psychonomic Society*, *12*, 35–38.
- Graesser, A. C., Gordon, S. E., & Sawyer, J. D. (1979). Recognition memory for typical and atypical actions in scripted activities: Tests of a script pointer + tag hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *18*, 319–332.
- Graesser, A. C., & Nakamura, G. V. (1982). The impact of a schema on comprehension and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 59–109). New York: Academic Press.
- Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: Defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 259–278.
- Guttentag, R. E., & Carroll, D. (1997). Recollection-based recognition: Word frequency effects. *Journal of Memory and Language*, *37*, 502–516.
- Heit, E. (1992). Categorization using chains of examples. *Cognitive Psychology*, *24*, 341–380.
- Hicks, J. L., & Cockman, D. W. (in press). The effect of general knowledge on source memory and decision processes. *Journal of Memory and Language*.
- Hicks, J. L., & Marsh, R. L. (2000). Toward specifying the attentional demands of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1483–1498.
- Hicks, J. L., Marsh, R. L., & Ritschel, L. (2002). The role of recollection and partial information in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 503–508.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28.
- Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67–85.
- Johnson, M. K., Raye, C. L., Foley, H. J., & Foley, M. A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology*, *94*, 37–64.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.
- Jacoby, L. L., & Dywan, J. (1990). Effects of aging on source monitoring: Differences in susceptibility to false fame. *Psychology & Aging*, *5*, 379–387.
- Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in source monitoring. *Journal of Experimental Child Psychology*, *52*, 297–318.
- Marsh, R. L., Cook, G. I., & Durley, J. R. (2002). Gender and orientation stereotypes bias source-monitoring attributions. Unpublished manuscript.
- Marsh, R. L., & Bower, G. H. (1993). Eliciting cryptomnesia: Unconscious plagiarism in a puzzle task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 673–688.
- Marsh, R. L., Landau, J. D., & Hicks, J. L. (1997). Contributions of inadequate source monitoring to unconscious plagiarism during idea generation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 886–897.
- Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, *16*, 437–458.
- Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*, 250–256.
- Perfect, T. J., Mayes, A. R., Downes, J. J., & Van Eijk, R. (1996). Does context discriminate recollection from familiarity in recognition memory? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *49*, 797–813.

- Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source ROCs are (typically) curvilinear: Comment on Yonelinas (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1110–1115.
- Ratcliff, R., Sheu, C., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, *99*, 518–535.
- Sherman, J. W., & Bessenoff, G. R. (1999). Stereotypes as source-monitoring cues: On the interaction between episodic and semantic memory. *Psychological Science*, *10*, 106–110.
- Thorndike, E. L. (1920). A constant error on psychological rating. *Journal of Applied Psychology*, *4*, 25–29.
- Yonelinas, A. (1999). The contribution of recollection and familiarity to recognition and source memory judgments: A formal dual-process model and an ROC analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1415–1434.