

## Source monitoring is not always enhanced for valenced material

GABRIEL I. COOK

Claremont McKenna College, Claremont, California

JASON L. HICKS

Louisiana State University, Baton Rouge, Louisiana

AND

RICHARD L. MARSH

University of Georgia, Athens, Georgia

Source monitoring for valenced materials has received very little attention from researchers interested in the residual effects that emotion can have on memory. The three previous studies that examined memory for valenced material found a source-monitoring enhancement effect. By contrast, we used two different combinations of sources and found a novel, consistent source-monitoring deficit for valenced words as compared with neutral ones. In addition, this memory deficit for contextual details did not consistently covary with item memory. We assert that it is possible to obtain an effect in which heightened attention toward valenced material reduces the binding of contextual details into memory.

Conventional wisdom assumes, and many laboratory studies have confirmed, that affectively valenced materials are usually better remembered than otherwise comparable neutral materials. This enhancement in memory is not stimulus bound, insofar as studies report similar effects for words, pictures, sentences, and narrated videos (reviews can be found in Buchanan & Adolphs, 2002; Hamann, 2001). There are many possible reasons for the memorial enhancement, including the engagement of biological systems such as the amygdala (e.g., LeDoux, 1996; Metcalfe & Jacobs, 1998), general and/or idiosyncratic associations with valenced materials (e.g., Kensinger & Corkin, 2003), or simple distinctiveness of the items in comparison with their neutral counterparts (e.g., Dewhurst & Parry, 2000). Still other theories argue for other mechanisms, such as differential rehearsal and/or valenced material capturing more attention (Cahill & McGaugh, 1995). None of these explanations are necessarily mutually exclusive. If one believes that estimates from the remember-know procedure can index recollection and familiarity, valenced items are often imbued with more recollective details, and this is true of both pictures (Ochsner, 2000) and words (Dewhurst & Parry, 2000; Kensinger & Corkin, 2003). Moreover, Kensinger and Corkin, as well as Ochsner, found that negative words elicited both increased recollection and familiarity.

The purpose of the present article was to investigate source memory for valenced words. Many researchers associate source-monitoring processes with recollective pro-

cesses, as in *remembering versus knowing* (e.g., Guttentag & Carroll, 1997; Perfect, Mayes, Downes, & Van Eijk, 1996; Yonelinas, 1999). Although we (and others) believe that source monitoring can be performed even in the presence of only partial, incomplete information (e.g., Dodson, Holland, & Shimamura, 1998; Hicks, Marsh, & Ritschel, 2002), the source-memory-as-recollection stance predicts that source monitoring will be better for valenced words than for neutral words, because valenced materials are associated with greater recollection. In fact, this is exactly what Doerksen and Shimamura (2001) found, using blue and yellow colored words. Kensinger and Corkin (2003) have recently replicated the better color memory, using both negative and taboo (curse) words rather than neutral words. Our study was motivated, in part, by the fact that D'Argembeau and Van der Linden (2004) had some difficulty replicating the basic color memory advantage ostensibly conferred on valenced items, as originally reported by Doerksen and Shimamura. Two of D'Argembeau and Van der Linden's four experiments did not replicate the enhancement effect, and only when they used four colors rather than two, under incidental learning procedures, was the source-monitoring enhancement found. However, Doerksen and Shimamura as well as Kensinger and Corkin used intentional learning procedures.<sup>1</sup> In unpublished experiments from our own laboratory, we have not been able to replicate the source-memory advantage using Doerksen and Shimamura's stimuli with incidental encoding. Although D'Argembeau and Van der

---

R. L. Marsh, rlmars@uga.edu

---

Linden extended the source-monitoring advantage to spatial location of words on a computer monitor, we feel that it would be fair to say that the evidence is far from strong that valenced words more generally increase source memory for a variety of contextual attributes.

The source-monitoring framework specifies a rich set of memorial and decision processes by which assessments are made concerning the original context surrounding learning (Johnson, Hashtroudi, & Lindsay, 1993). Many pieces of memorial information—among them perceptual details, cognitive operations from elaboration and organization, spatiotemporal information, semantics, and of course affective information, to name a few—can be used to determine in what context a piece of information was learned. The previous studies assessing the effect of valence on source memory have manipulated perceptual details within a single modality (i.e., color and spatial location). In everyday life, the characteristics associated with emotional information are likely to be somewhat richer and the contrast among competing candidate sources is likely to be more differentiated than is the case for decisions made within a single modality. For example, whether one receives a devastating piece of news by reading it or by hearing it is likely to be a very different judgment than the ones that have been assessed to date in this literature. For this reason, and for the explicit purpose of generalizing our knowledge about valence to new sources, we used a seen versus a heard source in the majority of experimental conditions in the present study.

In order to decide whether an item was read on a computer monitor or spoken by the experimenter, participants will have to weigh the relative perceptual details of each source for each item. They may do this in one of several ways (see Johnson et al., 1993; Johnson & Raye, 1981). One way is to assume a decision axis bounded on one end by strong and rich auditory details and on the other by strong and rich visual details. On the assumption that the item is old, participants may set a single criterion along this axis to make their decisions. Another way is to assume two orthogonal axes, one for visual details and the other for auditory details. In three-dimensional space, the criterion is a plane through, say, a bimodal distribution of evidence. When orthogonality is relaxed and other assumptions are made, one can derive a multivariate signal detection model (in two-dimensional space) for making source decisions (e.g., Banks, 2000). In all of these cases, participants will be assessing whether auditory characteristics in the memory trace outweigh visual ones (and vice versa). If valenced material attracts attention and more contextual details are bound into memory, this decision process should be more successful for valenced items than for neutral ones.

Other than null results, two basic alternative predictions can be made. One is discussed in this paragraph and the other in the next. The emotional enhancement effect is generally attributable to one of two classes of processes. With items that are arousing (running the gamut from calming to exciting or agitating), ample evidence exists that the enhancement effect is biological (Kensinger, 2004). Because we are holding arousal constant in the present study, the other set of mechanisms creating an emotional enhance-

ment effect are elaborations in memory. Consequently, if we replicate the emotional enhancement effect on source memory, it would likely be attributable to great elaboration and processing of valenced items.

By contrast, the mere fact that three previous studies have found effects that appear somewhat consistent with enhanced source memory on valenced items does not preordain that outcome. In what has been referred to (with little consistency) as *weapon focus*, the *trauma-memory argument*, and the *Easterbrook hypothesis* (Easterbrook, 1959), emotional events, while they are being experienced, can cause a narrowing of the focus of attention (see Kihlstrom, 1995; Read & Lindsay, 1997). As described below, this narrowing of attention can increase memory for central details but decrease memory for peripheral ones. However, Kihlstrom cautiously noted that the types of events that might cause such attentional narrowing are difficult, if not impossible, to bring into the laboratory. If such an effect were found here, memory for auditory and visual contextual details might be worse, because these aspects of the encoding experience might be less likely to be bound into the resulting memory trace. Along the lines of valenced items receiving more elaboration, if the central aspects of those items are what receive that extra processing, more peripheral details such as context information may receive less processing.

In favor of the source-memory-deficit outcome, Jurica and Shimamura (1999) found worse source memory (i.e., of the person speaking) when, at encoding, people were asked questions rather than listening to statements. Jurica and Shimamura also found a trade-off between item and source memory, such that when item memory was better (e.g., for the question asked), source memory was worse. Presumably, the cognitive effort in generating an answer to the question reduced binding of which person asked the question. Moreover, Johnson, Nolde, and De Leonardis (1996) found a similar effect with emotional statements. In a regression analysis that was subsidiary to their primary aims, they reported that item memory for more highly rated emotional items was better than for neutral statements, but that source memory was worse for the emotional statements. Consequently, these two studies suggest that a reduction in source memory for valenced materials is a theoretically viable alternative prediction to the account from the three published studies that have shown a source-memory enhancement for color and location of words.

By way of overview, we conducted four experiments using valenced and neutral words. The first experiment mixed positive, negative, and neutral items and tested participants' memories for whether items were seen or heard. To anticipate our results, we found worse source memory for negative items than for neutral words. Because Dewhurst and Parry (2000) found that memory for neutral material was affected by whether or not valenced material was intermixed with the studied materials, in Experiments 2 and 3 we compared source memory using a variety of list compositions. Three conditions were tested; in Experiment 2, a pure neutral list, a pure negative list, and a mixed negative and neutral list were compared between subjects. In Experiment 3, the comparable three conditions were

tested, except that positive and neutral materials were compared. Finally, for generality, we changed the type of valenced words we used and also tested two new sources (anagrams and frequency rating) to assess whether the materials or the procedure accounted for our finding of worse source memory with valenced material.

## EXPERIMENT 1

In Experiment 1, we assessed source monitoring for a word list containing equal numbers of positive, negative, and neutral items. Doerksen and Shimamura (2001) used such a list, but Kensinger and Corkin (2003) used only negative and neutral items. Doerksen and Shimamura reported no source-memory differences between positive and negative items, with both studies showing better source memory than with neutral items. All other things being equal, our use of seen versus heard sources could have produced the same pattern of results. But if attentional narrowing had occurred, or if elaborating on the central details of valenced materials had compromised encoding of contextual details, a decrement to source memory for valenced items might have been obtained.

### Method

**Participants.** Undergraduate students from the University of Georgia volunteered in exchange for partial credit toward a research appreciation requirement. Each of the 34 participants was tested individually in sessions that lasted approximately 25 min. Selection criteria for signing up for all experiments reported in this article required English to be the participant's native language, although we did not assess this during the experiment proper.

**Materials and Procedure.** The stimuli used in Experiments 1–3 were selected from the Affective Norms for English Words list (ANEW; Bradley & Lang, 1999). We selected a pool of 90 positive, 90 negative, and 90 neutral items, under the constraints that valence statistically differed across the three classes of items but arousal, word frequency, and number of letters did not. On a 9-point scale, average valence was 7.70, 2.31, and 5.18 for the positive, negative, and neutral items [ $F(2,267) = 798.78$ ], whereas average word frequency was 22.47, 22.34, and 23.58, respectively [ $F(2,267) < 1$ , n.s.]. The respective means for arousal (on a 9-point scale) were 5.71, 5.89, and 5.70 [ $F(2,267) = 1.48$ , n.s.]. Average word length was approximately six letters. Because different numbers of items from each class are used across the different experiments, the software (written in-house) randomly selected items from the relevant pool of 90 items in each experiment. In each experiment, 60 items were studied and 30 items were new (Marsh & Hicks, 1998, used the same procedure). In this experiment, 20 items were studied from each of the three classes of items; half were presented visually, and the other half were spoken by the experimenter. At test, 30 new items were randomly intermingled anew with the 60 studied items, with 10 new items coming from each of the positive, negative, and neutral classes. All randomization procedures were performed online for each participant tested, which allowed us to avoid any potential item selection effects at the level of conditions and experiments.

Participants were asked to study a list of words, but they were not informed prior to study that their source memory would be tested later. In addition, they were not informed that some items might be affectively valenced. For the seen source, we used Marsh and Hicks's (1998) procedure of presenting items visually in the center of the computer monitor for 3 sec each; the heard-source item was written by the participant's computer to the experimenter's computer, at which point it was spoken aloud by the experimenter. Participants

could not view the experimenter's monitor, and trial duration for heard stimuli was also 3 sec. Prior to each studied item, a short warning tone signaled the participant that a word was about to be presented. At the conclusion of the encoding phase, we administered a brief distractor task (solving puzzles) before the test instructions were presented. The test instructions indicated that participants should press one labeled key if the item was heard, a different labeled key if the item was seen, and a third key if the item was brand new. (The home keys and the space bar, respectively, were used for this purpose.) All instructions for encoding and test were read by the participant from the computer monitor, but were then verbally reiterated by the experimenter. No participant was allowed to proceed with any phase of the experiment unless the experimenter was convinced that he or she understood the task at hand.

### Results and Discussion

Across the experiments, we first report the inferred recognition hit rate, which is the proportion of studied items labeled *old* without regard to source accuracy. Also reported is the false alarm rate and a measure of corrected recognition, which is the difference between the hit and false alarm rates. Subsequently, we report source-monitoring performance as averaged conditionalized source identification measures (ACSIMs). For a given source—for example, heard—we computed the proportion of heard items correctly called *heard* from all heard items attributed to the study list (i.e., those labeled *seen* or *heard*). The ACSIM score for heard items was averaged with the comparable score for seen items. This measure of source memory is entirely adequate for the present purposes and is much less cumbersome for the reader than other approaches. However, for the sake of completeness, nine cell tables that show the full complement of attributions for Experiments 1–3 are provided in the Appendix.

Table 1 summarizes the recognition measures, and Table 2 summarizes the ACSIM scores for all four experiments. A repeated measures ANOVA found that the inferred recognition hit rate did not differ across the three classes of items [ $F(2,66) = 1.65$ , n.s.]. Thus, contrary to previous studies, there was no recognition advantage for valenced items. However, there was a large difference in the false alarm rate [ $F(2,66) = 13.95$ ], largely owing to more negative new items being called *old*. Consequently, when these were combined to obtain a measure of corrected recognition, valenced materials resulted in worse recognition memory [ $F(2,66) = 5.40$ ]. In comparison with the neutral items, this was true of the negative items [ $t(33) = 3.46$ ], and marginally true of the positive items [ $t(33) = 1.87$ ,  $p = .07$ ]. Rather than an emotional enhancement effect, these results show a pattern more consistent with Talmi and Moscovitch's (2004) argument that valenced materials can act like a category, at least in the false alarm rate.

In an analogous ANOVA, source-monitoring performance was different across the three classes of items [ $F(2,66) = 3.89$ ]. As is evident from the means (see Table 2), source memory was worse for negative items than for neutral items [ $t(33) = 2.77$ ], but source memory was equivalent for positive and neutral items [ $t(33) < 1$ , n.s.]. Consequently, we did not find a source-memory enhancement, as did Doerksen and Shimamura (2001) and Kensinger and Corkin (2003). Rather, we found a source-

**Table 1**  
**Average Inferred Recognition Hit Rates, False Alarm Rates, and Corrected Recognition for Experiments 1–4**

Experiment and Condition	Inferred Recognition			False Alarm Rate			Corrected Recognition		
	Positive	Negative	Neutral	Positive	Negative	Neutral	Positive	Negative	Neutral
Experiment 1	.75	.79	.77	.24	.34	.18	.52	.45	.58
Experiment 2									
Negative and neutral mixed		.85	.81		.37	.15		.49	.65
Pure negative		.82			.32			.50	
Pure neutral			.77			.18			.59
Experiment 3									
Positive and neutral mixed	.76		.71	.24		.13	.52		.59
Pure positive	.77			.20			.57		
Pure neutral			.78			.18			.60
Experiment 4									
Seen and heard	.76	.75	.70	.20	.21	.05	.56	.55	.65
Anagram and frequency	.78	.79	.82	.21	.18	.04	.57	.61	.78

Note—Empty cells indicate that items from that class of items were not tested in that particular experiment or condition.

monitoring deficit localized to the negatively valenced class of items.

As discussed earlier, the data for negative items are most consistent with a theory of narrowing of attention in which contextual details of the study episode are bound into memory more poorly for negative items (Christianson, 1992; Read & Lindsay, 1997). Alternatively, perhaps more elaborated encoding of the central details of concepts like *cancer*, or idiosyncratic thoughts about such concepts, causes less attention to be paid to contextual details of the experience. The effect bears some resemblance to Jurica and Shimamura's (1999) finding that attention drawn toward answering a question will also result in poorer binding of contextual information. Why the same result does not hold for color memory in previous studies is unclear. However, modality of presentation may not be in the focus of attention on every trial in the same way that the color of a word is always in the focus of attention. In that sense, modality may be a more peripheral piece of context information. The subsequent experiments explored whether this novel source-monitoring deficit is replicable.

## EXPERIMENT 2

The deficit to source memory for negatively valenced items was not expected from the three articles on valence

**Table 2**  
**Average Conditionalized Source Identification Measures (ACSIMs) for Experiments 1–4**

Experiment and Condition	ACSIM		
	Positive	Negative	Neutral
Experiment 1	.79	.74	.81
Experiment 2			
Negative and neutral mixed		.76	.83
Pure negative		.77	
Pure neutral			.81
Experiment 3			
Positive and neutral mixed	.76		.82
Pure positive	.76		
Pure neutral			.81
Experiment 4			
Seen and heard	.74	.70	.87
Anagram and frequency	.75	.75	.88

and source memory published to date. One issue that may be relevant is the composition of the study list. Dewhurst and Parry (2000) found some evidence that neutral items behaved differently when valenced items were on the study list. They argued that valenced items captured attention from the neutral items, causing the latter to be less well remembered. Although the inferred recognition results from Experiment 1 suggest otherwise, we have no evidence for what source memory would be in pure lists of neutral items rather than of valenced items, because no previous report has tested this aspect of performance. Because there is ample evidence that some variables such as word frequency behave differently in between-subjects versus within-subjects designs, we thought it would be important to assess this variable and to ascertain whether the source-monitoring deficit obtained with seen and heard sources could be replicated. Therefore, in Experiment 2, we tested a mixed list of neutral and negative items and compared that within-subjects effect to a between-subjects effect, in which a pure list of neutral versus a pure list of negative items was studied and tested.

## Method

**Participants.** Undergraduates from the University of Georgia volunteered in exchange for partial credit toward a research appreciation requirement. Each participant was tested individually in sessions that lasted approximately 25 min. Participants were quasirandomly assigned to the three experimental conditions. Sample sizes in the mixed and pure negative conditions were 29 and 30 participants, respectively, and 27 participants were tested in the pure neutral condition. An additional 34 participants were tested using low-arousal negative items as described in the Results and Discussion section.

**Materials and Procedure.** Except for list composition, the procedures were identical to those in Experiment 1. In the mixed condition, 30 negative and 30 neutral items were studied, half being seen and the other half being heard. An additional 15 new negative and 15 new neutral items were added to the test list. In the pure negative and pure neutral conditions, 60 items from the respective word pool were studied, with half from each of the seen and heard sources, and 30 new items from the correct word pool were added to the test as new items. Thus, the entire pool of 90 items was used in the pure negative and pure neutral conditions.

## Results and Discussion

The inferred recognition hit rates (see Table 1) in the mixed condition showed the traditional memorial ben-

efit to negatively valenced information as compared with neutral [ $t(28) = 2.22$ ], but when the extraordinarily high false alarm rate is factored in, corrected recognition for the negative items was far below the level for neutral items [ $t(28) = 4.36$ ]. The same was true of the between-subjects comparison in the pure lists of negative and neutral items for the hit rates [ $t(55) = 1.92, p = .06$ ], although the effect is just shy of conventional standards of significance. However, the corrected measure of recognition showed a memory deficit for the negative words relative to the neutral words [ $t(55) = 2.67$ ]. The deficit to source memory for negative items as compared with neutral items was found in the mixed-list condition [ $t(28) = 3.96$ ]. The same was true in the pure list of negative items compared with the pure neutral list [ $t(55) = 3.41$ ]. Therefore, the novel source-monitoring deficit for negative items found in Experiment 1 was replicated in this experiment.

As we mentioned earlier, arousal rather than valence cannot have affected these findings, because arousal was equated across the three stimulus classes. However, arousal and valence are based on subjective ratings, and they are correlated such that negative valence tends to evoke more arousal than positive valence. To be absolutely sure that arousal was not influencing our novel finding of a source-monitoring deficit for negative items, we tested an additional 34 people in the mixed-list condition with low-arousal negative items. Their mean arousal rating was 2.43, compared with 5.89 in Experiments 1 and 2. The inferred recognition hit rates for negative and neutral items were .77 and .72, respectively, thereby showing an item memory advantage for valenced over neutral items [ $t(33) = 2.27$ ]. But with the false alarms corrected, item memory was equivalent at .55 and .58, respectively [ $t(33) = 1.1, n.s.$ ]. Nonetheless, source memory was still lower for negative than for neutral items: .76 and .81, respectively [ $t(33) = 2.98$ ]. In three out of three comparisons, therefore, we found a source-monitoring deficit to negatively valenced words that did not replicate the source-memory enhancement previously reported in the literature. The data suggest that a narrowing of attention on negative items that reduces binding of contextual modality information is a viable theoretical mechanism that can affect subsequent source memory.

### EXPERIMENT 3

We designed this next experiment to parallel Experiment 2, but this time we compared any effect obtained with mixed lists of positive and neutral items to any effect obtained with pure lists of positive versus neutral items. Kensinger and Corkin (2003) did not use positive items, whereas Doerksen and Shimamura (2001) reported an equal source-memory enhancement for both positive and negative items. If positive items behave as negative items did in Experiment 2, then source memory should be worse for positive as compared with neutral words. By contrast, the results from Experiment 1 suggest that positive items may behave like neutral items, but that was in a test with all three classes of items. As Kensinger (2004) has recently claimed, positive items have received very little empirical scrutiny in this literature.

### Method

**Participants.** Undergraduates from the University of Georgia volunteered in order to receive credit for a research appreciation requirement. In the two conditions with pure positive and pure neutral words, 35 people were tested. In the condition testing the mixed list of positive and neutral items, 36 people were tested.

**Procedure.** The compositions of the study and test lists were identical to those reported for Experiment 2, with the exception that positively valenced items were substituted for the negative items used in that previous experiment. In all other procedural respects, this experiment was identical to Experiment 2.

### Results and Discussion

As in the previous experiments, the key results are summarized in Tables 1 and 2. In the mixed-list condition, an item memory advantage was observed for the positive items over the neutral items, because the inferred recognition hit rate was higher for positive items [ $t(35) = 2.86$ ]. However, the same was not true in the between-subjects comparison of pure lists [ $t(68) < 1.0$ ]. After correcting for the false alarm rates, the results were identical to those from Experiment 2. Corrected recognition was lower for positive items in the mixed condition [ $t(35) = 2.28$ ], but the comparable between-subjects outcome failed to materialize [ $t(68) < 1.0$ ]. As can be seen in the means, inferred recognition for neutral words was better in the pure-list condition than in the mixed-list condition [ $t(68) = 2.02$ ], but after correction, that difference disappeared. Once again, these results are reminiscent of Dewhurst and Parry's (2000) claim that valenced items studied alongside neutral items can change cognitive processing of the latter. In this case, mixed-list study conditions lowered the hit rate, presumably because the valenced items received more residual processing.

That processing appears to have come at a cost to source memory, because, as in Experiment 2, source monitoring for positive items was worse than for neutral items in the mixed-list condition [ $t(35) = 2.79$ ], and also worse in the pure-list comparison [ $t(68) = 2.45$ ]. We can now add two more comparisons to the list of evidence in favor of a source-monitoring deficit to valenced items. The only evidence not consistent with this finding comes from the positive items tested alongside both negative and neutral items in Experiment 1. One viable hypothesis is that positive items do not garner as much attention when presented with negative items, because the negative items themselves capture more attention. As a consequence, positive items behave like neutral items when tested alongside the other two classes of items. We will return to this point in the General Discussion section.

### EXPERIMENT 4

Against the backdrop of the three studies finding a source-monitoring enhancement, our consistent finding of a deficit to source memory for valenced material alerted us to the possibility that there was something peculiar about the stimuli that we chose. The extra comparison tested in Experiment 2 using low-arousal negative stimuli gave us some solace that the materials were not causing this effect. Nevertheless, in Experiment 4 we used a dramati-

cally different set of materials. Shaver, Schwartz, Kirson, and O'Connor (1987) performed both cluster analyses and multidimensional scaling analyses to arrive at a set of five key clusters of emotion words corresponding to five of the seven basic emotions: *love*, *joy*, *fear*, *sadness*, and *anger*. We used words relating to these concepts as our stimuli, acknowledging at the outset that we would be testing more negative items than positive. Because we had been testing only a combination of two external sources (seen and heard), we took this opportunity to test a combination of two internal sources as well. One source was solving the item as an anagram, and the other was rating how frequently the concept had been encountered in the previous 2 weeks (Marsh, Hicks, & Davis, 2002, used these two internal sources as well).

### Method

**Participants.** University of Georgia undergraduates volunteered in exchange for partial credit toward a research appreciation requirement. Each participant was tested individually in sessions that lasted approximately 30 min. Thirty-three participants were assigned quasirandomly to the seen–heard condition, and 35 were assigned to the anagram–frequency condition.

**Materials and Procedure.** The stimuli from the Shaver et al. (1987) study were reduced to a pool of items representing each emotion as follows (numbers in parentheses indicate pool size): love (16), joy (32), fear (28), sadness (36), and anger (16). To these we added a sixth, neutral class of items (32). A random half of these items from each class were studied, and the remainder were reserved as new items on the source test. Thus, unlike Experiments 1–3, and most of our previous source-monitoring work, 25% were studied from one source, 25% were studied from the other source, and 50% were new during the memory test. This brings the old–new percentages in line with Doerksen and Shimamura (2001), although in practice we have never found this slight change in list composition (from 33% for each source) to matter on a source-monitoring test.

The seen–heard manipulation of sources was carried out in the same manner as for Experiments 1–3. When the source was rated for frequency in the anagram–frequency manipulation, a query appeared under the studied word prompting participants to rate on a 1–7 Likert scale how frequently they had encountered the concept in the previous 2 weeks, using the anchors *not at all* (1) to *very frequently* (7). The anagram manipulation is one we have used on many previous occasions, and it required participants to interchange two letters to identify the studied word. Carets appeared under the two letters to be interchanged, and these letters were chosen under the constraints that the letters not be the same or adjacent, and that neither be the first letter of the word. As in previous experiments reported herein, all instructions were read on the computer monitor and reiterated aloud by the experimenter. Participants used labeled keys to make their source-monitoring judgments. Thus, the general procedure was virtually identical to that reported previously.

### Results and Discussion

Although we had thought source memory might vary depending on the specific type of emotion tested, it did not. Therefore, for the purposes of this article, we have pooled over the love and joy items to form the positive class of items, and pooled over the fear, sadness, and anger items to form the negative class of items. We tested a 2 (condition)  $\times$  3 (valence) mixed ANOVA model on the inferred recognition hit rates. Neither the main effect of condition nor the main effect of valence was statistically significant. There was, however, an interaction

[ $F(2,132) = 3.42$ ]. The seen and heard combination of sources evidenced a pattern reminiscent of Experiments 2 and 3, in which the inferred recognition hit rate was higher for valenced words than for neutral ones. By contrast, there were no differences among the three classes of items with the anagram and frequency sources, and perhaps a nominal benefit to the neutral items. As the reader must have anticipated, when the high level of false alarms is factored into the corrected recognition scores, item memory for valenced material was much worse than for neutral material [ $F(2,132) = 8.92$ ].

In the identical analysis on source performance (see Table 2), the deficit for valenced items relative to neutral replicated the previous experiments [ $F(2,132) = 43.28$ ]. There was neither a main effect of condition nor an interaction. Therefore, unlike Experiment 1, which also tested all three classes—positive, negative, and neutral items—simultaneously, this experiment showed a decrement in source monitoring for positive items. Even using new stimulus materials and an internal–internal combination of sources, a deficit to source performance was obtained.

We are now confident that the results from the previous experiments were not affected by stimulus characteristics or by our choice of the combination of sources. Considering the larger picture, Experiments 1–3 each had two tests of valenced materials against neutral ones, and Experiment 4 had two analogous tests in each condition for a total of four more comparisons. In sum, out of ten comparisons, nine showed that source monitoring for valenced items was worse than it was for neutral ones. Assuming chance accuracy (i.e., no association between valence and source monitoring), the binomial probability of this occurring is .0098, or less than 1% (Hays, 1994). In addition, the one failure (positive items in Experiment 1) contained means that were nominally in the correct direction of a source-monitoring deficit. Therefore, the evidence is strong that the valence of material can reduce subsequent source-monitoring accuracy. We turn now to some key observations that can be made on the basis of this study.

### GENERAL DISCUSSION

The literature on memory for valenced material can sometimes be rather tricky to navigate. As in the present article, the effect sizes are often small and sometimes do not replicate cleanly (Bower & Mayer, 1985; D'Argembeau & Van der Linden, 2004). When viewed together, the sheer weight of empirical evidence from the four experiments in this study clearly indicated that a deficit to source monitoring on valenced material can be found. This outcome stands in sharp contrast to previously published reports that a source-memory enhancement had been found for valenced material. We found this memory enhancement in the inferred recognition hit rates, but mainly when rates for either positive items alone or negative items alone were compared with those for neutral ones (Experiments 2 and 3), but not consistently when all three classes of items were tested together. What was consistent, however, was that false alarm rates for valenced material were much

higher, leading to consistent deficits in corrected recognition for valenced material. Thus, by corrected recognition, there was an emotion-related deficit in memory for valenced items. Given the “sometimes yes and sometimes no” enhancement for item memory but the very consistent deficit to source memory, one general implication of the present findings is that item memory and source memory do not always trade off as a function of valence (see the introduction of Doerksen & Shimamura, 2001; Jurica & Shimamura, 1999). In other words, valence does not necessarily confer its disadvantage to source memory because item memory is being enhanced. This observation is actually important, because some researchers (e.g., Glanzer, Hilford, & Kim, 2004) believe that a given manipulation will affect item memory and source memory in the same manner. By contrast, Johnson et al. (1993) argue quite strongly that the relationship between item and source memory is variable; one manipulation may act to increase item memory and decrease source memory, whereas another may have the opposite effect or move them in the same direction. The empirical results from this study support the assertion that the relationship is variable.

We believe that in theory valenced items can attract more attention. The heightened attention paid to them can reduce binding of the contextual details surrounding the study episode. In other words, valenced items can attract more attention or elaborate processing, but the resulting cognitive processing may not leave memory details that are criterial and diagnostic to the memory test (Yonelinas & Jacoby, 1996). For example, noncriterial attributes would be laid down in the memory trace when the item attracts attention, and that attention brings to mind stimulus-related thoughts that subsequently do not help on the memory test. Kensinger and Corkin (2003) speculated that participants performing a remember-know test might be recollecting idiosyncratic and personal associations to valenced items at a greater rate than for neutral items. Given a remember-know test, those thoughts and associations are indeed criterial for that type of test. In the present case, with a source-monitoring test, those thoughts and associations may not necessarily come to mind (or be relevant) at test, because the query is about the contextual details of the study episode. Nevertheless, their occurrence at encoding has the ability to reduce the contextual information stored in the first place, as we found.

The present article highlights the need to understand when a source-memory enhancement, as opposed to a decrement, will be found with valenced materials. With so little empirical evidence available, and with all of the previous studies using only memory for color (and one experiment on spatial location), we are not likely in this one article to pinpoint what variable or variables may account for the dramatically different results. However, color memory has been argued to be encoded relatively automatically (e.g., Chalfonte & Johnson, 1996). With more attention being paid to the item, perhaps more color information is stored and better subsequent source memory is observed. By this argument, any contextual attribute stored relatively automatically will enjoy better source

memory for valenced material than for neutral material. One problem with this argument, however, is that better memory for spatial location has been found for valenced material; and Chalfonte and Johnson argue that location is not bound automatically. However, our own work on spatial location suggests that it is bound more automatically, because attention directed toward another contextual attribute does not reduce memory for location, and may even improve it (Marsh, Hicks, & Cook, 2004). Thus, we acknowledge that there is some debate about what contextual features are bound more automatically, and this places some limitations on the foregoing account.

A slightly different account is based on the fact that older adults clearly have deficits on spatiotemporal source memory (Spencer & Raz, 1995). Such contextual details tend to be extrinsic to processing the items, and tend to require more resources for binding that information into memory. If the same is true in younger adults, perhaps contextual details extrinsic to processing, such as modality information, are most affected by valence, whereas an intrinsic property like color is not. Unfortunately, neither of these accounts adequately addresses the fact that worse source monitoring was found in a pair of internal-internal sources (anagram and frequency) in Experiment 4. Until more empirical data can be gathered on this issue, perhaps the best, albeit very general, explanation is that any contextual detail whose learning benefits from more available resources during learning has the potential to be negatively affected when processing valenced material.

The results presented here also speak to Kihlstrom's (1995) concern that it will be difficult to show weapon-focus- or trauma-memory-like deficits in the laboratory. That concern is probably appropriate for item memory, where valence often confers an advantage. That concern is obviously not true of memory for contextual details, because we have found a remarkably consistent deficit in source monitoring for valenced items across very different stimulus sets. The present results are more in line with findings that peripheral details of emotional pictures are remembered less well than details for neutral pictures (e.g., Burke, Heuer, & Reisberg, 1992). Those same studies tend to find better memory for central details of emotional rather than neutral material. One major difference is that most studies intermingle valenced and neutral items; it is not clear that pure lists would lead to the same effect. Dewhurst and Parry (2000) have argued that intermingling items depresses memory for neutral material, presumably through a combination of making valenced items distinct and disrupting residual rehearsal processes on the neutral material. By contrast, we have shown the same effects both in mixed-list and pure-list designs.

During the review process, a recent study by Anderson and Shimamura (2005) was brought to our attention. In that study, the authors showed people positive, negative, arousing, and neutral (control) film clips, none with audio, while playing words in the background. Later memory for which film clip was associated with the words was poorest for items associated with the negative film clip. That outcome is obviously slightly different from standard

source memory tests, but nevertheless is conceptually the same result that we found in the four experiments presented here. Consequently, we are not the only ones to find deficits associated with valenced materials.

In conclusion, the memory enhancement for contextual details that was found by others is an interesting effect indeed. We were not able to replicate that effect, but rather, found consistent decrements to source-monitoring performance. That outcome cannot be due to differential rehearsal, changes in distinctiveness, or item characteristics, because we found the same effect using different materials and different experimental designs. We only tested two combinations of sources (seen and heard and anagram and frequency), and perhaps the particular combination of sources determines the effect that valence has on memory for contextual details (i.e., within-modality produces enhancements, whereas cross-modality produces decrements). Regardless what variables are eventually discovered to modulate the enhancement versus decrement outcome, the present results argue that a decrement can be consistently found. Our present analysis argues that valenced material can attract attention and spontaneous elaboration, and can result in cognitive processing that may or may not be beneficial for item memory but that can, in any case, reduce the available resources for binding context information into a memory trace. Not only is this theoretically possible, the present results are evidence that it can be empirically demonstrated.

#### AUTHOR NOTE

We thank Todd Lindsey, David Cloud, Colleen Sloan, and Field Vanmeter for their invaluable assistance in collecting the data. Correspondence concerning this article should be addressed to R. L. Marsh, Department of Psychology, University of Georgia, Athens, GA 30602-3013 (e-mail: rlmash@uga.edu).

*Note—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.*

#### REFERENCES

- ANDERSON, L., & SHIMAMURA, A. P. (2005). Influences of emotion on context memory while viewing film clips. *American Journal of Psychology*, **118**, 323-337.
- BANKS, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science*, **11**, 267-273.
- BOWER, G. H., & MAYER, J. D. (1985). Failure to replicate mood-dependent retrieval. *Bulletin of the Psychonomic Society*, **23**, 39-42.
- BRADLEY, M. M., & LANG, P. J. (1999). *Affective norms for English words (ANEW)*. Gainesville, FL: NIMH Center for the Study of Emotion & Attention, University of Florida.
- BUCHANAN, T. W., & ADOLPHS, R. (2002). The role of the human amygdala in emotional modulation of long-term declarative memory. In S. C. Moore & M. Oaksford (Eds.), *Emotional cognition: From brain to behaviour* (pp. 9-34). Amsterdam: Benjamins.
- BURKE, A., HEUER, F., & REISBERG, D. (1992). Remembering emotional events. *Memory & Cognition*, **20**, 277-290.
- CAHILL, L., & MCGAUGH, J. L. (1995). A novel demonstration of enhanced memory associated with emotional arousal. *Consciousness & Cognition*, **4**, 410-421.
- CHALFONTE, B. L., & JOHNSON, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, **24**, 403-416.
- CHRISTIANSON, S.-Å. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, **112**, 284-309.
- D'ARSEMBEAU, A., & VAN DER LINDEN, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, **4**, 173-188.
- DEWHURST, S. A., & PARRY, L. A. (2000). Emotionality, distinctiveness and recollective experience. *European Journal of Cognitive Psychology*, **12**, 541-551.
- 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, & Cognition*, **24**, 1121-1136.
- DOERKSEN, S., & SHIMAMURA, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, **1**, 5-11.
- EASTERBROOK, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, **66**, 183-201.
- GLANZER, M., HILFORD, A., & KIM, K. (2004). Six regularities of source recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 1176-1195.
- GUTTENTAG, R. E., & CARROLL, D. (1997). Recollection-based recognition: Word frequency effects. *Journal of Memory & Language*, **37**, 502-516.
- HAMANN, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, **5**, 394-400.
- HAYS, W. L. (1994). *Statistics* (5th ed.). Fort Worth, TX: Harcourt Brace.
- HICKS, J. L., MARSH, R. L., & RITSCHER, L. (2002). The role of recollection and partial information in source monitoring. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 503-508.
- JOHNSON, M. K., HASHTRUDI, S., & LINDSAY, D. S. (1993). Source monitoring. *Psychological Bulletin*, **114**, 3-28.
- JOHNSON, M. K., NOLDE, S. F., & DE LEONARDIS, D. M. (1996). Emotional focus and source monitoring. *Journal of Memory & Language*, **35**, 135-156.
- JOHNSON, M. K., & RAYE, C. L. (1981). Reality monitoring. *Psychological Review*, **88**, 67-85.
- JURICA, P. J., & SHIMAMURA, A. P. (1999). Monitoring item and source information: Evidence for a negative generation effect in source memory. *Memory & Cognition*, **27**, 648-656.
- KENSINGER, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, **15**, 241-251.
- KENSINGER, E. A., & CORKIN, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, **31**, 1169-1180.
- KIHLSTROM, J. F. (1995). The trauma-memory argument. *Consciousness & Cognition*, **4**, 63-67.
- LEDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Simon & Schuster.
- MARSH, R. L., & HICKS, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1137-1151.
- MARSH, R. L., HICKS, J. L., & COOK, G. I. (2004). Focused attention on one contextual attribute does not reduce source memory for a different attribute. *Memory*, **12**, 183-192.
- MARSH, R. L., HICKS, J. L., & DAVIS, T. T. (2002). Source monitoring does not alleviate (and may exacerbate) the occurrence of memory conjunction errors. *Journal of Memory & Language*, **47**, 315-326.
- METCALFE, J., & JACOBS, W. J. (1998). Emotional memory: The effects of stress on "cool" and "hot" memory systems. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 38, pp. 187-222). San Diego: Academic Press.
- OCHSNER, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General*, **129**, 242-261.
- PERFECT, T. J., MAYES, A. R., DOWNES, J. J., & VAN ELK, R. (1996). Does context discriminate recollection from familiarity in recognition memory? *Quarterly Journal of Experimental Psychology*, **49A**, 797-813.
- READ, J. D., & LINDSAY, D. S. (1997). *Recollections of trauma: Scientific evidence and clinical practice*. New York: Plenum.
- SHAYER, P., SCHWARTZ, J., KIRSON, D., & O'CONNOR, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality & Social Psychology*, **52**, 1061-1086.
- SPENCER, W. D., & RAZ, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology & Aging*, **10**, 527-539.
- TALMI, D., & MOSCOVITCH, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition*, **32**, 742-751.



YONELINAS, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 1415-1434.

YONELINAS, A. P., & JACOBY, L. L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. *Consciousness & Cognition*, **5**, 131-141.

**NOTE**

1. The use of incidental encoding is generally required in source-monitoring studies if one wants to claim that one is studying memory for context. Otherwise, with intentional encoding of source information, that information is in focal attention, which certainly detracts from being able to claim that one is studying contextual information. However, we would agree that, even with intentional encoding, one is still studying source memory.

**APPENDIX**

Experiment, Condition, and Claim	Valence and Source								
	Positive Items			Negative Items			Neutral Items		
	Seen	Heard	New	Seen	Heard	New	Seen	Heard	New
Experiment 1									
Seen	.59	.22	.14	.54	.19	.16	.60	.19	.12
Heard	.11	.59	.09	.22	.63	.19	.12	.62	.07
New	.30	.19	.76	.24	.17	.66	.28	.19	.81
Experiment 2									
Negative and Neutral									
Seen				.60	.17	.17	.63	.12	.08
Heard				.25	.69	.20	.16	.70	.07
New				.15	.14	.60	.21	.18	.85
Pure Negative									
Seen				.61	.19	.15			
Heard				.19	.65	.17			
New				.20	.17	.68			
Pure Neutral									
Seen							.62	.15	.09
Heard							.14	.64	.09
New							.24	.21	.82
Experiment 3									
Positive and Neutral									
Seen	.55	.19	.13				.58	.14	.05
Heard	.18	.61	.11				.11	.58	.08
New	.27	.20	.76				.31	.28	.87
Pure Positive									
Seen	.56	.19	.09						
Heard	.17	.61	.11						
New	.27	.20	.79						
Pure Neutral									
Seen							.63	.14	.08
Heard							.15	.65	.08
New							.24	.21	.84

Note—Some subcolumns will not add up to 1.0 because of rounding error.

(Manuscript received December 14, 2004;  
revision accepted for publication November 16, 2005.)