ACADEMIC PRESS

Journal of Memory and Language 47 (2002) 315–326

Journal of Memory and Language

www.academicpress.com

Source monitoring does not alleviate (and may exacerbate) the occurrence of memory conjunction errors

Richard L. Marsh,[a,*] Jason L. Hicks,[b] and Taylor T. Davis[a]

[a] *University of Georgia, Athens, GA 30602-3013, USA*
[b] *Louisiana State University, Baton Rouge, LA 70803, USA*

Received 24 January 2001; revision received 24 August 2001

Abstract

Four experiments investigated whether memory conjunction errors could be reduced when memories were inspected in a source-monitoring task as opposed to the standard recognition tests that have traditionally been used in this literature. A conjunction stimulus is composed of pieces that have been experienced previously but in different contexts. We found that source monitoring did not generally reduce conjunction errors, but rather increased them in some cases. The data are consistent with theoretical accounts that characterize conjunction materials as evoking high degrees of familiarity. We also present an alternative account in which partial recollective details of context can act as familiarity in order to account for the above-chance source monitoring when pieces of conjunctions were studied in the same context. © 2002 Elsevier Science (USA). All rights reserved.

Keywords: Source monitoring; Recognition; Conjunction errors

A memory conjunction error occurs when pieces of separate past experiences become mingled and conjoined, thereby causing a rememberer to believe that the amalgam was previously experienced as a single episode. For example, if one meets Helen at an office party, and Marion at a birthday party, a conjunction error might place Helen at the birthday party despite the fact that she was never there. As another example, being introduced to new people with names such as Mrs. Chaplan and Mrs. Goldman, a conjunction error would occur if one of the two woman was later addressed as Mrs. Chapman (Reinitz, Lammers, & Cochran, 1992). Such errors can be particularly compelling to the rememberer because the information that is retrieved has indeed been experienced previously, but not necessarily in the same context or at the same point in time. From this perspective, most source-monitoring errors are a form of conjunction errors because they place previously experienced material in an incorrect context. Therefore, our goal in this study was to determine whether a memory test that placed a premium on accessing the details of the original context (i.e., source monitoring) would reduce these conjunction errors. In other words, we tested the general hypothesis that source-monitoring conditions at test might highlight the discrepancy in the two different studied contexts and thereby

* Corresponding author. Fax: +706-542-3275.
E-mail address: rlmarsh@uga.edu (R.L. Marsh).

PII: S0749-596X(02)00005-0

reduce these erroneous claims. Before describing our rationale in detail, we briefly review the experimental work that has been reported to date.

Conjunction errors occur when syllables from old words are misconjoined into new words (Underwood & Zimmerman, 1973) and when word phrases are intermingled with one another and claimed to have been experienced (Underwood, Kapelak, & Malmi, 1976). More recently, Reinitz and his colleagues have shown that features of faces studied as line drawings can be interchanged at test and the novel face recognized as old (e.g., Reinitz et al., 1992; Reinitz & Hannigan, 2001; Reinitz, Morrisey, & Demb, 1994). Busey and Tunnicliff (1999) obtained similar results when studied face stimuli were electronically morphed together and participants claimed that the morph was old (see also Bartlett & Searcy, 1998). However, the pictorial aspect of this phenomenon is not confined to faces because separately experienced pictures of scenes that are blended together at test will cause conjunction errors as well (Albert, Reinitz, Buesmans, & Gopal, 1999). In the paradigm that we will use in this article, studied compound words such as *keyboard* and *foxhole* can lead to the conjunction error of claiming that *keyhole* was experienced earlier (e.g., Jones & Jacoby, 2001; Reinitz & Demb, 1994; Reinitz, Verfaelli, & Milberg, 1996). Therefore, conjunction errors occur quite readily with all manner of stimuli that have been tested.

Although these errors appear to present themselves quite readily, there is no agreement on the underlying cognitive mechanisms which lead to their occurrence. Reinitz and his colleagues have argued that they are a consequence of intentional memory processes that are being used during binding at study and/or at retrieval (e.g., Reinitz et al., 1992). According to this account, the stimulus parts are encoded but their relational binding is poor, which can then foster erroneous binding of the parts to one another. By contrast, Jones and Jacoby (2001) have argued that conjunction errors arise primarily as a result of familiarity-based retrieval processes (see also Jones, Jacoby, & Gellis, 2001). According to their dual-process model of recognition, familiarity of the pieces of the compound should drive one toward claiming that the conjunction stimulus is old, whereas recollection should oppose this influence.

In the present study, we attempted to reduce conjunction errors by using a source-monitoring test as compared to the recognition tests that have been used previously. On a source test, participants must make judgments about the original context in which an item appeared (Johnson, Hashtroudi, & Lindsay, 1993). For example, some items might be studied visually, whereas others are heard. At test, the qualitative characteristics of memories must be inspected in order to decide whether more visual versus auditory details are present to claim that learning took place in one modality versus another. We undertook the present study because the cognitive processing associated with source monitoring has been shown in the past to reduce memory errors. For example, Lindsay and Johnson (1989) reduced the eyewitness suggestibility effect with source judgments as compared with recognition judgments. Multhaup (1995) reduced attributions of false fame in a population of older adults using Jacoby's false-fame paradigm (Jacoby & Dywan, 1990; Jacoby, Kelley, Brown, & Jasechko, 1989; Jacoby, Woloysn, & Kelley, 1989). We have also used conditions of source tests to reduce unconscious plagiarism (e.g., Marsh, Landau, & Hicks, 1997) and to curtail the amount of false recall in the DRM paradigm (Hicks & Marsh, 1999). Therefore, we believed that source monitoring might be able to reduce memory conjunction errors as well.

If the cognitive processes associated with source monitoring are more stringent because of the requirement to determine how an item was studied (Marsh et al., 1997), then conjunction errors should occur less often than they do in standard recognition. In addition, if Jones and Jacoby (2001) are correct that conjunction stimuli foster relatively few semantic recollective details, then perhaps the forced recollection of contextual details that is required by a source test may be enough to reduce these errors as compared to standard recognition. According to our analysis, the recollection involved in recognition may not affect conjunction errors, but the recollective processes involved in source monitoring could. As such, we had three specific research questions to address.

The first question was whether a source test in general would reduce conjunction errors. Theoretically, if a source test can reduce conjunction errors, then Jones and Jacoby's (2001) claim that recollection can be recruited to oppose them would receive support. On the other hand, the failure of source monitoring to reduce these errors would add empirical support to their claim that familiarity gives rise to these types of errors. The second question was whether the parent stimuli from which the conjunction stimuli were formed had to be experienced from different sources or modalities for such a reduction to occur. We reasoned that the discrepancy in context information associated with the pieces of the conjunction might be apparent enough only if they were studied in different contexts and the test required a discrimination between these specific memorial details. Jones et al. (2001) did not find a difference when the pieces of the conjunction were studied in different modalities (i.e., seen and heard) as compared to when they were studied in the same modality. That outcome suggests that a discrepancy in context will not affect recognition, but we wished to extend such a test to standard source-monitoring conditions.

A third question that we needed to address was what potential combinations of sources might lead to a reduction in conjunction errors. Various forms of source monitoring are not equivalent to one another (e.g., Gruppuso, Lindsay, & Kelley, 1997; Lindsay, Johnson, & Kwon, 1991). In fact, a particular source (e.g., visually studied items) might not behave in the same fashion when tested in combination with one particular source versus another (Marsh & Hicks, 1998). In other words, sources vary in their relative diagnosticities to one another, as well as in the absolute sense that some sources will generally lead to better or worse memory for that particular context. Therefore, we appealed to the idea that there are three general types of source combinations (e.g., Brown, Jones, & Davis, 1995). When an item is generated, elaborated on, or a decision is made about it at study, then more internal cognitive operations are assumed to be associated with it (Johnson, Raye, Foley, & Foley, 1981). By contrast, when an item is presented intact with no orienting task, then it is believed to have primarily external perceptual details associated with it. Therefore, whether source monitoring can affect the incidence of conjunction errors is likely to be a function of the combination of internal versus external sources that are tested (e.g., Hicks & Marsh, 1999, 2001). We tested one type of each of the external–external and internal–internal source-monitoring conditions. We also tested two types of internal–external source-monitoring conditions (i.e., reality monitoring). Consequently, the four experiments presented here differed from one another only in the particular combination of sources tested.

In each experiment, four conditions were tested. Two conditions tested standard recognition and two conditions tested source memory. For each type of test, the parent stimuli of conjunction lures either came from the same source or came from different sources. Thus, the two factors were crossed orthogonally in a between-subjects design. To reiterate, the source versus recognition test manipulation addressed whether source monitoring in general would reduce conjunction errors. The manipulation of presenting the parents from the same versus different sources addressed whether source tests would only be able to reduce these errors when a discrepancy in context information was maximally evident because the two pieces of the compound were learned differently (i.e., in different contexts). Across experiments, the particular combination of sources was manipulated to assess whether the utility of a source test in reducing conjunction errors varied as a function of the type of source-monitoring processes that were recruited.

Experiment 1

This first experiment tested external–external source monitoring in which items were learned either by being heard spoken aloud or by reading them from a computer monitor. In the source-monitoring conditions, this seen–heard combination of sources required that participants distinguish between visual and auditory details associated with the test items. As control conditions, standard item recognition tests were administered for each of the same and different source conditions. Thus, type of test (recognition vs source) was crossed orthogonally with the type of conjunction lure (formed from parents that were learned in the same vs different modalities).

Method

Participants. Eighty University of Georgia undergraduates volunteered in exchange for partial credit toward a course requirement. Based on their order of arrival at the laboratory, participants were assigned to the four conditions on a rotating basis until 20 volunteers had been tested in each condition. Two participants were replaced because it was obvious to us that they did not understand what was being asked of them.

Materials and procedure. Ninety compound word triplets were created, which meant that the total size of the word pool was 270. We will denote items within a given triplet by the consecutive letters A–C. The A and B compound words were the parent stimuli (e.g., *deadbolt* and *neckline*) that could be recombined into the C compound word, which was the conjunction lure (i.e., *deadline*). For a given participant, 40 triplets were randomly selected anew from the pool of 90 triplets and the A and B items were studied. Consequently, each participant studied a list of 80 items. Half of these were seen intact and half were read aloud by the experimenter. In the same source conditions, both A and B compound words from a given triplet were presented in the same modality (i.e., both seen or both heard). Once the software had selected 40 triplets, half were randomly designated to be presented from the heard source and half from the seen source. In the different source conditions, all of the A items and the B items for each of the triplets were presented from different sources. Half of the A items came from the heard source and half were seen, which meant that the same was true of the B items. The proximity of one parent stimulus to its mate of the triplet (i.e., the A and B terms) was not controlled, but rather, was left to the randomization procedure. At test, the two studied items (A and B items) from a random 20 triplets were tested as old words (half were seen and half were heard). Therefore, 40 words were old on the test. For the remaining 20 triplets, the compound C word was tested as a conjunction lure. To balance the number of targets and lures, an additional 20 C words were randomly chosen from among the 50 triplets in the word pool that were not being used for that particular participant, and thus, these served as brand new items. However, across participants the conjunction lures and unrelated lures always came from the identical pool of 90 items (i.e., the C terms). On the 80-item test every participant received the same numbers of items: 40 words were old, 20 were conjunctions, and 20 were brand new.

Two computers were linked together via their serial ports. One machine was the participant's machine and one was the experimenter's. The experimenter's machine was located behind the participant, who could not see it. During learning, words were presented randomly to one of the two computers at a 5-s rate. If a word appeared on the participant's monitor, it was to be read silently. If the word appeared on the experimenter's machine, then it was spoken out loud to the participant. Our instructions indicated that all words should be learned for an unspecified memory test. At test, all of the words from the three classes of items were randomly intermingled anew. A brief fixation point and warning tone appeared for 250 ms followed by the test item in the center of the participant's monitor.

Depending on their assigned test condition, participants in the recognition conditions pressed one of two labeled keys to indicate whether the item was old or new. Participants in the source-monitoring conditions pressed one of three labeled keys to indicate whether the item was seen, was heard, or was new. As described before, the same source versus different source conditions designated whether the software presented the two parents (A and B compounds) from the same or from different sources during study. A 5-min distractor period was introduced in between the end of the study list and the commencement of the test. During this time, participants worked on solving mazes with paper and pencil. Instructions for all phases of the experiment were read by the participant from the computer monitor and then verbally reiterated by the experimenter.

Results and discussion

Unless specified otherwise by a p value, statistical significance does not exceed chance by the conventional 5% throughout this article. Following Jones and Jacoby (2001), we analyzed only the hit rates on the old, studied items and false alarms to the conjunction lures that were labeled old. Both of these measures have been corrected for the baseline false alarm rate to the new items that were unrelated to any studied items. Rather than using the simpler two-high threshold model of $H - FA$, we used the one-high threshold model of $(H - FA)/(1 - FA)$. However, all methods of correction yielded very similar results. Both of these dependent variables were then analyzed with a 2×2 analysis of variance (ANOVA) model with the two between-subjects factors of test type (recognition vs source) and source of the parent items (same vs different). In the source conditions, the hit rate has been inferred from studied items claimed to be old regardless of the source to which they were attributed (e.g., Bink, Marsh, & Hicks, 1999; Hoffman, 1997). The average proportions that address the basic questions are summarized as the last two columns of Table 1. The preceding three columns in that table contain the uncorrected data that were used to derive the corrected rates on which the analyses were conducted.

The corrected hit rate under source monitoring was equivalent to standard recognition testing, $F(1,76) < 1.0$, $MSE = .01$, n.s. Oddly, the hit rate was higher for those conditions where the parents

Table 1
Proportion of claims that items were studied in uncorrected form and corrected for baseline false alarm rates in each of the conditions for Experiments 1–4

Experiment and condition	Uncorrected rates			Corrected rates	
	Old items	New items	Conjunctions	Old items	Conjunctions
Experiment 1					
Recognition					
Same source	.72	.15	.35	.66	.22
Different source	.75	.12	.33	.72	.24
Source monitoring					
Same source	.76	.25	.45	.68	.25
Different source	.80	.18	.42	.75	.27
Experiment 2					
Recognition					
Same source	.86	.06	.26	.85	.21
Different source	.90	.05	.24	.89	.20
Source monitoring					
Same source	.91	.08	.31	.91	.25
Different source	.90	.08	.27	.89	.21
Experiment 3					
Recognition					
Same source	.71	.20	.43	.62	.28
Different source	.71	.19	.40	.64	.26
Source monitoring					
Same source	.80	.32	.64	.70	.45
Different source	.76	.24	.46	.68	.28
Experiment 4					
Recognition					
Same source	.72	.15	.43	.67	.33
Different source	.67	.13	.38	.63	.28
Source Monitoring					
Same source	.78	.22	.56	.71	.44
Different source	.73	.15	.46	.68	.37

were studied from different sources than when they were presented in the same source, $F(1,76) = 4.98$, $MSE = .01$. Because the same versus different source manipulations of the conjunction stimuli should have no bearing on studied items, we attribute this meaningless effect to variation in sampling. This same pattern is not evident in the remaining experiments. The two factors did not interact. The incidence of conjunction errors was not affected by the type of test (source vs recognition) or the source of the parent items (same vs different). Moreover these factors did not interact, all $F(1,76)$'s < 1, n.s. Therefore, contrary to our predictions, conjunction errors remained stable under conditions of source monitoring, and conjunctions composed of pieces that were experienced in different contexts were not able to be avoided better under source monitoring as compared to recognition. These results support the contention that conjunctions evoke high levels of familiarity, and this, in turn, causes them to be erroneously labeled old at very high rates.

In subsidiary analyses of the two source-monitoring conditions, we do know what participants claimed to be true about how they studied the conjunctions. For these claims, the data do not need to be corrected or otherwise adjusted. In the same source group, more conjunctions were labeled heard (26%) than were labeled seen (19%). The same was true in the different source group, with 25% being labeled heard and only 17% being called seen. Pooling over the two conditions, more conjunctions were labeled heard, $t(39) = 2.53$. These claims suggest that participants may believe that the heard source is less diagnostic or that they have weaker memory for that source. However, in the same source condition we can also calculate how accurate participants were to label the conjunctions as having come from the source in which the parents were truly experienced (e.g., call the conjunction heard when both parents were heard). Correct source attributions (30%) were significantly greater than misattributions (15%), $t(19) = 3.74$. Therefore, these results suggest that participants are inspecting these conjunction stimuli for details of the original source, and when there is consistent information to find they more often label it the context in which the parents were studied. However, when the source information conflicts, or when they are not sure, there was a general propensity to label them heard.

We now briefly consider the overall source-monitoring performance pooling over the same and different source conditions. New items were best identified ($M = .79$), heard less so ($M = .65$), and worst identification was for the seen items ($M = .59$), $F(2, 78) = 14.18$, $MSE = .03$. Misattributions of new items, like the conjunction errors, were more often attributed to the heard source ($M = .14$) than the seen source ($M = .08$), $t(39) = 2.97$. That outcome supports the interpretation that participants may believe that the heard source is weaker, or less diagnostic, despite the fact that they had better source-monitoring accuracy for those items than the seen items.

By way of summary, with a combination of seen and heard sources we found no evidence that source monitoring could reduce conjunction errors. That result certainly supports Jones and Jacoby's (2001) contention that conjunction errors are driven largely by familiarity-based processes because a source task ostensibly requires more stringent cognitive processes and the use of recollective details. However, the fact that participants can find accurate source information in the conjunctions when the parents were presented from the same source suggests that they are using recollective processes. If one were to assume (for the sake of argument) that an old–new detection judgment preceded a source judgment, then perhaps the high familiarity value of the conjunctions leads to a decision that an item was old prior to a discrimination of source. We will return to these considerations once we have gathered more evidence in the subsequent experiments. We also found that having the parents of the conjunctions be experienced in different contexts did not reduce conjunction errors, even under source-monitoring test conditions (cf. Jones et al., 2001). Perhaps the combination of two external sources that are both perceptual in nature does not help very much because the two sources can be easily confused with one another.

Experiment 2

In Experiment 1, very little evidence was found that source monitoring at test was effective at inoculating people against claiming that conjunctions were old. However, Experiment 1 tested only conditions of external–external source monitoring. In the next experiment, internal–internal source monitoring was tested. An internal source leaves records of cognitive operations and these memorial details tend to be quite diagnostic (i.e., "strong") about the origin of a memory (Brown et al., 1995; Johnson et al., 1981). Therefore, we expected to find much higher hit rates, better source accuracy, and consequently, fewer conjunction errors. Most importantly, with sources whose memorial characteristics lead to better source monitoring, they might foster a reduction in conjunction errors, even if it is only for the different source condition in which parent stimuli are studied in different contexts. The two internal sources involved participants judging how easily an item was placed into a larger, superordinate category and judging how easily it was to form a mental image of the referent.

Method

Participants. Eighty University of Georgia undergraduates volunteered in exchange for partial credit toward a course requirement. Each participant was tested individually in sessions that lasted approximately 25 min. The procedure of assignment to the four between-subjects conditions was identical to Experiment 1.

Materials and procedure. The same word pool of 90 triplets was used again, as well as all of the same randomization procedures that were described under the Method section of Experiment 1. The only procedural aspects that were different were the two orienting instructions given to participants during study and the instructions for the source conditions at test. During study, all of the items appeared in the center of the participant's computer monitor. Each studied compound word had one of two queries below it: *How Imageable*? or *How Categorizable?*. Participants had been instructed to rate items on a 1–7 Likert scale for how easily they could form a mental image of the referent under imagability instructions. By contrast, they were asked to rate on the same scale how easily they could place the item into some superordinate category under categorizability instructions. Both of these orienting instructions result in cognitive operations being stored in memory. During the test, participants receiving source instructions pressed one of three keys to indicate whether they made an imagability judgment, whether they made a categorizability judgment, or whether the item was new. In all other respects including the same versus different source manipulations of the parent stimuli used for the conjunction lures, this experiment was identical to Experiment 1. Recall that a 5-min distractor was included between study and test.

Results and discussion

The data are summarized in Table 1. As predicted, the two internal sources led to a markedly higher hit rate in this experiment as compared with the two external sources tested in Experiment 1. In the 2×2 ANOVA on the corrected hit rates, neither the main effect of test type (recognition vs source) nor the source manipulation concerning the parents of conjunctions (same vs different) was statistically significant, larger $F(1,76) = 2.26$, $MSE = .01$, n.s. In addition, there was no significant interaction. For the sake of brevity, these same null outcomes were observed in the ANOVA model for the conjunction errors themselves. Thus, source monitoring led to comparable outcomes as testing old–new recognition. Simply put, the incidence of conjunction errors was not responsive to source testing and not responsive to compounds whose pieces were originally experienced in the same versus different contexts (even under source testing instructions).

We conducted the same subsidiary analyses as performed in Experiment 1. Participants in this experiment could successfully identify the qualitative characteristics that were associated with the parents of conjunctions. For those in the same source condition that took a source test, more conjunctions were labeled correctly to the source of the parents (19%) than were misattributed to the other source (11%), $t(19) = 4.36$. This result shows that participants were attempting to inspect conjunction items and they found evidence for the type of orienting task performed on the parent items. Moreover, there was no bias to attribute items to the categorized or to the imageable sources because in the same source condition, source claims averaged 17% for categorizability and 14% for imagability, whereas the direction of these numeric differences was opposite in the different source conditions, at 11 and 16%, respectively. Therefore, unlike the bias to label conjunctions heard in the previous experiment, no bias was present in this experiment. Nevertheless, participants in both experiments identified more conjunctions with the source of their parent items in the same source condition.

The data for overall source-monitoring performance (pooled over the two source conditions) showed that performance was best on the new items (M $= .92$), worse for the items judged for imagability ($M = .78$), and worst for the categorized items ($M = .60$), $F(2,78) = 87.10$, $MSE = .01$. However, there was no difference in erroneously assigning new items to either old source (both $M = .04$), $t(36) < 1.0$. On the whole, source monitoring was generally quite good, and therefore, source-monitoring test conditions should have been able to reduce conjunction errors if that effect was to be found with this particular combination of sources.

Experiment 3

Together, the first two experiments certainly do not constitute a resounding endorsement for the ability of source-monitoring processes to reduce conjunction errors despite the fact that when these cognitive processes are used in other situations they can reduce other types of memory errors. In all fairness, however, in using source monitoring to reduce errors in the DRM paradigm (Hicks & Marsh, 1999), we were not successful when the combinations of sources were highly similar to one another (e.g., male versus female voices). Success at reducing false recall was obtained under reality-monitoring conditions only when the two sources were very different from one another. Therefore, in this next experiment we tested reality monitoring in which an internal–external combination of sources was used. Half of the items were generated from anagrams and half were spoken aloud by the experimenter. The anagrams should primarily contain cognitive operations from solving them during study and the heard items should primarily contain auditory details as was argued earlier under Experiment 1. These two sources should vary quite a bit in their discriminability from one another, and perhaps under these conditions conjunction errors will be more malleable to source-monitoring test manipulations.

Method

Participants. Eighty University of Georgia undergraduates volunteered in exchange for partial credit toward a course research requirement. Each participant was tested individually in sessions that lasted approximately 30 min. Three participants were replaced for obvious failure to understand what was being asked of them. Assignment to four between-subjects conditions was performed in the same manner as before.

Procedure. The general procedures that were used in the previous two experiments were used again, except in the following respects. For all participants, studied compound words either were solved as an anagram or were spoken aloud by the experimenter. The same experimental setup involving two connected computers that was used in Experiment 1 was used again. Heard items were presented to the experimenter's computer for him to read aloud. Rather than items seen intact as in Experiment 1, anagrams were written to the participant's monitor. The anagram manipulation was identical to our previous uses of it (e.g., Hicks & Marsh, 1999, 2001). As randomly generated anew online, an anagram of a compound word had two letters interchanged under the constraints that the two letters be nonadjacent, not be the same letter, and not be in the first letter position of the word. A caret (∧) was placed underneath each of the interchanged letters and participants were asked to mentally interchange the two letters to discover the identity of the words, which were then spoken aloud. At test, recognition versus source tests were administered, and within these test conditions the parents of conjunctions were manipulated as having been experienced from the same or different sources. All of the randomization measures that were in force in Experiments 1 and 2 were operative here as well.

Results and discussion

The data are summarized in Table 1, and 2 × 2 ANOVA models were tested on the corrected hit rates and conjunction errors separately. The hit rate was higher under source testing as compared to recognition, $F(1,76) = 4.21$, $MSE = .02$. However, it did not vary as a function of the whether parents were from the same source or not, and there was no interaction, both $F(1,76)$'s < 1, n.s. The higher hit rate under source instructions may reflect the fact that the increased discriminability of the two sources has somehow affected the old–new detection component of source monitoring, which is an issue that we revisit later.

Higher rates of conjunction errors occurred in the source conditions, resulting in a main effect of test type, $F(1,76) = 5.28$, $MSE = .04$, but a marginally significant two-way interaction was also present, $F(1,76) = 3.25$, $MSE = .04$, $p = .07$. As can be seen in Table 1, the conjunction error rate was quite large in the same source condition under source-monitoring test instructions. There are two interpretations of the statistically significant difference in conjunction error rates between the two source-monitoring conditions, $t(38) = 3.63$. Participants could be particularly compelled by the match of the qualitative characteristics in a conjunction whose parents were studied from the same source or the mismatch of these memorial details in the different source condition has prevented the conjunction error rate from being even higher. In other words, it is unclear whether participants are more compelled to experience the conjunctions as real in the same source condition or whether participants are finding contradictory evidence (as we predicted) in the different source condition. Before returning to this question, we report the other effects that are likely to inform the interpretation of this outcome.

The claims about the conjunction errors under source instructions demonstrated that more were labeled heard as compared with generated in the same source condition (40% called heard and 24% called generated), whereas in the different source condition the claims were more equal but still favored the heard source (26% called heard and 21% called generated). In addition, as was found in the previous two experiments, the conjunctions in the same source condition were generally ascribed to the context from which the parents were studied. Correct attributions (47%) far outnumbered misattributions (17%), $t(19) = 10.98$. These results demonstrate that participants were examining the qualitative characteristics of test items and the conjunctions were ascribed to the source in which the parents had been experienced originally. Given that participants are ascribing sources to conjunction errors that match the source from which parents were studied (and did so consistently in the previous two experiments as well), we believe that they are finding source specifying details in the conjunctions in the same source condition and that is what has caused the extraordinarily high conjunction rate. The overarching point is that source monitoring did not reduce these errors, but rather, has appeared to exacerbate their occurrence when the qualitative characteristics of their parents match.

In terms of overall source-monitoring performance, accurate performance was best for the new items ($M = .73$) and generated items ($M = .79$) and worst for the heard items ($M = .56$), $F(2, 78) = 20.06$, $MSE = .03$. On the assumption that cognitive operations serve as a very diagnostic cue to origin, this outcome was expected. When new items were falsely believed to have been studied, they were claimed to have been heard ($M = .23$) rather than generated ($M = .05$), $t(39) = 7.98$. This result represents the it-had-to-be-you effect (Johnson et al., 1981) and accords nicely with the conjunctions being labeled heard in the absence of consistent diagnostic evidence of the parent's source in the same source condition. In other words, conjunctions were ascribed to the weaker class of items (cf. Bink et al., 1999).

Experiment 4

The results of Experiment 3 suggested that under reality-monitoring conditions conjunction errors are increased when source judgments were required at test. Because this outcome is somewhat counterintuitive, the goal of this next experiment was to replicate conceptually the results of Experiment 3 with a slightly different combination of sources. We replaced the heard source with the seen source used in Experiment 1. This change meant that the generated and seen sources used in this experiment would share some visual details with one another and, therefore, make source monitoring slightly more challenging. In other words, we attempted to make the two sources more similar, albeit still testing reality monitoring.

Method

Participants. Eighty University of Georgia undergraduates volunteered in exchange for partial credit toward a course research requirement. Each participant was tested individually in sessions that lasted approximately 30 min. One participant with very low performance was replaced in order to maintain consistency with the previous experiments.

Procedure. The general procedure followed in Experiment 3 was used for this experiment except in the following respects. The heard source used previously was replaced with a seen source in which participants silently read intact items from the center of the computer monitor. Anagrams were also presented in the center of the monitor and were constructed and solved according to the rules described earlier. At test, participants in the two source test conditions made judgments by pressing one of three labeled keys to indicate whether the item was generated, was seen, or was new. In all other respects, this experiment conformed to those that have been described previously.

Results and discussion

In the 2×2 ANOVA, the corrected hit rate was again numerically higher in the two source conditions than in the two recognition conditions, but not statistically so, $F(1, 76) = 2.57$, $MSE = .02$, n.s. In addition, the same versus different source manipulation neither affected the corrected hit rate nor was there an interaction, both $F(1, 76)$'s < 2.00, n.s. As in Experiment 3, more conjunction errors were produced by source testing than recognition testing, $F(1, 76) = 7.79$, $MSE = .02$, and conjunctions formed from the same source were called old more than those formed from different sources, but this failed to reach significance in the ANOVA model, $F(1, 76) = 2.51$, $MSE = .02$, n.s. The interaction was not significant. Therefore, the important result was that source monitoring increased rather than reduced source errors.

In the previous experiment, more conjunctions were labeled heard overall, and in this experiment they were labeled seen. In the same source condition, 33% were called seen but only 24% were called generated, $t(19) = 2.50$. In the different source condition, 28% were called seen and 19% were called generated, $t(19) = 2.03$. Therefore, once again, participants attributed the conjunctions to the weaker, less diagnostic source (seen). Nevertheless, participants were attempting to analyze the qualitative characteristics associated with the memories because in the same source condition they were well above chance at labeling the conjunctions the source in which the parents had been studied insofar as correct attributions (38%) outnumbered misattributions (18%), $t(19) = 4.96$. By way of summary, the reality monitoring conditions tested in this experiment do not support the notion that conjunction errors can be reduced with a source test. In fact, conjunction errors increased under source monitoring and slightly more so when the parents of the conjunctions were studied in the same context.

In terms of overall source monitoring, the results replicated those of Experiment 3. Source monitoring was best for new ($M = .82$) and generated ($M = .78$) items and worse for seen items ($M = .48$), $F(2,78) = 17.49$, $MSE = .03$. The it-had-to-be-you effect was present in the more frequent false alarm of new items to the seen source ($M = .15$) as compared to the generated source ($M = .04$), $t(39) = 6.39$. Thus, the generated source was the more diagnostic source.

General discussion

We undertook this investigation with three interrelated research questions. The first concerned whether the cognitive processes recruited to perform source monitoring would reduce conjunction errors as compared to the processes used in standard recognition. The answer to that question appears to be that they cannot. The second question addressed whether the context information associated with the pieces of the conjunctions had to be different for a such reduction to be found. The answer to that question is more equivocal. The same versus different source manipulation had no effect on conjunction rates in Experiments 1 and 2 with external–external and internal–internal source monitoring, respectively. But in Experiment 3 (and numerically in Experiment 4), which used reality monitoring, conjunctions whose pieces were experienced in the same context were found to be quite compelling and resulted in higher error rates. In some circumstances participants can be driven by the consistency of the information that they found in conjunctions whose pieces came from the same source. Therefore, conjunction errors can be influenced in some limited cases by how their parents are learned. Our third research question concerned whether the combination of sources tested across experiments would matter, and of course, the answer is that they do affect the rate of conjunction errors.

On the whole, the current data support Jones and Jacoby's (2001) theoretical proposition that conjunctions are called old because of the high level of familiarity that they evoke. Because undifferentiated familiarity cannot be easily used as a basis for a source judgment (e.g., Bink et al., 1999), source monitoring during test is not effective at inoculating people against these errors. Moreover, the high degree of familiarity in the conjunctions may cause people to believe that they are old, and then once this assessment is made, a source is chosen for the conjunctions after the old–new decision has been made largely on the basis of familiarity. We are not suggesting that source judgments are composed of an old–new detection component that obligatorily precedes a source discrimination. Rather, we are suggesting that in the case of conjunctions, their high degree of familiarity influences old–new detection even under source test instructions. However, there must also be another effect operating because source tests exacerbated the incidence of conjunction errors in both Experiments 3 and 4.

To explain that outcome, consider the relative diagnosticities of the sources used in each experiment. In Experiments 1 and 2, the two sources were relatively similar to one another. By contrast, Experiments 3 and 4 tested sources that should have been very different from one another, with one source being "stronger" and one being "weaker." In this case, the qualitative characteristics were quite identifiable, and when they matched in a conjunction stimulus, participants found that to be very compelling evidence that it was experienced. However, just as in Experiment 1, participants in Experiments 3 and 4 also had a bias to label conjunctions to the weaker, external source. Therefore, one general theme may be that when candidate sources differ in their absolute diagnosticity, any bias to label items to a particular source could exacerbate the rate of conjunction errors because they can be attributed to the source with the weaker diagnosticity.

If this analysis is true, then it paints a somewhat bleak picture for reducing these errors in everyday life. For example, if the veracity or plausibility of a memory conjunction is being evaluated and the candidate contexts differ in their relative diagnosticity, then the conjunction is likely to be attributed to the less diagnostic source. In this case, conjunctions are likely to occur more frequently, and perhaps with more confidence, because the rememberer can assign an erroneous recollective detail to the conjunction. The good news is that if the contextual details do not match, as was the case in the different source conditions of Experiments 3 and 4, then a conjunction error is probably no more likely than when an erroneous memory is evaluated less stringently as is probably occurring in recognition testing. The foregoing analysis highlights one cost associated with source monitoring. That cost is that when source-specifying information can be found in a memory, and when one context among several alternatives seems plausible, people will often assume that the event was experienced in that context. In an attempt to reduce false recognition in the DRM paradigm, we found a similar effect (Hicks & Marsh, 2001).

However, there is also an alternative theoretical account for why source monitoring can sometimes increase the incidence of conjunction errors. Incomplete source information in the conjunction lures could be compelling enough to increase the rate of these errors under source monitoring in Experiments 3 and 4. In other words, incomplete source information can be used to render a source judgment (e.g., Dodson, Holland, & Shimamura, 1998; Hicks, Marsh, & Ritschel, 2002; Gruppuso et al., 1997). The term "incomplete information" only means that the quality of the details that are retrieved are vague and not particularly vivid (Hicks et al., 2002). The more of these partial details that match a particular context, the higher the probability that conjunction will be ascribed to that source. These ideas are consistent with the fact that participants in all four experiments attributed same-source conjunctions to the source of their parents. And arguably, that result demands an explanation that must consider information beyond undifferentiated familiarity.

Therefore, it seems that partial information can be used to accurately determine how an item was encountered during encoding. However, that same information in its partial form is not diagnostic enough to reject highly plausible lures such as conjunctions as being new. In other words, partial source information can used in source discrimination, but it does not seem to be as useful in the old–new detection component of source-monitoring processes, at least with the paradigm tested here. To our knowledge, no one has investigated directly how the old–new detection component of source monitoring is carried out. Rather, the relationship between old–new detection as it is carried out in standard recognition and the old–new detection component of source monitoring has been described as variable (e.g., Johnson et al., 1993). If lawful relationships are to be found, then perhaps our conjectures about partial information may be one way to proceed in studying these issues.

The results from this study suggest that conjunctions are evaluated at test to be consistent enough with past experience to be called old at similar rates under recognition and source testing. A recognition test decision may require only familiarity in the absence of all-or-none recollection, but partial details of the context are probably nevertheless being recollected that act like familiarity. However, on a source test, those partial details support accurate context judgments (see also Qin, Raye, Johnson, & Mitchell, 2001; Slotnick, Klein, Dodson, & Shimamura, 2000). The ironic effects of higher conjunction errors under some forms of source monitoring are really not that ironic at all. Rather, contextual details that vary in their clarity are serving as the basis for judgmental processes to claim that the conjunction was experienced earlier. This alternative account assumes that partial recollective details can act like familiarity but those same details that are useful in assigning a source are not powerful or diagnostic enough to be useful to an editing process that could potentially reduce conjunction errors.

Acknowledgments

We thank L. Patrick Shippey, Jeremy Newton, Matthew Serafin, Christy Anderson, and LeRoya Chester for their dedicated help in collecting the data. We also thank Mark Reinitz, Bob Greene, and an anonymous reviewer for helpful comments on previous versions of this paper.

References

Albert, W., Reinitz, M. T., Buesmans, J., & Gopal, S. (1999). Role of attention in spatial learning during simulated route navigation. *Environment and Planning A, 31*, 1459–1472.

Bartlett, J. C., & Searcy, J. H. (1998). *Configural processing and binding in recognition of faces*. Presented at the 39th Annual Meeting of the Psychonomic Society, Dallas, TX.

Bink, M. L., Marsh, R. L., & Hicks, J. L. (1999). An alternative conceptualization to memory "strength" in reality monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 804–809.

Brown, A. S., Jones, E. M., & Davis, T. L. (1995). Age differences in conversational source monitoring. *Psychology & Aging, 10*, 111–122.

Busey, T. A., & Tunnicliff, J. L. (1999). Accounts of blending, distinctiveness, and typicality in the false recognition of faces. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1210–1235.

Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 24, 1121–1136.

Gruppuso, V., Lindsay, D. S., & Kelley, C. M. (1997). The process-dissociation procedure and similarity: defining and estimating recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 259–278.

Hicks, J. L., & Marsh, R. L. (1999). Attempts to reduce the incidence of false recall with source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1195–1209.

Hicks, J. L., & Marsh, R. L. (2001). False recognition occurs more frequently during source identification than during old–new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*.

Hicks, J. L., Marsh, R. L., & Ritschel, L. (2002). The role of recollection and partial information in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 503–508.

Hoffman, H. G. (1997). Role of memory strength in reality monitoring decisions: evidence from source attribution biases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 371–383.

Jacoby, L. L., & Dywan, J. (1990). Effects of aging on source monitoring: differences in susceptibility to false fame. *Psychology & Aging, 5*, 379–387.

Jacoby, L. L., Kelley, C. M., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: limits on the ability to avoid unconscious influences of the past. *Journal of Personality & Social Psychology, 56*, 326–338.

Jacoby, L. L., Wolosyn, V., & Kelley, C. (1989). Becoming famous without being recognized: unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General, 118*, 115–125.

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

Johnson, M. K., Raye, C. L., Foley, H. J., & Foley, M. A. (1981). Cognitive operations and decision bias in reality monitoring. *American Journal of Psychology, 94*, 37–64.

Jones, T. C., & Jacoby, L. L. (2001). Feature and conjunction errors in recognition memory: evidence for dual-process theory. *Journal of Memory and Language, 45*, 82–102.

Jones, T. C., Jacoby, L. L., & Gellis, L. A. (2001). Cross-modal feature and conjunction errors in recognition memory. *Journal of Memory and Language, 44*, 131–152.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition, 17*, 349–358.

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

Marsh, R. L., & Hicks, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 1137–1151.

Marsh, R. L., Landau, J. D., & Hicks, J. L. (1997). The contribution of inadequate source monitoring during idea generation to unconscious plagiarism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 886–897.

Multhaup, K. S. (1995). Aging, source, and decision criteria: when false fame errors do and do not occur. *Psychology & Aging, 10*, 492–497.

Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source ROCs are (typically) curvilinear: comment on Yonelinas (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1110–1115.

Reinitz, M. T., & Demb, J. (1994). Implicit and explicit memory for compound words. *Memory & Cognition, 22*, 687–694.

Reinitz, M. T., & Hannigan, S. L. (2001). Effects of simultaneous stimulus presentation and attention switching on memory conjunction errors. *Journal of Memory and Language, 44*, 206–219.

Reinitz, M. T., Lammers, W. J., & Cochran, B. P. (1992). Memory conjunction errors: miscombination of stored features can produce illusions of memory. *Memory & Cognition, 20*, 1–11.

Reinitz, M. T., Morrisey, J., & Demb, J. B. (1994). The role of attention in face encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 161–168.

Reinitz, M. T., Verfaelli, M., & Milberg, W. (1996). Memory conjunction errors in normal and amnesic subjects. *Journal of Memory and Language, 35*, 286–299.

Slotnick, S. D., Klein, S. A., Dodson, C. S., & Shimamura, A. P. (2000). An analysis of signal detection and threshold models of source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1499–1517.

Underwood, B. J., Kapelak, S. M., & Malmi, R. A. (1976). Integration of discrete verbal units in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 293–300.

Underwood, B. J., & Zimmerman, J. (1973). The syllable as a source of error in multisyllable word recognition. *Journal of Verbal Learning and Verbal Behavior, 12*, 701–706.