Rate of false source attributions depends on how questions are asked.

by Chad S. Dodson and Marcia K. Johnson

Lindsay and Johnson (1989) and Zaragoza and Koshmider (1989) report evidence indicating that "eyewitness" subjects are much less likely to falsely claim to have seen information suggested to them verbally when they receive a source monitoring test than when they receive a recognition test requesting only identification of the seen information. The present study reports additional evidence that source misattributions are affected by the nature of the test. Intraub and Hoffman (1992) recently reported the results of a study in which subjects claimed to have seen pictures corresponding to scenes that had only been described in paragraphs they had read. With this paradigm, we found a similar effect using their test, but source confusions were reduced with a test patterned after the one used by Lindsay and Johnson. We attribute this difference in performance to the different decision criteria evoked by these two tests.

Source monitoring errors occur when we confuse the origin of our memories, such as believing that we learned about a new restaurant from Bill when actually Dave told us about it. Johnson and her colleagues (Johnson, 1988; Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981) have proposed a source monitoring framework that describes the processes involved in attributing a memory to a particular source. According to their model, source monitoring attributions are influenced by two factors: (a) the characteristics of the memory, that is, the amount of perceptual, contextual, semantic, and affective information, as well as information about cognitive operations (i.e., cognitive processes involved while experiencing the event); and (b) the kind of decision processes used at the time of test to make a judgment about the source of the memory.

The characteristics of the memory depend on how one experienced the event (e.g., whether it was imagined or perceived) and what one was thinking and feeling at that moment of acquisition. Typically, for example, self-generated or imagined memories have less perceptual information than externally derived memories (Suengas & Johnson, 1988). When we monitor the sources of our memories, we make use of such qualitative differences to make correct source attributions (Johnson et al., 1992; Johnson & Raye, 1981). However, sometimes these differences are subtle, and source confusions occur.

In fact, a number of studies have shown that as the qualities of our externally derived (e.g., having seen a picture) and our internally generated (e.g., having imagined a picture) memories increase in similarity, an increase in source confusions occurs (Durso & Johnson, 1980; Johnson, Foley, & Leach, 1988; Johnson, Raye, Wang, & Taylor, 1979; Lindsay, Johnson, & Kwon, 1991; Rabinowitz, 1989). For instance, Durso and Johnson (1980) showed subjects line drawings and verbal labels of pictures (e.g., knife). Some subjects (intentional imagers) were asked to imagine the presented items as pictures, whereas other subjects (incidental imagers) answered questions about the function of each item (e.g., "What is the object used for?"). On a subsequent recognition test, subjects heard labels of new items and of items that had been presented initially as line drawings and words. Durso and Johnson (1980) found that the intentional imaging subjects were more accurate than the incidental imaging subjects at identifying which items had been presented as line drawings and words. Subjects receiving the incidental orienting task experienced more source confusions because they were presumably less aware that they were constructing images of the items, producing little information about cognitive operations in the imagined memory. Typically, one basis for distinguishing between memories for perceived and imagined items is the presence of more information about cognitive operations in imagined memories than in perceived memories. However, in this experiment, the relative absence of this information in the incidental group increased the similarity in the phenomenal qualities of their imagined and perceived memories, and thereby increased source confusions (see also Intraub & Hoffman, 1992).

The frequency of source confusions is also affected by the judgment processes involved when remembering, especially the criteria used to evaluate the characteristics of the memory. That is, the purpose of remembering, the cost of making a mistake, what information one uses, and so forth, influence the accuracy of making source attributions. For example, Raye, Johnson, and Taylor (1980) found that subjects' judgments of how often a word had been presented increased as a function of how often the subjects had generated the word. However, if subjects' judgments of frequency of appearance were constrained by allowing only estimates that were less than 10, they gave more accurate frequency estimations. Evidently, this constraint strengthened the criteria that subjects used in evaluating the origin of memories (i.e., determining

Rate of false source attributions depends on how questions are asked.

whether they were perceived or generated), so that subjects could edit out the generated occurrences and give more accurate estimates of frequency of perceived occurrences. In sum, the source monitoring framework and the preceding studies point to two factors that cause source confusions: (a) similarities in the characteristics of subjects' memories from various sources; and (b) subjects' use of loose criteria when remembering.

The fact that criteria affect accurate remembering is especially important in the eyewitness testimony paradigm (Lindsay & Johnson, 1989). In this paradigm (Loftus, Miller, & Burns, 1978), subjects are given to-be-remembered material, usually in the form of a slide or slide sequence. They then read an ostensibly accurate written description of the material, but the description contains misleading information, either referring to peripheral items that were seen in the slides as something else or implying that entirely new items were present in the slides. For example, a wrench that had been presented in the slides is referred to as a hammer in the text. On a final recognition test for information from the slides, subjects often respond on the basis of the misleading information, claiming to have seen items, such as a hammer, that were only read about. This occurs even though subjects have been instructed to rely only on the information in the slides. In essence, subjects confuse the sources of the information, mistakenly reporting that the misinformation was acquired during the slide sequence.

Although the misinformation effect is robust using standard recognition tests, other studies using this paradigm have found few source confusions when subjects were given a source monitoring test that orients them to assess their memory in terms of the possible sources of the learned information. That is, for each item on the test, subjects must indicate whether it had been seen, read, both seen and read, or was new (Lindsay & Johnson, 1989; Zaragoza & Koshmider, 1989). Lindsay and Johnson (1989) suggested that criterial differences between the recognition and source monitoring tests may explain the differences in the probability of making source confusions. Most recognition tests query subjects only about whether the target item was in the slide sequence and may be answered according to the item's degree of familiarity; the most familiar items are marked as having been seen in the slides. In contrast, most source tests require subjects to use different, perhaps stricter, criteria and to evaluate memory in terms of all the possible sources for this target item (e.g., was it read, seen, both read and seen, or new?). Judgments of this sort cannot be based on familiarity, but must rely on the characteristics of the memories (e.g., how much and what type of perceptual detail and cognitive operations it includes). Although source confusions occur on most recognition tests, confusions are reduced with the use of source monitoring tests.

A paradigm used by Intraub and Hoffman (1992) presents a further opportunity for examining the importance of decision processes and criteria for making source attributions. In their study, subjects initially viewed a sequence of unrelated slides; they then read a series of paragraphs and answered a question accompanying each paragraph. A few of the paragraphs described a few of the slides, and the rest described novel scenes. Then, for each paragraph presented at test, subjects answered two yes/no questions: "Read before?" and "If yes, then was it seen?" Intraub and Hoffman's test appears similar to typical source monitoring tests in that subjects were explicitly asked to consider two potential sources (verbal and pictorial) of their memory of each item. Even though subjects were presumably oriented to examine the sources of their memories on this test, they frequently made source confusions, claiming to have seen items that were only read about.

The present experiment examined whether a test more like the one used by Lindsay and Johnson (1989) would reduce false source attributions in the Intraub and Hoffman (1992) paradigm. There are two major differences between the tests used by Intraub and Hoffman and by Lindsay and Johnson: the size of the cue, and the format of the response options. Lindsay and Johnson's test used a brief phrase as a cue, whereas Intraub and Hoffman's test employed a full paragraph that was identical, in the case of targets, to the previously read paragraphs. The full cue, however, should be a better reinstatement cue than a partial cue; therefore, we would expect fewer errors with full cues than with partial cues.

We think the more critical factor is the format of the response options. The response format on Lindsay and Johnson's (1989) test required subjects to check, for each cue, one of four source columns (i.e., slide, text, both, new) that corresponded to their memory of the source for that cue; this is in contrast to the binary question format used by Intraub and Hoffman (1992). Because the Lindsay and Johnson format makes all the possible sources of memories more explicit than does the Intraub and Hoffman format, it is conceivable that this difference will induce subjects receiving the Lindsay and Johnson format to employ stricter criteria when assessing their memories, resulting in fewer source confusions.

As a way of clarifying how test features influence the rate of false source attributions, our experiment examined the

Rate of false source attributions depends on how questions are asked.

separate roles of the size of the cue and the type of response format. Half of the subjects were tested with full paragraph cues, and the remaining half with brief phrasal synopses. Further, half of the subjects received tests with the binary question response format used by Intraub and Hoffman (1992), and the remaining subjects received tests with a four alternative forced-choice (4AFC) source-listing response format used by Lindsay and Johnson (1989).

We were also interested in how processing information in an emotional (or apperceptive) manner influences the likelihood of source confusions. A study by Suengas and Johnson (1988) showed that thinking about a previously performed or imagined event in an apperceptive manner (i.e., thinking about one's thoughts and feelings) reduced subjective ratings of the amount of perceptual information (e.g., objects, colors) associated with the event, relative to subjects who thought about the previous event in a perceptual manner. Insofar as perceptual information implies that a memory was perceptually derived, a reduction in perceptual information would result in memories more like those of imagined events. Consistent with the idea that emotional processing might reduce accurate source monitoring, Hashtroudi, Johnson, Vnek, and Ferguson (1992; also described in Johnson, 1991) reported that older adults (mean = 70 years) showed less accurate source monitoring than younger subjects (mean = 20 years) after thinking about their affective reactions to events but not after thinking about factual aspects of events. However, source monitoring of young adults was not influenced by the type of focus. Thus, the impact of emotion on source monitoring is only beginning to be investigated, and examining this variable under a range of conditions would be useful. In our study, half of the subjects processed the paragraphs in a perceptual manner by answering a factual question about each paragraph; this condition replicated Intraub and Hoffman's (1992) study. The remaining subjects processed the paragraphs in an emotional manner by generating an emotional response after reading each paragraph; no question preceded the paragraphs given to this group. The point of interest was the effect on source monitoring of emotional relative to perceptual processing of the paragraphs.

In short, our study used the same stimuli and much of the procedure from Intraub and Hoffman (1992). Subjects first saw a series of 60 unrelated pictures, and then read aloud 50 unrelated paragraphs (10 of the paragraphs described 10 pictures). At acquisition, before reading each paragraph, half the subjects were presented with a question which they read aloud (perceptual processors). The question asked about some perceptual detail that was implied by the paragraph. The remaining subjects (emotional processors) did not receive any questions, but read the paragraph aloud and then generated an emotional response. All the subjects were then given a test containing 60 items: 40 probed items that had been read only, 10 probed items that had been read and that had matched a previously seen picture, and 10 probed items that were never experienced. At the test, we varied size of cue (partial or full) and response option (binary of 4AFC).

EXPERIMENT

METHOD

Subjects

Participants were 96 male and female paid volunteers who were either Princeton students or visiting summer students.

Materials

The 110 slides and 60 paragraphs were the same stimuli as those used by Intraub and Hoffman (1992). All subjects were tested individually and were randomly assigned to one of eight conditions. The eight conditions were formed by three between-subjects variables: type of paragraph processing (perceptual or emotional); size of test cue (full or partial); and type of response option on the test (binary question or 4AFC source-listing).

Initially, all subjects saw a randomly mixed set of 60 unrelated slides that depicted various ordinary scenes (e.g., a mountain behind a farm, a living room scene, a boy and his father at the zoo). Fifty of the slides were the same for everybody. The remaining 10 slides came from a set of 60 slides that corresponded to 60 paragraphs. This set of 60 slides was randomly divided into six groups, A through F, that were randomly assigned to subjects. The slides were shown at 5 s/slide, and were rear-projected onto a screen 0.9 m from the subject, with a projected size of 19.0 cm by 12.7 cm.

After viewing the slides, subjects read one of six different groups (A through F) of 50 paragraphs, each differing in the particular set of 10 paragraphs that corresponded to 10 of the previously seen slides. Each paragraph had two to four

Rate of false source attributions depends on how questions are asked.

sentences and was presented on a Macintosh SE computer. Half of the subjects (perceptual paragraph processors) received a question before each paragraph that asked about some strongly implied detail in the paragraph. For example:

What are the steps made of? / To get to the house you walk up the set of steps and turn to the left. The mason did a good job laying them. Each one was perfectly aligned to form the three red steps.

The subject would respond with "brick." The remaining subjects (emotional paragraph processors) saw no question before each paragraph, but generated an emotional response instead. They were told:

This could be a statement of the emotional state that is depicted in the item, or this could be whatever emotion that the item evokes for you. By emotion, I refer to anything from happy or sad to respectful or nostalgic. Try to give an emotional response for each passage, but if you really can't generate one then say "indifferent." When you see the passage, please read it aloud and then tell me your emotional response. I will record your answer.

After reading the paragraphs, subjects received one of four tests which factorially combined two variables: the size of the cue, and the format of the response options. Half of the tests used full paragraphs, and the remaining tests used brief phrasal synopses of the paragraphs. Most of the phrasal synopses included the answer to the question, such as, "brick steps lead up to a house." Further, half of the tests (binary question tests) asked two questions for each cue: "Read before?" and "If yes, was it also a picture?"; subjects circled yes or no for each question. On the remaining tests (4AFC), subjects selected for each cue a column associated with four possible sources: slide, text, both, new. The test used by Intraub and Hoffman (1992) had a full cue with a binary-question response format. On each test, 40 cues probed items that had been only read before, 10 cues probed items that had been read and that corresponded to a previously seen picture, and 10 cues probed items that had never been experienced. Cues were presented in a random order.

Procedure

All subjects were told that we were studying how people process information from pictures and from written text and that they would participate in two experiments: a picture memory experiment, and a reading comprehension experiment. Subjects were informed that in the picture memory experiment, they would see a series of unrelated pictures and then they would be asked some questions about the pictures. The final source monitoring test was not mentioned. To give credence to our cover story of a separate picture memory experiment, all subjects were asked four questions after viewing the slides (e.g., "Do you remember seeing a baby playing in the sand?"). Two questions were correctly answered yes and two no.

After the questions, all subjects were informed that in the reading comprehension experiment they would see a series of short descriptions which must be read aloud. The subjects in the perceptual group were informed that each of the descriptions was preceded by a question that asks about some detail in the description. These subjects were told to "read everything aloud--first the question then the paragraph--and then tell me the answer to the question which I will record." The subjects in the emotional group did not receive a question with the paragraph, but were told to read each paragraph aloud and then to generate an emotional response to each description.

After reading the last paragraph, all subjects were told that "during the reading comprehension experiment there was an occasional paragraph that corresponded to one of the previously seen pictures." All subjects were then given one of four tests; for each one, the instructions read as follows:

In the following list there are brief descriptions of information that you previously saw in a slide, read about in a paragraph or did not experience before. For each brief description indicate how you experienced it.

Those subjects who received the 4AFC source test were told to indicate the source of their memory for the description by checking the appropriate column. The subjects receiving the binary question test were told that for each description they should indicate whether they had read it before. If they had read the description, they were told to indicate whether they saw a corresponding slide.

RESULTS

Rate of false source attributions depends on how questions are asked.

Unless otherwise noted, results are significant at p |is less than~ .05 or better.

False picture errors

The primary result of interest concerns how often subjects failed to source monitor accurately, claiming to have seen items that were only read about. In other words, how frequently were there false picture errors for the paragraphs that were only read and did not have a corresponding picture? On the 4AFC source test, an item was scored as having been seen if the subjects checked either the slide or the both columns. On the binary question test, an item was scored as having been seen if the subjects responded yes to the question "If yes, was it also a picture?" In all of the following 2 x 2 x 2 independent groups analyses, the factors are processing (emotional or perceptual paragraph processing), cue (full or partial), and response option (binary question or 4AFC source listing).

Table 1 presents the proportion of "read only" items recognized as "old" that were incorrectly marked as having been seen. A 2 x 2 x 2 ANOVA yielded significant main effects of response option and cue and no other significant main effects or interactions. Subjects were more likely to claim mistakenly to have seen items that had only been read when they received the binary question test than when they received the 4AFC source test, 18.5% and 11.2%, respectively; F(1, 88) = 9.56, |MS.sub.e~ = 136.01. Moreover, subjects made more of these false picture errors on tests with the partial cue (17.7%) than on tests with the full cue (12.0%), F(1, 88) = 5.81. The important finding from this analysis is that there were fewer source confusions with the 4AFC source format than with the binary question format.

Table 1. Percentage of read only items recognized as old that were false picture errors

	Response option	
Cue	**Binary**	**4AFC(a)**
Full	**14.7**	**9.3**
Partial	**22.4**	**13.1**

a 4AFC = Four alternative forced choice.

A second analysis was conducted to correct for the baseline rate at which subjects claim to have seen pictures. That is, to what extent do subjects claim to have seen items that were neither seen nor read about before? A high false picture rate for the new paragraphs would suggest that some of the false picture errors for the read only paragraphs do not reflect confusion about which items were read and which were seen, but rather a more general failure to discriminate old from new information.

A 2 x 2 x 2 ANOVA was performed on the difference between the false picture rate for the read only paragraphs that were recognized as old and the false picture rate for the new paragraphs. This ANOVA revealed a main effect for response option, F(1, 88) = 6.69, |MS.sub.e~ = 225.95, and no other significant main effects or interactions. The subjects with the binary question format made false picture errors at a magnitude of 8.5% more than the baseline rate. This is in contrast to the subjects with the 4AFC format who performed no differently from the baseline level (0.55% from baseline rate). In other words, false picture reports were eliminated with a source monitoring test that explicitly oriented subjects to assess their memories in terms of all the possible source options.

Old/new paragraph recognition

Another question of interest is the rate at which subjects correctly recognized those paragraphs that were only read and correctly rejected those paragraphs that were new. A 2 x 2 x 2 analysis was conducted on the correct recognition rate for the read only paragraphs. This analysis yielded significant main effects of cue, processing, and response option but no significant interactions. Because, for targets, the full cue was identical to the previously read paragraphs, it is not surprising that subjects receiving tests with the full cue (86.9%) had a higher recognition rate than subjects receiving tests with the partial cue (73.1%), F(1, 88) = 28.05, |MS.sub.e~ = 162.26. The two processing groups differed in their ability to recognize previously read paragraphs. Subjects who gave emotional responses (76.5%) recognized fewer of these previously read paragraphs than the perceptual processing subjects (83.6%), F(1, 88) = 7.39. Subjects receiving the

Rate of false source attributions depends on how questions are asked.

4AFC source test (84.3%) recognized more of these read only items than subjects receiving the binary question test (75.7%), F(1, 88) = 10.79.

Table 2. Percent corrected false picture rate |i.e., read only (false picture rate) - new (false picture rate)~

	Response option	
Cue	Binary	4AFC(a)
Full	7.2	0.9
Partial	9.8	0.2

(a) 4AFC = Four alternative forced choice.

A 2 x 2 x 2 ANOVA was performed on the correct rejection rates of the novel items. This analysis revealed a main effect for cue, F(1, 88) = 5.87, |MS.sub.e~ = 259.44, and no other significant main effects or interactions. As with the read only items, it is not surprising that subjects receiving tests with the full cue (90.6%) correctly rejected more novel items than did subjects receiving tests with the partial cue (82.7%). This analysis also indicates that although the emotional processing subjects recognized fewer read only items than the perceptual processing subjects, they were not just saying no to more items in general than were the perceptual processing subjects. This strategy would result in the emotional processing subjects correctly rejecting more novel items than the perceptual processing subjects. This was not the case. The perceptual and emotional processing groups performed no differently at rejecting items that were new, 87.3% vs. 86.0%, respectively. Likewise, there was no difference between the subjects receiving the 4AFC source and the binary question test at rejecting these novel items, 86.7% vs. 86.6%, respectively.

Items both seen and read

A 2 x 2 x 2 ANOVA was performed on the proportion of items that subjects correctly recognized as both seen and read. There was a main effect of response option that approached significance, F(1, 88) = 2.93, |MS.sub.e~ = 356.12, p |is less than~ .09, and no other main effects or interactions. The subjects receiving the 4AFC source test (67.9%) recognized fewer of these both seen and read items than the subjects receiving the binary question test (74.5%). A 2 x 2 x 2 ANOVA was conducted on the proportion of both seen and read items that subjects incorrectly marked as read only. This analysis showed a main effect of cue, F(1, 88) = 4.46, |MS.sub.e~ = 284.07, and an effect of response option that approached significance, F(1, 88) = 3.50, p |is less than~ .06; there were no other main effects or interactions. Subjects receiving tests with the full cue (25.8%) more often responded incorrectly that these both seen and read items had been read only than did subjects receiving tests with the partial cue (18.6%). More important for interpreting our false picture errors is that subjects receiving the 4AFC source test (25.4%) more often said that items presented as both pictures and paragraphs had only been read than did subjects receiving the binary question test (19.0%). There was no difference between the 4AFC source (2.3%) and the binary question (3.2%) test in the rate at which subjects incorrectly referred to these both seen and read items as new, F(1, 88) = .52.

DISCUSSION

The important result from this experiment is that although subjects receiving the binary question test claimed falsely to have seen items that were only read, replicating Intraub and Hoffman's (1992) results, these false picture errors did not occur on the 4AFC source test. We do not interpret this result to mean that there will never be false picture errors on a 4AFC source test under any condition; the important point is that a 4AFC source test can be expected to produce fewer false picture errors than a binary question test.

It might at first seem that relative to the binary question test, the 4AFC source test simply prompts subjects to adopt a higher threshold for saying yes on the basis of familiarity--thereby making fewer false picture errors. However, they did not recognize read only paragraphs at a lower rate (in fact, they performed better at this than the subjects receiving the binary question test), nor did they have higher correct rejection rates on new items. Perhaps most persuasive, it is not clear how an interpretation that says that the 4AFC test induced subjects to evaluate an item's familiarity more strictly than did the binary question test could explain the pattern of results on the both seen and read items. Subjects receiving the 4AFC test

Rate of false source attributions depends on how questions are asked.

were incorrect more often on both seen and read items, but they did not call such items new more often. Rather, they more often incorrectly claimed that items both seen and read had been read only. Such a pattern suggests that the 4AFC subjects did not use a higher familiarity criterion, but rather had a higher criterion for some more specific type of information. One reasonable interpretation is that the 4AFC subjects were using a higher criterion for evaluating their memory for perceptual information than were the subjects taking the binary question test.

Perhaps, on the first question of the binary question format, subjects use familiarity to decide whether a paragraph was read before, and then on the second question subjects rely on the presence or absence of perceptual information to decide whether they have seen a picture of the item. Any perceptual information or imagery that is associated with memory for the item may be sufficient "evidence" for subjects to respond that they have seen the item. This loose criterion causes these subjects to report that items that were only read had been seen in the slides. That is, they take memories for images that arose from having read the item as signifying that the item had been seen. In contrast, subjects receiving the 4AFC test evidently require a greater amount of perceptual detail in order to attribute an item to the slides. Apparently, they realize that perceptual detail alone does not mean that the item has been seen, and thus they examine the amount or type of perceptual detail in their memory more carefully.

Because the subjects taking the 4AFC source test employ a stricter criterion for evaluating perceptual information, they make fewer false picture errors and they more often fail to correctly recognize the both seen and read items than do the subjects taking the binary question test. However, as emphasized above, those subjects receiving the 4AFC test do not more often fail to recognize the both seen and read items as old, because there is no difference between the two tests in the rate at which subjects incorrectly refer to these items as new. In short, both tests demand that subjects assess the source of their memories; the difference between these tests is how they induce subjects to make this assessment. Perhaps considering possible sources simultaneously rather than sequentially (as in the binary question test) makes the overlap in qualitative characteristics of memories from various sources more salient and hence induces stricter criteria for using any particular cue as diagnostic of a specific source.

The present study expands Lindsay and Johnson's (1989) argument about decision processes and criteria in source attributions. Although the term criteria has generally referred to the amount of a particular aspect of memory, such as the degree of familiarity, required for a given judgment, we want to emphasize that criteria also refer to the types of mnemonic information, such as perceptual or cognitive operations, that individuals rely on for a memory judgment (Lindsay & Johnson, 1989). Those investigators note that source confusions are likely to occur on a recognition test in the eyewitness testimony paradigm because subjects may rely on familiarity to make a judgment about the source of an item. They argue that source confusions are less likely when subjects take the 4AFC source test because subjects use criteria different from familiarity; instead, they examine the characteristics of memory. The present study further suggests that subjects may examine particular qualitative aspects of their memory, using either relatively loose or relatively strict criteria. Specific characteristics of a test may, then, orient subjects to assess familiarity or to assess more specific memory qualities; but even various memory qualities can be assessed using criteria that differ in strictness. For identifying the origin of pictorial information, we think that tests similar to the 4AFC format should orient subjects to assess memory for perceptual information more strictly than yes/no recognition tests or source monitoring tests similar to the binary question format, which orient subjects to rely on familiarity or to use a loose criterion when assessing perceptual information. In general, more accurate source monitoring should occur when all of the potential sources for a memory are simultaneously salient to the rememberer.

Finally, it is interesting that subjects in the emotional processing group recognized fewer paragraphs than did those in the perceptual processing group. This result is consistent with the finding that an emotionally arousing event can impair memory for information, especially peripheral information, that was associated with the arousing event (Christianson, 1991). Apparently, emotional arousal limits the subjects' ability to remember information that is outside their focal attention; memory is impaired for information on the periphery. Although the paragraphs in our study described ordinary events and were generally not especially arousing, thinking about emotions that are associated with the paragraphs may have an impact on memory similar to being emotionally aroused. That is, as suggested by Suengas and Johnson (1988), in both cases, subjects may focus on their feelings at the expense of focusing on more central factual information. In support of this idea of a trade-off between emotional and factual information, Suengas and Johnson reported that the subjective ratings of the amount of perceptual information associated with an event were reduced when subjects rehearsed thoughts and feelings associated with that event compared with when they rehearsed perceptual aspects of that event. In our study, emotional processing may impair memory for perceptual details, such as colors and objects in the

scenes, that are described by the paragraphs.

At test, subjects must differentiate new scenes from previously described scenes. Memory for specific perceptual details should be more helpful for this discrimination than more abstract memories that include affective associations. In other words, the factual question that accompanied each paragraph in the perceptual processing condition evidently provided a more effective technique for remembering specific perceptual details and for discriminating old from new paragraphs than did generating an emotional response. In essence, answering the question induced subjects to focus on the gist and on perceptual details described or implied in the paragraphs. In contrast, subjects in the emotional condition were allowed to produce whatever feelings were triggered by the paragraphs; these feelings need not have related directly to the gist or to specific details of the paragraph. Because the partial and full test cues were either brief paraphrases (essentially the gist) or full paragraphs, they probably were more effective in reinstating the gist than in reinstating the affective response, producing better recognition cues for perceptual than for emotional processing subjects. Moreover, generating an emotional reaction may have been a relatively ineffective basis for discriminating old from new paragraphs because the new paragraphs might have produced some of the same affective reactions as did the old paragraphs.

In contrast to our results, not all studies have shown that thinking about thoughts and feelings reduces old/new recognition memory for aspects of a previous event. In the previously mentioned study by Hashtroudi et al. (1992), pairs of subjects participated in an interactive task (e.g., each played a role in a short play) and then rehearsed this interactive task in one of three ways: rehearsed what they had been feeling (affective group); rehearsed what had been said (factual group); rehearsed anything associated with the task (control group). Hashtroudi et al. found no difference between the groups in their old/new recognition ability for statements from the task. In other words, thinking about feelings did not, in this case, impair recognition memory for objective information, such as specific statements, from the event.

Why would we find in the present study an old/new recognition decrement from emotional processing, whereas Hashtroudi et al. (1992) did not? There were many differences in the two studies, so we can only speculate at this point. One interesting possibility to explore in the future is that in the Hashtroudi et al. study there was a clear separation between the initial interactive task phase and the subsequent rehearsal (and manipulation of focus) phase. In our study, there was no clear line separating encoding from rehearsal. That is, our subjects very likely encoded the paragraphs differently when they were prepared to answer a factual question (perceptual processors) than when they were to generate an emotional response (emotional processors). In the Hashtroudi et al. study, the differential focus at rehearsal came after an initial encoding phase in which the groups of subjects presumably encoded events in a similar fashion. Perhaps, for old/new recognition, emotional focus at acquisition is more disruptive than subsequent emotional focus in thinking back over events.

Hashtroudi et al. (1992) did find that their affective group recalled fewer items from the initial interactive task than did their factual group. This greater sensitivity of recall than recognition to the effects of emotional focus is consistent with the generally greater sensitivity of recall than recognition to manipulations affecting encoding, retrieval conditions, or both (e.g., Alba, Alexander, Hasher, & Caniglia, 1981; Godden & Baddeley, 1975, 1980; Smith, Glenberg, & Bjork, 1978). According to the idea that focusing on emotions at acquisition is more disruptive than emotional focus at rehearsal, had we included recall conditions in the present study, we would expect the decrement from emotional processing relative to perceptual processing to have been even greater in recall than what we obtained in recognition.

The equivalent source monitoring performance by our two processing groups is consistent with the performance of the young adult subjects in the study by Hashtroudi et al. (1992). Those subjects did not experience reduced source monitoring performance after having thought about events in an emotional rather than a factual manner. We are not surprised (in retrospect) that our subjects in the emotional processing condition did not show poorer source monitoring performance than our subjects in the perceptual processing condition. The ability to source monitor accurately relies on exploiting the dissimilarities between memories, assuming that it is easier to distinguish between memories that are dissimilar than similar (e.g., Johnson et al., 1993; Lindsay et al., 1991). In the present experiment, subjects must discriminate paragraphs that had a corresponding picture from those that did not. The perceptual processing subjects' memories of the paragraphs and pictures are likely to be more similar than these same memories in the emotional processing subjects. The perceptual processing subjects are distinguishing memories of pictures from memories of paragraphs that very likely have a pictorial representation, as they imaged the scene to answer the question. The emotional processors, on the other hand, are distinguishing memories of pictures from memories of paragraphs that may be represented more abstractly and with more affective associations. Whether this difference between the emotional

processors' memories of the pictures and paragraphs would actually show up as a beneficial effect on source monitoring (in contrast to the detrimental effect on old/new recognition) would depend on how the pictures were processed. Picture processing in the present study was not controlled with an orienting task, so our best guess is that it included some combination of perceptual and more affective processing.

In summary, the present finding of clear differences in the rate of false picture attributions depending on test conditions provides further evidence that source monitoring depends on assessment of various qualitative characteristics of memories according to flexible criteria for attributing memories to their sources. In addition, the results of old/new paragraph recognition provide further evidence that, relative to perceptual focus, emotional focus may disrupt memory. However, as the discussion implies, the relation between memory and affect needs clarification. A useful strategy would be to explore the potential differential effects of time of emotional focus (encoding, subsequent rehearsal, at test) on various memory measures (recall, recognition, source monitoring).

Notes

This article is based upon work supported by a National Science Foundation Graduate Fellowship and National Institute on Aging (NIA) Grant AG09253. We thank Helen Intraub for providing the stimuli used in this study.

Correspondence concerning this article should be addressed to Chad S. Dodson or Marcia K. Johnson, Department of Psychology, Princeton University, Princeton, NJ 08544-1010. Received for publication March 5, 1992; revision received July 8, 1992.

References

Alba, J. W., Alexander, S. G., Hasher, L., & Caniglia, K. (1981). The role of context in the encoding of information. Journal of Experimental Psychology: Human Learning and Memory, 4, 283-292.

Christianson, S.-A. (1991). Remembering emotional events: Potential mechanisms. In S.-A. Christianson (Ed.), The handbook of emotion and memory: Current research and theory (pp. 307-340). Hillsdale, NJ: Erlbaum.

Durso, F. T., & Johnson, M. K. (1980). The effects of orienting tasks on recognition, recall, and modality confusion of pictures and words. Journal of Verbal Learning and Verbal Behavior, 19, 416-429.

Godden, D., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. British Journal of Psychology, 66, 325-331.

Godden, D., & Baddeley, A. D. (1980). When does context influence recognition memory? British Journal of Psychology, 71, 99-104.

Hashtroudi, S., Johnson, M. K., Vnek, N., & Ferguson, S. (1992). Aging and the effects of affective and factual focus on source monitoring and recall. Manuscript submitted for publication.

Intraub, H., & Hoffman, J. E. (1992). Reading and visual memory: Remembering scenes that were never seen. American Journal of Psychology, 105, 101-114.

Johnson, M. K. (1988). Discriminating the origin of information. In T. F. Oltmanns & B. A. Maher (Eds.), Delusional beliefs: Interdisciplinary perspectives (pp. 34-65). New York: Wiley.

Johnson, M. K. (1991, June). Source monitoring: The case of emotion, aging, and reality monitoring. Paper presented at the meeting of the American Psychology Society, Washington, DC.

Johnson, M. K., Foley, M. A., & Leach, K. (1988). The consequences for memory of imagining in another person's voice. Memory & Cognition, 16, 337-342.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. Psychological Bulletin, 114, 3-28.

Rate of false source attributions depends on how questions are asked.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. Psychological Review, 88, 67-85.

Johnson, M. K., Raye, C. L., Wang, A. Y., & Taylor, T. H. (1979). Fact and fantasy: The roles of accuracy and variability in confusing imaginations with perceptual experiences. Journal of Experimental Psychology: Human Learning and Memory, 5, 229-240.

Lindsay, D. S., & Johnson, M. K. (1989). Eyewitness suggestibility and memory for source. Memory & Cognition, 17, 349-358.

Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. Journal of Experimental Child Psychology, 52, 297-318.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. Journal of Experimental Psychology: Human Learning and Memory, 4, 19-31.

Rabinowitz, J. C. (1989). Judgments of origin and generation effects: Comparisons between young and elderly adults. Psychology and Aging, 4, 259-268.

Raye, C. L., Johnson, M. K., & Taylor, T. H. (1980). Is there something special about memory for internally generated information? Memory & Cognition, 8, 141-148.

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. Memory & Cognition, 6, 342-353.

Suengas, A. G., & Johnson, M. K. (1988). Qualitative effects of rehearsal on memories for perceived and imagined complex events. Journal of Experimental Psychology: General, 117, 377-389.

Zaragoza, M. S., & Koshmider, J. W., III. (1989). Misled subjects may know more than their performance implies. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 246-255.