

The influences of partner accuracy and partner memory ability on social false memories

Katya T. Numbers · Michelle L. Meade · Vladimir A. Perga

Published online: 18 July 2014
© Psychonomic Society, Inc. 2014

Abstract In this study, we examined whether increasing the proportion of false information suggested by a confederate would influence the magnitude of socially introduced false memories in the social contagion paradigm Roediger, Meade, & Bergman (Psychonomic Bulletin & Review 8:365–371, 2001). One participant and one confederate collaboratively recalled items from previously studied household scenes. During collaboration, the confederate interjected 0 %, 33 %, 66 %, or 100 % false items. On subsequent individual-recall tests across three experiments, participants were just as likely to incorporate misleading suggestions from a partner who was mostly accurate (33 % incorrect) as they were from a partner who was not at all accurate (100 % incorrect). Even when participants witnessed firsthand that their partner had a very poor memory on a related memory task, they were still as likely to incorporate the confederate’s entirely misleading suggestions on subsequent recall and recognition tests (Exp. 2). Only when participants witnessed firsthand that their partner had a very poor memory on a practice test of the experimental task itself were they able to reduce false memory, and this reduction occurred selectively on a subsequent individual recognition test (Exp. 3). These data demonstrate that participants do not always consider their partners’ memory ability when working on collaborative memory tasks.

Keywords Group memory · False memory

Despite common misconceptions of memorial accuracy, research suggests that reconstructions of the past are often

unreliable. We forget information, confuse aspects of different events, and critically, we are influenced by what other people say. Research on social false memory suggests that when another person recollects inaccurate details about a shared event, individuals often incorporate those inaccurate suggestions into their own memories (e.g., Roediger, Meade, & Bergman, 2001; Wright, Self, & Justice, 2000; see Harris, Paterson, & Kemp, 2008; Hirst & Echterhoff, 2012; and Rajaram, 2011, for reviews). Importantly, research on social false memory typically embeds only a small proportion of erroneous details into the total suggestions made by one’s partner. The implicit assumption is that memories are fairly accurate, and so to avoid participant suspicion, the experimenter must slip in relatively few incongruent items. But is this really the case? In the present study, we examined how the proportion of inaccurate items suggested by one’s partner influences the likelihood that individuals will falsely remember inaccurate partner suggestions.

Social false-memory paradigms that utilize low proportions of inaccurate suggestions include the memory conformity paradigm and the social contagion paradigm (see, too, the related misinformation paradigm; e.g., Loftus, Miller, & Burns, 1978). In the memory conformity paradigm, participants are presented with images of an event (e.g., 21 pictures depicting a wallet theft; Wright et al., 2000). Of these images, 20 are identical between participants, and one critical image differs between participants (e.g., an accomplice was present or not present at the time of the robbery). In the social contagion paradigm (Roediger et al., 2001), participants study images of household scenes and then recall the scenes in collaboration with a confederate, who introduces both correct and incorrect items. Specifically, the confederate “recalls” 36 items in collaboration with the participant, and only six (or 17 %) are incorrect. Research in both paradigms has demonstrated that participants are likely to incorporate these small percentages of suggested erroneous items into their

K. T. Numbers · M. L. Meade (✉) · V. A. Perga
Department of Psychology, Montana State University,
P.O. Box 173440, Bozeman, MT 59717, USA
e-mail: mlmeade@montana.edu

subsequent individual recall and/or recognition tests (e.g., Allan & Gabbert, 2008; Bodner, Musch, & Azad, 2009; Davis & Meade, 2013; Gabbert, Memon, & Allan, 2003; Gabbert, Memon, Allan, & Wright, 2004; Huff, Davis, & Meade, 2013; Skagerberg & Wright, 2008; Wright et al., 2000). Of interest in the present study was whether or not increasing the proportion of false items suggested by the confederate would modulate the social contagion effect.

Higher proportions of false items should decrease the magnitude of social false memories, because hearing primarily (or entirely) incorrect items may alert participants to discrepancies between their original learning and the confederate's suggestions and/or to the confederate's inferred credibility (see Tousignant, Hall, & Loftus, 1986). According to the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), additional processing of discrepancy during encoding should lead to greater discriminability between items that had appeared in the original event and items that had appeared in the postevent misinformation. The source-monitoring framework also suggests that, if participants notice the confederate's errors, they might employ a more stringent decision criterion when attributing the confederate's suggestions to the study episode. Greater discriminability and/or a stricter decision criterion predict lower levels of social false memories from increasingly inaccurate confederates.

Previous research has been consistent with the idea that when participants in social memory paradigms are made aware of misleading items, the magnitude of the effect is reduced (Echterhoff, Hirst, & Hussy, 2005; Greene, Flynn, & Loftus, 1982; Paterson, Kemp, & Ng, 2011; although see Muller & Hirst, 2010, for an exception). It is important to note, however, that studies demonstrating reductions in false memories have typically employed explicit, experimenter-issued warnings that directed attention to inaccurate partner suggestions, and they also have held confederate accuracy constant across conditions (i.e., the confederate suggested the same errant items in both the warning and no-warning conditions; see, e.g., Echterhoff, Groll, & Hirst, 2007; Meade & Roediger, 2002). That is, participants' reductions in false memory were the result of experimenter-issued directives of partner accuracy, rather than of partner accuracy per se. False memories can be diminished when participants are made aware, via explicit warnings, that items suggested by their partner may be inaccurate, even when the warnings do not correspond to actual partner accuracy.

In the present experiments, we provided no warning or explicit instructions to participants, but rather examined whether higher proportions of false items alone would reduce the social contagion effect. Put another way, in the present study we manipulated actual confederate accuracy, rather than perceived (or experimenter-issued warnings about) confederate accuracy. Arguably, accuracy is one of the most important characteristics to consider when adopting information from

another person. The present study tested the assumption that individuals spontaneously consider the accuracy of their partner's contributions when working together on a memory task.

To our knowledge, just one prior study has examined the effects of partner accuracy on memory performance. Jaeger, Lauris, Selmecky, and Dobbins (2012) presented participants with responses from a fictitious previous participant who was either reliable or unreliable (75 % vs. 50 % correct in Exp. 1; 75 % vs. 25 % correct in Exp. 2). Interestingly, participants treated even the unreliable sources as being generally informative, especially when their own confidence was low. Jaeger et al. concluded that there are strong metacognitive assumptions that others will provide useful, valid information on memory tests. The paradigm utilized by Jaeger et al.—a focus on veridical recognition using an implied confederate—is quite different than the social contagion paradigm used in the present research, but their findings suggest that without an explicit warning, participants may not spontaneously consider partner accuracy.

However, growing evidence also indicates that participants do spontaneously consider partner characteristics and that partner characteristics can influence memory conformity. For example, Gabbert, Memon, and Wright (2006) demonstrated that the person who speaks first on a memory conformity test is less likely to conform to his or her partner's memory. Likewise Cuc, Ozuru, Manier, and Hirst (2006) demonstrated that the person in the group considered to be a dominant narrator has a disproportionately large influence on the subsequent memory of the event. In both studies, the experimenter provided no explicit instructional manipulations, but rather, the perceived or inferred characteristics of the partners influenced the magnitude of social false memories. Importantly, partner accuracy was not actually manipulated in either study, but participants nonetheless spontaneously inferred accuracy on the basis of partner characteristics (see, too, Allan, Midjord, Martin, & Gabbert, 2012, and French, Garry, & Mori, 2011).

In the present study, we systematically manipulated the accuracy of the confederate's responses in the social contagion paradigm. In Experiment 1, partner accuracy varied from 0 % incorrect (the control condition) to 33 %, 66 %, and 100 % incorrect. This last condition is especially novel, in that participants recalled an event with a partner who suggested entirely inaccurate information throughout the duration of the experiment. Of interest was whether or not participants would pick up on the confederates' inaccuracy when not explicitly instructed to do so. In Experiments 2 and 3, we examined whether partner accuracy was more influential when participants witnessed firsthand that their partner had good or poor memory on a related memory task (Exp. 2) or on the very same experimental task (Exp. 3) that they were about to complete together. Given previous evidence that subtle confederate characteristics can modulate memory conformity

effects (Cuc et al., 2006; Gabbert et al., 2006), we predicted that confederate accuracy would reduce the magnitude of misinformation adopted. However, it was also possible that participants might not notice their partner's accuracy without explicit instructions (see Harris et al., 2008; Jaeger et al., 2012).

As a final note, we were interested in both item expectancy effects and metacognitive judgments as they related to confederate accuracy. Past research had shown that high-expectancy items (typical of a given scene) are more prone to contagion than are low-expectancy items (items less typical of a given scene; see, e.g., Roediger et al., 2001). However, we hypothesized that low-expectancy items would be more influenced by confederate accuracy; that is, participants should be especially likely to discredit low-expectancy items suggested by highly inaccurate confederates.

Metacognitive judgments are also interesting in relation to confederate accuracy, because in addition to influencing items recalled, partner accuracy may influence participants' metacognitive assessments of items recalled. As such, we collected "remember"/"know" judgments in the present experiments (see Gardiner, 1988; Rajaram, 1993; Tulving, 1985). "Remember" responses indicate that participants recollect something specific about the item, whereas "know" responses indicate that participants lack specific recollected details about the item. We hypothesized that participants would be less likely to provide "remember" responses for items produced by highly inaccurate confederates.

Experiment 1

In the present experiment, we examined whether participants were as likely to incorporate misleading suggestions from a partner who was mostly accurate as from a partner who was *never* accurate.

Method

Participants The participants were 82 Montana State University undergraduates who participated for course credit. Ten of them were excluded because of suspicion, lack of English proficiency, or failure to follow instructions. The final analysis included the remaining 72 participants.

Design This experiment was based on a 2×4 mixed design. Expectancy of the contagion items (high or low expectancy) was manipulated within subjects, and the proportion of false information suggested by the confederate (0 %, 33 %, 66 %, or 100 % incorrect) was manipulated between subjects. The primary dependent variables were false recall and false recognition of the critical suggested items.

Materials The materials included six slides from Roediger et al. (2001) depicting schematically consistent household scenes (toolbox, bathroom, kitchen, bedroom, closet, and desk). Each scene contained an average of 23.8 items that were either highly typical of the scene (high expectancy) or less typical of the scene (low expectancy; as determined by Roediger et al., 2001). High- and low-expectancy items were then purposely excluded from each slide so that they could be used as contagion items. *Contagion items* refer to items suggested by the confederate that were not present in the scenes. The contagion items generated by Roediger et al. were used to construct our 33 %-incorrect condition. To generate additional contagion items for the 66 % and 100 % conditions, we ran a pilot study using the same materials as Roediger et al., to determine four new contagion items and four alternate contagion items for each scene (see the [Appendix](#)). High-expectancy contagion items were always suggested in Positions 1, 4, and 5, and low-expectancy contagion items were always suggested in Positions 2, 3, and 6. Other materials included a filler task composed of addition problems, individual recall sheets, a final individual recognition task (not reported here, due to experimenter error), and final manipulation check and demographic questionnaires, both developed locally.

Procedure One participant and confederate independently studied six household scenes for 15 s each in preparation for a later memory test. The slides were presented in the same order for every participant (toolbox, bathroom, kitchen, bedroom, closet, and desk) and were verbally labeled by the experimenter. After viewing all six scenes, the participant and confederate completed a 4-min filler task. Next, the confederate and participant took turns recalling items from each scene, one scene at a time, until each had recalled six items (12 total) for all six scenes. It was during this collaborative recall phase that the confederate interjected 0 %, 33 %, 66 %, or 100 % errant items.

Following group recall, the participant and confederate were separated into two rooms to complete their individual recall tests (the confederate did not actually complete a recall test). Participants were given a sheet of paper with the title of the scene (e.g., bedroom) and were told that they had 2 min to write down the items that they remembered from the scene. The experimenter then collected their responses and gave them the next recall sheet, until the participant had recalled all six scenes. Participants were also asked to make "remember"/"know" judgments for the items that they wrote down (Gardiner, 1988; Rajaram, 1993; Tulving, 1985). A "remember" response indicated that participants recollected something specific about the item that they were recalling, whereas a "know" response indicated that participants lacked memory of specific details about the item, but nonetheless were sure that the item had been presented.

Finally, participants were given a locally developed manipulation check that required them to indicate (on 5-point Likert scales) how credible they thought their partner was, and how accurate they thought that their partner's memory was. They were also asked whether they would choose to work with their partner again if monetary compensation (\$50) was contingent on partner accuracy (a yes–no question). All participants were debriefed.

Results

False recall The mean proportions of critical contagion items falsely recalled are shown in Table 1. Critical contagion items are defined as the two false items offered for each scene by confederates in the 33 %-incorrect condition. These same two contagion items per scene were tracked across all conditions, so as to control for both item effects and different baseline numbers of false items across conditions. The top four rows in the table show recall of the critical contagion items on the individual recall test, the next four rows illustrate “remember” judgments, and the final four rows report “know” judgments for critical contagion items. Statistical significance is set at $p < .05$ unless otherwise noted, and p values are reported only for nonsignificant effects.

A 4 (proportion incorrect: 0 %, 33 %, 66 %, 100 %) \times 2 (item expectancy: high, low) mixed factorial analysis of variance (ANOVA) conducted on critical contagion revealed significant social contagion effects that were not moderated by partner accuracy, $F(3, 68) = 9.23$, $MSE = .02$. Participants

in the 33 %-incorrect condition ($M = .38$) were more likely to falsely recall critical contagion items than were those in the control (0 %-incorrect) condition ($M = .15$), $t(34) = -4.72$, $SEM = .05$; participants in the 66 %-incorrect condition ($M = .39$) also showed more contagion than those in the control condition, $t(34) = -4.56$, $SEM = .02$; and even participants in the 100 %-incorrect condition ($M = .32$) showed more contagion than those in the control condition, $t(34) = -4.42$, $SEM = .03$. Critically, we found no significant differences for recall of contagion items between the 33 %-, 66 %-, and 100 %-incorrect conditions ($ts \leq 1.58$, $ps > .05$). Because this conclusion is based on null effects, we report also the Bayesian information criteria (BICs; Masson, 2011). The p_{BIC} represents the Bayesian estimated probability that the null hypothesis is preferred over the alternate hypothesis. The estimated probability that the null-effect model was preferred for the false recall difference between the 33 %, 66 %, and 100 % conditions was $p_{BIC} = .86$, which is considered positive evidence in support of the null hypothesis (see Raftery, 1995).

Participants were also more likely to recall high-expectancy contagion items ($M = .42$) than low-expectancy contagion items ($M = .20$), $F(1, 68) = 86.96$, $MSE = .02$. The interaction between proportion incorrect and expectancy was not significant, $F < 1.4$. Considered together, these results suggest that increasing the proportion of false information suggested by the confederate (to the point at which every item suggested by the confederate was false) had little, if any, effect on participants' false recall of the high- and low-expectancy critical contagion items.

“Remember” judgments It is possible that even though recall data did not vary as a function of partner accuracy, participants' metamemorial judgments were influenced by partner accuracy. A 4 (proportion incorrect: 0 %, 33 %, 66 %, 100 %) \times 2 (item expectancy: high, low) mixed ANOVA on “remember” responses revealed significant main effects of proportion incorrect, $F(3, 68) = 4.22$, $MSE = .02$, and item expectancy, $F(1, 68) = 19.62$, $MSE = .02$, but no significant interaction, $F(3, 68) = 1.30$, $MSE = .02$, $p > .05$, $p_{BIC} = .98$. Follow-up tests revealed that participants in the 33 %-incorrect condition ($M = .11$) gave more “remember” responses for critical contagion items than did those in the control condition ($M = .04$), $t(34) = -2.64$, $SEM = .03$. The same was true for the participants in the 66 %-incorrect condition ($M = .16$), $t(34) = -3.79$, $SEM = .03$, and the 100 %-incorrect condition ($M = .12$), $t(34) = -2.26$, $SEM = .03$. No significant differences in “remember” judgments were apparent between the 33 %-, 66 %-, and 100 %-incorrect conditions ($ts \leq 1.38$, $ps > .05$). These results suggest that participants were more likely to give “remember” responses for false items suggested by a confederate than for false items not suggested by a confederate, and for high-expectancy items than for low-expectancy items, and that the proportion of false items did not moderate these effects.

Table 1 Mean proportions of false recall, and “remember” or “know” responses for high- and low-expectancy items as a function of proportions of incorrect items suggested by the confederate in Experiment 1

Proportion incorrect	High-expectancy items	Low-expectancy items	M
Recall			
0 %	.23 (.14)	.06 (.08)	.15
33 %	.51 (.18)	.25 (.22)	.38
66 %	.49 (.26)	.29 (.24)	.39
100 %	.45 (.19)	.19 (.13)	.32
Remember			
0 %	.06 (.17)	.02 (.07)	.04
33 %	.16 (.16)	.05 (.19)	.11
66 %	.19 (.20)	.12 (.19)	.14
100 %	.17 (.21)	.06 (.13)	.12
Know			
0 %	.17 (.17)	.04 (.07)	.11
33 %	.35 (.16)	.20 (.19)	.28
66 %	.30 (.20)	.17 (.19)	.24
100 %	.28 (.21)	.13 (.13)	.21

Standard deviations are reported in parentheses

“Know” judgments A separate 4 (proportion incorrect: 0 %, 33 %, 66 %, 100 %) \times 2 (item expectancy: high, low) mixed ANOVA run on “know” responses revealed significant main effects of proportion incorrect, $F(3, 68) = 4.90$, $MSE = .02$, and item expectancy, $F(1, 68) = 35.06$, $MSE = .04$, with no interaction, $F < 1.0$, $p > .05$, $p_{BIC} = .90$. As with the false recall and remember analyses, participants were more likely to give “know” responses for high-expectancy than for low-expectancy items, and for items that were suggested by a confederate than for those not suggested by a confederate, $t_s > 2.6$, $p_s < .05$. Furthermore, participants in the 33 %-, 66 %-, and 100 %-incorrect conditions were equally likely to give “know” judgments for false items, $t_s \leq 1.67$, $p_s > .05$.

Final questionnaires Additional analyses were conducted to determine whether partner accuracy influenced participants’ willingness to work with their partner (the confederate) again, as well as their subjective ratings of partner credibility and memory abilities (see Table 2). Interestingly, no significant effects of proportion incorrect emerged, $F_s < 1.12$, $p_s > .05$. Partner accuracy alone may not have been a strong enough manipulation to warrant discounting a partner, even though the present study included the strongest possible manipulation of accuracy (i.e., a partner who was 100 % inaccurate throughout the duration of the entire experiment).

Experiment 2

Contrary to predictions, the results of Experiment 1 suggested that partner accuracy alone was not enough to alert participants to item discrepancies and/or to reduce false memories in the social contagion paradigm. Even when participants recalled alongside a confederate who was *entirely* inaccurate, they later incorporated the confederate’s erroneous suggestions. One explanation for such results is that we generally

Table 2 Mean proportions of participants who would choose to work with their partner again (if monetary reward was contingent on accuracy), mean ratings of partner memory, and mean ratings of partner credibility as a function of proportions of incorrect information suggested by the confederate in Experiment Experiment 1

Proportion incorrect	Choose again	Partner memory	Partner credibility
0 %	.94 (.24)	2.5 (1.2)	2.3 (1.1)
33 %	.89 (.32)	2.3 (1.1)	2.3 (1.1)
66 %	.94 (.24)	2.7 (1.1)	2.1 (0.83)
100 %	.77 (.43)	2.4 (0.97)	2.5 (1.2)

Standard deviations are reported in parentheses. Likert scales for partner memory and credibility were reverse ordered (1 = *best memory/most credible*, 5 = *worst memory/least credible*).

assume that memory partners are informative and recollecting as accurately as possible (see Harris et al., 2008; Jaeger et al., 2012). The purpose of Experiment 2 was to examine whether person credibility (operationalized as the partner’s memory ability) would interact with partner accuracy to influence the contagion effect. Would participants still be likely to adopt items in the 100 %-inaccurate condition when they witnessed firsthand that their partner had a “very poor” memory?

Research has established that person credibility affects the likelihood that participants will incorporate misleading suggestions into their memory reports. For example, Echterhoff, Higgins, and Groll (2005) found that both untrustworthy and incompetent sources of misinformation reduced false memories (see, too, Andrews & Rapp, 2012; Chambers & Zaragoza, 2001; Dodd & Bradshaw, 1980; Echterhoff et al., 2007; Echterhoff, Higgins, & Groll, 2005; Highhouse & Bottrill, 1995; Hoffman, Granhag, Kwong See, & Loftus, 2001; Smith & Ellsworth, 1987; Underwood & Pezdek, 1998).

The question of paramount concern for Experiment 2 was whether manipulating perceptions of person credibility would urge participants to more closely monitor the output of their partner, and thus alert participants to their partner’s large proportion of inaccurate suggestions. In contrast to hearing an explicit, experimenter-issued warning regarding a perceived partner’s general credibility, the participants in the present study witnessed firsthand the confederate’s *memory* ability on a related *memory* task. Furthermore, they were required to score the confederate’s recall and then categorize the confederate as having either a “very poor” or a “very good” memory. Again, we employed no specific instructions or public announcements of the confederate’s credibility, so participants had to infer partner credibility on the basis of the partner’s performance, rather than on what they were told about their partner.

In Experiment 2, we hypothesized that participants who believed they were recalling with a partner who had a “very good” memory should replicate the results of Experiment 1 and demonstrate no difference in false recall between the 33 %-inaccurate and 100 %-inaccurate conditions (the 66 %-incorrect condition was not included in Exp. 2, because in Exp. 1 its results had been equivalent to those in the 33 %-incorrect condition). In contrast, participants who believed that they were recalling with a partner who had a “very poor” memory should more closely monitor the confederate’s output, and thus reduce false recall of the critical contagion items, especially in the 100 %-inaccurate condition.

Experiment 2 also included a final source-monitoring recognition test. Any effects of partner credibility should be especially salient on the recognition test, because the test directs attention to the source of misleading suggestions and so may aid in the reduction of false memory (see Huff et al., 2013; Multhaup, 1995).

Method

Participants The participants were 115 Montana State University undergraduates who participated for course credit. Seven of them were excluded due to suspicion, insufficient English proficiency, or failure to follow instructions. The final analysis included 108 participants.

Design This experiment was based on a $2 \times 3 \times 2$ mixed design. Expectancy of the contagion items (high or low expectancy) was manipulated within subjects, and the proportion of false information (0 %, 33 %, or 100 %) and perceptions of partner credibility (very good memory or very poor memory) were manipulated between subjects. The primary dependent variables were again false recall and false recognition of the critical suggested items.

Materials The same materials used in Experiment 1 were also used in Experiment 2, with a few exceptions. First, the “pilot” study (which contained our manipulation of partner credibility) required a 15-item categorized list (developed from Battig & Montague, 1969), as well as a recall sheet. Second, a 36-item recognition task used by Meade and Roediger (2002) was included. Half of the items on the recognition test were previously studied items (three items from each of the six scenes), 12 of the items were potentially misleading (the high- and low-expectancy items from each scene), and the remaining six items were fillers. Finally, we reversed the order of the Likert scales used in our final manipulation check, so that 5 indicated *high scores* and 1 indicated *low scores*).

Procedure Participants were tested with a same-age confederate. Before participants viewed the scenes, they completed a pilot task that contained our manipulation of partner memory ability. The participants and confederates studied identical word lists on the computer (for 60 s), and then one of them (always the confederate) was “randomly” selected to recall the list aloud and the other (always the participant) was selected to record their responses. The participant was instructed to mark an “X” on a recall sheet next to the words that the confederate recalled. The recall sheet directed the participant to total the correct number of responses, and for additional salience, to circle a category that corresponded to the total number of words recalled. In the high-credibility condition, the confederate recalled 13 of the 15 words correctly, performance that fell within the “very good memory” category. In the low-credibility condition, the confederate recalled only three of the 15 words correctly and inserted one intrusion, performance that fell within the “very poor memory” category. All participants correctly categorized the confederates in accordance with their performance. The experimenter made no explicit or public announcements of performance.

The remainder of the procedure was identical to that of Experiment 1, with one exception: Following individual recall, participants were asked to complete a recognition test that required them to indicate the source of each item (scene, other participant, both the scene and the other participant, or never presented). There was no time limit for this task, and all participants completed it in less than 10 min.

Results

False recall False recall results for the critical contagion items are shown in Table 3. A 3 (proportion incorrect: 0 %, 33 %, 100 %) $\times 2$ (item expectancy: high, low) $\times 2$ (partner credibility: high, low) mixed factorial ANOVA was conducted on the mean proportions of critical contagion items recalled. Replicating Experiment 1, we found a significant social contagion effect, and the proportion of false items did not mediate this effect, $F(2, 102) = 26.52$, $MSE = 0.04$. Follow-up tests confirmed that participants in the 33 %-incorrect condition showed greater contagion ($M = .32$) than did participants in the control (0 %-incorrect) condition ($M = .11$), $t(70) = -7.23$, $SEM = .03$, as did participants in the 100 %-incorrect condition ($M = .34$), $t(70) = -7.11$, $SEM = .03$. However the 33 %- and 100 %-incorrect conditions did not differ from one another, $t(70) = -0.67$, $SEM = .04$, $p > .05$. Participants were also more likely to recall high-expectancy contagion items ($M = .37$) than low-expectancy contagion items ($M = .14$), $F(1, 102) = 151.06$, $MSE = 0.02$. Critically, the main effect of partner credibility was not significant, $F(1,$

Table 3 Mean proportions of false recall, and “remember” or “know” responses for high- and low-expectancy items as a function of proportions of incorrect items suggested by the confederate and perceived partner credibility in Experiment 2

Proportion incorrect	High partner credibility			Low partner credibility		
	High-expectancy items	Low-expectancy items	<i>M</i>	High-expectancy items	Low-expectancy items	<i>M</i>
Recall						
0 %	.19 (.17)	.03 (.06)	.11	.20 (.20)	.03 (.06)	.12
33 %	.48 (.14)	.19 (.15)	.34	.42 (.19)	.17 (.17)	.30
100 %	.47 (.26)	.24 (.22)	.36	.49 (.17)	.16 (.18)	.33
Remember						
0 %	.07 (.16)	.00 (.00)	.04	.02 (.05)	.01 (.04)	.02
33 %	.12 (.11)	.05 (.10)	.09	.13 (.18)	.05 (.13)	.09
100 %	.15 (.18)	.05 (.10)	.10	.14 (.16)	.06 (.11)	.10
Know						
0 %	.12 (.13)	.03 (.07)	.08	.19 (.15)	.02 (.05)	.11
33 %	.36 (.16)	.15 (.14)	.26	.29 (.19)	.12 (.15)	.21
100 %	.32 (.20)	.19 (.21)	.26	.35 (.21)	.10 (.12)	.23

Standard deviations are reported in parentheses

102) = 1.67, $MSE = 0.04$, $p > .05$, $p_{BIC} = .99$; participants were as likely to adopt false information from a partner who they believed had a “very poor memory” as they were to adopt false information from a partner who they thought had a “very good memory.” No interactions were significant (all $F_s \leq 2.81$, $p_s > .05$, $p_{BICs} \geq .86$).

“Remember” judgments A 3 (proportion incorrect: 0 %, 33 %, 100 %) \times 2 (item expectancy: high, low) \times 2 (partner credibility: high, low) mixed factorial ANOVA again confirmed main effects of proportion incorrect, $F(2, 102) = 6.59$, $MSE = 0.02$, and item expectancy, $F(1, 102) = 20.74$ $MSE = 0.02$, but no main effect of partner credibility, $F(1, 102) = .21$, $MSE = 0.02$, $p > .05$, $p_{BIC} = .90$, nor any interactions ($F_s \leq 1.20$, $p_s > .05$). Follow-up tests confirmed that all conditions differed significantly from the control condition [$t(70) = -3.02$, $SEM = 0.02$, for the 0 %- and 33 %-incorrect conditions; $t(70) = -3.78$, $SEM = 0.02$, for the 0 %- and 100 %-incorrect conditions], but again the remaining conditions did not differ from one another [$t(70) = -0.20$, $SEM = 0.08$, $p > .05$, for 33 % and 100 %]. Participants were more willing to give “remember” responses to items suggested by the confederate than to items not suggested by the confederate, and to high-expectancy items than to low-expectancy items, but manipulating the perceived memorial ability of the confederate did not moderate these effects.

“Know” judgments Separate analyses on “know” responses revealed main effects of proportion incorrect, $F(2, 102) = 16.85$, $MSE = 0.03$, and item expectancy, $F(1, 102) = 96.06$ $MSE = 0.03$, no main effect of partner credibility, $F(1, 102) = 0.42$, $MSE = 0.03$, $p > .05$, $p_{BIC} = .89$, and no interactions (all $F_s \leq 2.45$, $p_s > .05$, $p_{BICs} \geq .90$). Follow-up tests showed that the control condition differed significantly from the 33 %-incorrect condition, $t(70) = -5.55$, $SEM = 0.03$, and from the 100 %-incorrect condition, $t(70) = -5.39$, $SEM = 0.03$. However the 33 %- and 100 %-incorrect conditions did not differ significantly from one another, $t(70) = -0.22$, $SEM = 0.03$, $p > .05$. As with “recall” and “remember” responses, participants gave more “know” responses to items suggested by the confederate than to items not suggested by the confederate, and to high-expectancy items than to low-expectancy items, regardless of the perceived memorial ability of the confederate.

Recognition The mean proportions of participants’ responses on the final recognition/source-monitoring tests are displayed in Table 4. False recognition was defined as the proportion of critical contagion items that participants attributed to having occurred in the scenes (“scene only” responses plus “scene and other” responses). Accurate recognition was defined as the proportion of studied items that participants attributed to

Table 4 Mean proportions of false source judgments for critical contagion items and veridical source judgments for correct items as a function of proportions of incorrect items suggested by the confederate and perceived partner credibility in Experiment 2

Proportion incorrect	False recognition		Correct recognition	
	High partner credibility	Low partner credibility	High partner credibility	Low partner credibility
0 %	.33 (.29)	.31 (.18)	.63 (.26)	.64 (.26)
33 %	.54 (.32)	.50 (.35)	.65 (.14)	.66 (.20)
100 %	.44 (.34)	.63 (.37)	.56 (.21)	.57 (.16)

Standard deviations are shown in parentheses

having been presented in the scenes (“scene only” plus “scene and other” responses).

A 3 (proportion incorrect: 0 %, 33 %, 100 %) \times 2 (partner credibility: high, low) between-subjects ANOVA computed on false recognition revealed a significant main effect of proportion incorrect, $F(2, 102) = 4.95$, $MSE = .10$, but no main effect of partner credibility, $F(1, 102) = 0.54$, $MSE = .10$, $p > .05$, $p_{BIC} = .89$, and no interaction, $F(2, 102) = 1.63$, $MSE = .10$, $p > .05$, $p_{BIC} = .95$. Follow-up tests verified that both contagion conditions differed significantly from the control condition [$t(70) = -2.85$, $SEM = 0.07$, for the 0 %- and 33 %-incorrect conditions; $t(70) = -2.90$, $SEM = 0.07$, for the 0 %- and 100 %-incorrect conditions], but that the contagion conditions did not differ from one another, $t(70) = -0.20$, $SEM = 0.08$, $p > .05$. Importantly, partner credibility did not influence false recognition of contagion items.

Accurate source judgments for veridical items were equated across conditions, with no main effects or interactions (all $F_s \leq 1.80$, $p_s > .05$, $p_{BICs} \geq .94$). On the final recognition test, participants recognized correct items as having occurred in the scene, regardless of partner accuracy or partner credibility.

Final questionnaires Responses on the final questionnaire were analyzed to determine whether manipulations of partner accuracy and/or partner credibility had any effect on participants’ subjective judgments of partner accuracy or their willingness to work with their partner again (see Table 5). A 3 (proportion incorrect: 0 %, 33 %, 100 %) \times 2 (partner credibility: high, low) between-subjects ANOVA revealed no main effect of proportion incorrect, nor any significant interactions (all $F_s \leq 1.67$, $p_s > .05$). However, these data *did* show a significant main effect of partner credibility: Participants were more likely to choose to work again with a high-credibility partner, $F(1, 102) = 15.7$, $MSE = .17$, and they rated high-credibility partners as being more credible, $F(1, 102) = 11.30$, $MSE = .38$, and as having a better memory, $F(1, 102) = 15.96$, $MSE = .4$. These data suggest that participants were aware of,

Table 5 Mean proportions of participants who would choose to work with their partner again (if monetary consumption was contingent on accuracy), mean ratings of partner memory, and mean ratings of partner

credibility as a function of proportions of incorrect information suggested by the confederate and the perceived credibility (memorial ability) of the confederate in Experiment 2

Proportion of incorrect items	High partner credibility			Low partner credibility		
	Choose again	Partner credibility	Partner memory	Choose again	Partner credibility	Partner memory
0 %	.94 (.24)	4.1 (0.73)	4.2 (0.65)	.56 (.51)	3.7 (0.46)	3.7 (0.57)
33 %	.94 (.24)	4.1 (0.58)	4.3 (0.67)	.61 (.50)	3.4 (0.62)	3.4 (0.70)
100 %	.83 (.38)	3.7 (0.75)	3.9 (0.96)	.61 (.50)	3.5 (0.50)	3.6 (0.49)

Standard deviations are shown in parentheses. Likert scales: 1 = *worst memory/least credible*, 5 = *best memory/most credible*

and could later classify, their partner according to their manipulated credibility.

Experiment 3

The results of Experiment 2 replicated those of Experiment 1 by revealing that participants were likely to incorporate suggestions from a confederate who was entirely incorrect. Even more interesting is that, in Experiment 2, participants were explicitly aware that the confederate had a “very poor memory” and rated them as being lower on credibility, memory ability, and willingness to work with them again. This suggests that in spite of participants’ knowledge that their partner had a very poor memory, they still incorporated the confederate’s misleading suggestions into their own memories.

One possible explanation is that without the explicit instructions typically provided in credibility studies, participants did not spontaneously make the connection that poor performance on the “pilot” study was relevant to performance on the experimental task. We examined this possibility in Experiment 3 by manipulating credibility on the experimental task itself. Participants scored the confederate’s performance on a practice trial of the experimental task, so that there was no need to extrapolate confederate performance on one memory task to performance on another memory task. Again we manipulated the proportion of incorrect items (0 %, 33 %, or 100 % incorrect) to see whether the confederate’s performance on the experimental task would alert the participants to the fact that their partner was producing inaccurate items.

Method

Participants The participants were 130 Montana State University undergraduates who participated for partial course credit. Sixteen of them were excluded due to suspicion, insufficient English proficiency, or failure to follow instructions. The final analysis included 114 participants.

Design This experiment was based on a $2 \times 3 \times 2$ mixed design. Expectancy of the contagion items (high or low expectancy) was manipulated within subjects, and the proportion of false information (0 %, 33 %, or 100 %) and perceptions of partner credibility (very good memory or very poor memory) were manipulated between subjects. The primary dependent variables were again false recall and false recognition of the critical suggested items.

Materials and procedure The same materials and procedure were used as in Experiment 2, with one exception: One slide (the desk scene) was pulled out of the study presentation and used as a “practice test.” Participants studied the scene and then collaboratively recalled items from the scene with the confederate. As in Experiment 2, the confederate was scored either as having a “very poor memory” on this task (recalled zero correct and six incorrect intrusions) or as having a “very good memory” on this task (recalled six correct and zero incorrect intrusions). All participants correctly categorized the confederates in accordance with their performance.

The experimenter made no explicit or public announcements of performance.

Results

False recall The false-recall results, along with “remember” and “know” judgments, for critical contagion items are shown in Table 6. A 3 (proportion incorrect: 0 %, 33 %, 100 %) $\times 2$ (item expectancy: high, low) $\times 2$ (partner credibility: high, low) mixed factorial ANOVA was conducted on the mean proportions of critical contagion items recalled. Replicating the results of Experiments 1 and 2, participants were equally likely to incorporate misleading suggestions from accurate and inaccurate confederates, $F(2, 108) = 25.93$, $MSE = 0.05$. Participants in the 33 %-incorrect condition showed greater contagion ($M = .36$) than did participants in the control (0 %-incorrect) condition ($M = .12$), $t(74) = -6.00$, $SEM = .04$, as did participants in the 100 %-incorrect condition ($M = .34$), $t(74) = -7.27$, $SEM = .03$. However, the 33 %- and 100 %-incorrect conditions did not differ from one another, $t(74) = 0.38$, $SEM = .04$, $p > .05$. Participants were also more likely to

Table 6 Mean proportions of false recall and “remember” and “know” responses for high- and low-expectancy items as a function of proportions of incorrect items suggested by the confederate and perceived partner credibility in Experiment 3

Proportion incorrect	High partner credibility			Low partner credibility		
	High-expectancy items	Low-expectancy items	<i>M</i>	High-expectancy items	Low-expectancy items	<i>M</i>
Recall						
0 %	.17 (.25)	.03 (.07)	.10	.22 (.16)	.05 (.11)	.13
33 %	.53 (.23)	.29 (.19)	.41	.45 (.27)	.16 (.26)	.31
100 %	.48 (.20)	.27 (.22)	.38	.43 (.22)	.18 (.21)	.31
Remember						
0 %	.01 (.05)	.01 (.05)	.01	.03 (.07)	.01 (.05)	.02
33 %	.15 (.17)	.03 (.10)	.09	.12 (.20)	.03 (.10)	.08
100 %	.19 (.25)	.08 (.12)	.14	.11 (.15)	.05 (.13)	.08
Know						
0 %	.16 (.25)	.02 (.06)	.09	.19 (.17)	.04 (.11)	.11
33 %	.38 (.26)	.26 (.16)	.32	.33 (.20)	.13 (.20)	.23
100 %	.29 (.23)	.19 (.19)	.24	.32 (.21)	.13 (.20)	.23

Standard deviations are reported in parentheses

recall high-expectancy contagion items ($M = .38$) than low-expectancy contagion items ($M = .16$), $F(1, 108) = 76.20$, $MSE = 0.04$. Critically, the main effect of partner credibility was not significant, $F(1, 108) = 2.28$, $MSE = 0.05$, $p > .05$, $p_{BIC} = .98$. Participants were as likely to adopt false information from a partner with a very poor memory as they were to adopt false information from a partner with very good memory on the experimental task. No interactions were significant (all $F_s \leq 1.93$, $p_s > .05$, $p_{BICs} \geq .85$).

“Remember” judgments A 3 (proportion incorrect: 0 %, 33 %, 100 %) \times 2 (item expectancy: high, low) \times 2 (partner credibility: high, low) mixed factorial ANOVA again confirmed main effects of proportion incorrect, $F(2, 108) = 7.28$, $MSE = 0.03$, and item expectancy, $F(1, 108) = 20.73$, $MSE = 0.01$, along with a significant interaction between proportion incorrect and item expectancy, $F(2, 108) = 3.85$, $MSE = .01$. Follow-up t tests revealed that for high-expectancy items, all conditions differed significantly from the control condition [$t(74) = -3.63$, $SEM = 0.03$, for the 0 %- and 33 %-incorrect conditions; $t(74) = -3.70$, $SEM = 0.04$, for the 0 %- and 100 %-incorrect conditions], but again the remaining conditions did not differ from one another in terms of “remember” responses, $t(74) = -0.35$, $SEM = 0.05$, $p > .05$, for the 33 % and 100 % conditions. However, for low-expectancy items, “remember” responses only varied between the 0 % and 100 % conditions, $t(74) = -2.68$, $SEM = .02$. “Remember” responses for low-expectancy items did not differ between the 0 %- and 33 %-incorrect conditions, $t(74)$

$= -1.19$, $SEM = .02$, $p > .05$, nor between the 33 %- and 100 %-incorrect conditions, $t(74) = 1.42$, $SEM = .03$, $p > .05$. The main effect of partner credibility was not significant, $F(1, 102) = 0.21$, $MSE = 0.02$, $p > .05$, $p_{BIC} = .88$, nor did it interact with other variables, all $F_s \leq 1.20$, $p_s > .05$, $p_{BICs} \geq .98$.

“Know” judgments Separate analyses on “know” responses revealed main effects of proportion incorrect, $F(2, 108) = 12.79$, $MSE = 0.05$, and item expectancy, $F(1, 108) = 42.22$, $MSE = 0.03$, but no main effect of partner credibility, $F(1, 108) = 1.07$, $MSE = 0.05$, $p > .05$, $p_{BIC} = .86$, nor any interactions (all $F_s \leq 1.69$, $p_s > .05$, $p_{BICs} \geq .91$). Follow-up tests showed that the control condition differed significantly from the 33 %-incorrect condition, $t(74) = -5.10$, $SEM = 0.03$, and from the 100 %-incorrect condition, $t(74) = -3.83$, $SEM = 0.03$. However the 33 %- and 100 %-incorrect conditions did not differ significantly from one another, $t(74) = -1.09$, $SEM = 0.04$, $p > .05$. As with the “remember” responses, participants gave more “know” responses to items suggested by the confederate than to items not suggested by the confederate, and to high-expectancy than to low-expectancy items, regardless of the perceived memorial ability of the confederate.

Recognition The mean proportions of participants’ responses on the final recognition tests are displayed in Table 7. A 3 (proportion incorrect: 0 %, 33 %, 100 %) \times 2 (partner credibility: high, low) between-subjects ANOVA computed on false recognition revealed a significant main effect of proportion incorrect, $F(2, 108) = 4.72$, $MSE = .08$. Follow-up tests verified both contagion conditions differed significantly from the control condition [$t(74) = -3.06$, $SEM = 0.06$, for 0 %- and 33 %-incorrect conditions; $t(74) = -2.30$, $SEM = 0.07$, for 0 %- and 100 %-incorrect conditions], but that they did not differ from one another, $t(74) = -0.62$, $SEM = 0.07$, $p > .05$. Importantly, we found a marginal main effect of partner credibility, $F(1, 108) = 3.38$, $MSE = .08$, $p = .069$, $p_{BIC} = .55$, which is considered weak evidence in support of the null hypothesis (see Raftery, 1995). On the recognition test,

Table 7 Mean proportions of false source judgments for critical contagion items and of veridical source judgments for correct items as a function of proportions of incorrect items suggested by the confederate and perceived partner credibility in Experiment 3

Proportion incorrect	False recognition		Correct recognition	
	High partner credibility	Low partner credibility	High partner credibility	Low partner credibility
0 %	.42 (.24)	.44 (.25)	.88 (.12)	.69 (.22)
33 %	.69 (.32)	.55 (.29)	.81 (.13)	.67 (.18)
100 %	.67 (.31)	.48 (.31)	.66 (.18)	.62 (.16)

Standard deviations are shown in parentheses.

participants were marginally less likely to falsely recognize contagion items suggested by the low-credibility confederate ($M = .49$) than by the high-credibility confederate ($M = .59$). This effect is marginal and so should be interpreted with caution, but it may suggest that when participants witnessed the confederate's poor memory on the experimental task itself, they were able to use that information to reduce false recognition on the final source-monitoring test. The interaction between partner credibility and proportion incorrect was not significant, $F(2, 108) = 1.31$, $MSE = .08$, $p > .05$, $p_{BIC} = .87$.

To provide a more sensitive test of partner credibility, an additional analysis was conducted on only the 33 %- and 100 %-incorrect conditions. Only in these conditions did participants hear inaccurate suggestions from the confederate, and so have an opportunity to reject those items. A 2 (proportion incorrect: 33 % or 100 %) \times 2 (partner credibility: high or low) ANOVA revealed only a main effect of partner credibility, $F(1, 72) = 5.10$, $MSE = .09$, $p = .027$. This demonstrates that when participants were exposed to inaccurate confederate suggestions, they were less likely to falsely recognize those items when they were suggested by the low-credibility confederate ($M = .52$) than when they were suggested by the high-credibility confederate ($M = .68$).

A separate 3 (proportion incorrect: 0 %, 33 %, 100 %) \times 2 (partner credibility: high, low) between-subjects ANOVA computed on accurate recognition revealed significant main effects of both proportion incorrect, $F(2, 108) = 7.59$, $MSE = .03$, and partner credibility, $F(1, 108) = 15.31$, $MSE = .03$, but no interaction, $F = 1.89$, $p > .05$, $p_{BIC} = .94$. Follow-up tests confirmed that participants made fewer accurate source judgments in the 100 %-incorrect condition than in the 33 % condition, $t(74) = 2.51$, $SEM = .04$, and in the 0 % condition, $t(74) = 3.47$, $SEM = .04$, though accurate recognition did not vary between the 33 % and 0 % conditions, $t < 1.3$, $p > .05$. These data suggest that when participants were presented with 100 %-incorrect information from a partner, they reduced their accurate recognition as well.

Also important is the main effect of partner credibility on accurate recognition. Participants were less likely to accurately recognize information suggested by the low-credibility confederate ($M = .66$) than that suggested by the high-

credibility confederate ($M = .79$). Further support for the influence of partner credibility on accurate recognition comes from an additional analysis conducted only on the 0 % and 33 % conditions. These are the only conditions in which participants heard the confederate suggest accurate items, and so they offer a more sensitive test of whether or not participants rejected accurate suggestions by the low-credibility confederate. The analysis is conditional and includes only the specific items that the confederate recalled during collaboration. A 2 (proportion incorrect: 33 % or 100 %) \times 2 (partner credibility: high or low) ANOVA revealed only a main effect of partner credibility, $F(1, 72) = 14.73$, $MSE = .04$, $p = .00$. This suggests that when participants were exposed to correct confederate suggestions, they were less likely to accurately recognize items suggested by the low-credibility confederate ($M = .71$) than items suggested by the high-credibility confederate ($M = .86$). Considered together, the recognition data suggest that participants discounted both accurate and inaccurate suggestions from low-credibility confederates (although the reduction for false recognition was marginal).

Final questionnaires A 3 (proportion incorrect: 0 %, 33 %, 100 %) \times 2 (partner credibility: high, low) between-subjects ANOVA revealed that participants were more likely to choose to work again with a high-credibility partner, $F(1, 108) = 16.66$, $MSE = .17$, and they rated high-credibility partners as having more credibility, $F(1, 108) = 67.49$, $MSE = .51$, and better memory, $F(1, 108) = 58.72$, $MSE = .58$ (see Table 8). Interestingly, partner credibility also interacted with proportion incorrect for the average credibility ratings only, $F(2, 108) = 3.76$, $MSE = .54$. Follow-up tests confirmed that participants working with a high-credibility partner gave equivalent credibility ratings to partners in the 0 %-, 33 %-, and 100 %-inaccurate conditions, $t_s < 1.1$, $p_s > .05$. However, when working with a low-credibility partner, participants rated the confederate who was 100 % inaccurate as being less credible than both the partners who were 33 % inaccurate, $t(36) = 3.65$, $SEM = .29$, and 0 % inaccurate, $t(36) = 3.06$, $SEM = .29$ (credibility ratings for the 0 % and 33 % partners did not differ: $t < 1$, $p > .05$). Only when

Table 8 Mean proportions of participants who would choose to work with their partner again (if monetary consumption was contingent on accuracy), mean ratings of partner memory, and mean ratings of partner

credibility as a function of proportions of incorrect information suggested by the confederate and the perceived credibility (memorial ability) of the confederate in Experiment 3

Proportion of incorrect items	High partner credibility			Low partner credibility		
	Choose again	Partner credibility	Partner memory	Choose again	Partner credibility	Partner memory
0 %	.89 (.32)	4.5 (0.51)	4.6 (0.51)	.63 (.50)	3.5 (0.84)	3.4 (0.84)
33 %	.84 (.37)	4.5 (0.61)	4.5 (0.61)	.68 (.48)	3.7 (0.82)	3.6 (0.96)
100 %	.95 (.23)	4.3 (0.56)	4.5 (0.53)	.42 (.51)	2.6 (0.96)	3.1 (1.0)

Standard deviations are shown in parentheses. Likert scales: 1 = *worst memory/least credible*, 5 = *best memory/most credible*

participants were aware that their partner had a very poor memory on the exact task they were working on together did they notice that the 100 %-inaccurate partner was relatively less credible than more-accurate partners.

General discussion

The three experiments reported here were the first to systematically examine the influence of partner accuracy on socially introduced false memories. Importantly, these experiments revealed that manipulating partner accuracy (even in its most extreme version, by literally providing *only* incorrect information for the *duration* of the experiment) was not enough on its own to influence how likely participants were to incorporate misleading suggestions from the confederate. Across all three experiments, participants were just as likely to incorporate misleading suggestions from a mostly accurate confederate (33 % incorrect) as from an entirely inaccurate confederate (100 % inaccurate). Even when participants witnessed firsthand that the confederate had a poor memory on a related task (and reported that the confederate had poor memory and low credibility, and that they would not wish to work with the confederate in the future), they were still likely to adopt false items from the entirely inaccurate source (Exp. 2). Only when participants engaged in a practice trial of the experimental task itself (allowing them to witness firsthand the confederate's good/poor memory on the very task that they were about to perform together) were they less likely to attribute the low-credibility confederate's responses to having occurred in the original study episode (Exp. 3). Importantly, this knowledge did not influence false recall, but it selectively influenced recognition. Note that false and veridical recognition were not influenced by the confederate's actual accuracy on the experimental task, but instead were influenced by the confederate's performance on the practice task: The participants discounted low-credibility partners, regardless of whether they went on to be mostly accurate, and they did not discount high-credibility partners who went on to be mostly inaccurate.

The null effect of partner accuracy was obtained across all three experiments, and occurred both when participants had no information about the confederate's memory (Exp. 1) and when they knew that the confederate had a "very poor" memory (Exps. 2 and 3). Notably, significant levels of social contagion were obtained across experiments, suggesting that, consistent with past research (Meade & Roediger, 2002; Roediger et al., 2001), participants incorporated the confederate's misleading suggestions on both individual recall and recognition tests. One important finding from the present study is that even the strongest possible manipulation of partner accuracy (100 % inaccurate) did not modulate these effects. Partner accuracy also had no effect on participants'

metacognitive judgments of "remember" versus "know" responses: They were just as likely to report "remembering" items suggested by accurate and inaccurate confederates. Of course, null effects should be considered with caution, but the recall and metacognition results are consistent with those of Jaeger et al. (2012) in demonstrating that individuals view inaccurate partners as being informative on memory tests and that they do not spontaneously consider partner accuracy.

Partner credibility, however, did modulate the social contagion effect, but only when participants witnessed firsthand that the confederate performed poorly on the social contagion task itself (Exp. 3) *and* when the test directed participants to attend to the source (the recognition task). The partner credibility manipulation employed in this study was based on actual confederate performance rather than on the experimenter-issued instruction that had been used in many previous studies (e.g., Dodd & Bradshaw, 1980). Importantly, the participants in Experiments 2 and 3 reported on the final questionnaires that the low-credibility confederates had worse memory and were less credible than the high-credibility confederates. In spite of this, participants were just as likely to recall (and to report *remembering*) suggestions from the low-credibility confederates as from the high-credibility confederates. Only on the recognition test did participants spontaneously utilize their partner's memory performance on the exact task (Exp. 3) to later discredit confederate suggestions. Consistent with past research demonstrating reduced false-memory effects on recognition, but not recall, the source-monitoring recognition test used in the present study directed participants to utilize source information to reduce false recognition (Huff et al., 2013; Multhaup, 1995). Importantly, the reduction in false recognition was marginal, suggesting that even when participants were successfully able to discount inaccurate confederate suggestions, they were still unable to decisively reduce social contagion errors.

Source-monitoring theory (Johnson et al., 1993) can account for the finding that even when participants were aware that their partner had a poor memory, they still incorporated their partners' erroneous suggestions. The suggested items were schematically similar to the studied items, thus rendering source discrimination more difficult (see Johnson & Raye, 1998). In addition, strong metacognitive assumptions surround one's partner (e.g., Harris et al., 2008; Jaeger et al., 2012). Such metacognitive assumptions could reduce the processing of suggested items, thereby making discrimination between studied and suggested items more difficult at test. It is also possible that metacognitive assumptions about partner accuracy could result in a more lenient decision criterion being used to incorporate suggestions (see Allan & Gabbert, 2008; Davis & Meade, 2013; Gabbert et al., 2003; Gabbert et al., 2004; Harris et al., 2008; Hoffman et al., 2001; Jaeger et al., 2012; Meade & Roediger, 2002; Paterson & Kemp, 2006). It

is likely that both reduced discriminability and lenient decision criteria influenced participants' memory decisions.

In conclusion, the present study revealed several interesting and unexpected findings related to socially transmitted false memories. Importantly, participants were just as likely to incorporate misleading suggestions from a partner who was mostly accurate (33 % incorrect) as from a partner who was not at all accurate (100 % incorrect). Second, even when participants were aware that the person they were remembering with had a "very poor" memory on a related memory task, this information was not enough to induce them to discredit that person's output. Only when participants were aware that their partner had a "very poor" memory on the experimental task itself were they able to marginally discredit the confederate's suggestions on the final recognition test. The findings of the present study highlight the robust nature of socially suggested memory errors, and suggest that participants spontaneously consider their partner's memory ability only when it is tied exactly to the task at hand and when the test encourages participants to consider the source of information. More generally, individuals do not spontaneously differentiate suggestions from accurate and inaccurate partners on social memory tests.

Author Note Experiments 1 and 2 were completed to fulfill the Master's thesis requirements at Montana State University for the first author. We thank Francisco Velasco, Chase Ladd, Meg Huntoon, Carla Schwarting, Laura Clark, Rachel Dunkel, and Bethany Spring for help running participants. Additionally, we thank Keith Hutchison and Wes Lynch for helpful discussion regarding the methodology of this project and for comments on earlier drafts.

Appendix A: confederate scripts

Condition 1		
<u>Toolbox</u>	<u>Bathroom</u>	<u>Kitchen</u>
Knife (metal supports)	Shampoo (Tissue Box)	Coffee maker (Wire whisk)
Bungee cord (tool box)	Toothpaste (sink)	Phone (Books)
Extension cord (table)	Razor (Mirror)	Cutting Board (metal spoons)
Wrench (Hammer)	Magazine (Cup)	Magnets (Spices)
Washers (Saw)	Deodorant (Medicine Cabinet)	Paper towel (stove)
Hair clip (Banana)	Shaving Gel (Scissors)	Tea Kettle (Spice rack)
<u>Bedroom</u>	<u>Closet</u>	<u>Desk</u>
Candle (wine glass)	Suitcase (Gloves)	Plant (monitor)
Fan (video)	Tie (Hat)	Tape dispenser (pen)
Pillow (window)	Flashlight (helmet)	Planner (dictionary)
Bed (Picture)	Cooler (Soup)	Filing cabinet (Chair)

Condition 1		
Vacuum (Comforter)	Umbrella (Hangers)	Keyboard (Lamp)
Laundry bag (Dresser)	Frisbee (shirts)	Mouse Pad (Photos)

(0 % false information).

Condition 2		
<u>Toolbox</u>	<u>Bathroom</u>	<u>Kitchen</u>
Knife (metal supports)	Shampoo (Tissue Box)	Coffee maker (Wire whisk)
Bungee cord (tool box)	Toothpaste (sink)	Phone (Books)
Extension cord (table)	Razor (Mirror)	Cutting Board (metal spoons)
Screws (Pliers)	Toothbrush (Shower)	Toaster (Cups)
Washers (Saw)	Deodorant (medicine Cabinet)	Paper towel (stove)
Ruler (Pencil)	Hair Brush (Contact Solution)	Oven Mitts (Napkins)
<u>Bedroom</u>	<u>Closet</u>	<u>Desk</u>
Candle (wine glass)	Suitcase (Gloves)	Plant (monitor)
Fan (video)	Tie (Hat)	Tape dispenser (pen)
Pillow (window)	Flashlight (helmet)	Planner (dictionary)
Night Stand (Mirror)	Shoes (Boxes)	Printer (Paper Clips)
Vacuum (Comforter)	Umbrella (Hangers)	Keyboard (Lamp)
Cologne (Quilt)	Ball (Gym Shorts)	Rolodex (Radio)

(33 % false information). Note: bolded words denote false items.

Condition 3		
<u>Toolbox</u>	<u>Bathroom</u>	<u>Kitchen</u>
Knife (metal supports)	Shampoo (Tissue Box)	Coffee maker (Wire whisk)
Chisel (Duct Tape)	Hairspray (Dental floss)	Blender (Spatula)
Extension cord (table)	Razor (Mirror)	Cutting Board (Metal spoons)
Screws (Pliers)	Toothbrush (Shower)	Toaster (Cups)
Tape Measure (Bolts)	Soap (Bath rug)	Knives (Pots)
Ruler (Pencil)	Hair Brush (Contact Solution)	Oven Mitts (Napkins)
<u>Bedroom</u>	<u>Closet</u>	<u>Desk</u>
Candle (wine glass)	Suitcase (Gloves)	Plant (monitor)
Water glass (Rug)	Slippers (Boots)	Briefcase (Glue)
Pillow (window)	Flashlight (helmet)	Planner (dictionary)
Night Stand (Mirror)	Shoes (Boxes)	Printer (paper Clips)
TV (Lamp)	Belt (Bags)	Calendar (Phone)
Cologne (Quilt)	Ball (Gym Shorts)	Rolodex (Radio)

(66 % false information). Note: bolded words denote false items.

Condition 4		
Toolbox	Bathroom	Kitchen
Bolts (Electric Tape)	Mouthwash (Bathtub)	Dishwasher (Forks)
Chisel (Duct Tape)	Hairspray (Dental floss)	Blender (Spatula)
Band-aid (Coins)	Nail clippers (Plunger)	Toaster (Cups)
Tape measure (Bolts)	Soap (Bath rug)	Knives (Pots)
Ruler (Pencil)	Hair Brush (Contact Solution)	Oven Mitts (Napkins)
Bedroom	Closet	Desk
Alarm Clock (Books)	Coats (Socks)	Stapler (Books)
Water glass (Rug)	Slippers (Boots)	Briefcase (Glue)
Reading Glasses (Backpack)	Suit (Lotion)	Money (Candy)
Night Stand (Mirror)	Shoes (Boxes)	Printer (Paper Clips)
TV (Lamp)	Belt (Bags)	Calendar (Phones)
Cologne (Quilt)	Ball (Gym Shorts)	Rolodex (radio)

(100 % false information). Note: bolded words denote false items.

References

Allan, K., & Gabbert, F. (2008). I still think it was a banana: Memorable “lies” and forgettable “truths.”. *Acta Psychologica*, *127*, 299–308. doi:10.1016/j.actpsy.2007.06.001

Allan, K., Midjord, J. P., Martin, D., & Gabbert, F. (2012). Memory conformity and the perceived accuracy of self versus other. *Memory & Cognition*, *40*, 280–286. doi:10.3758/s13421-011-0141-9

Andrews, J. J., & Rapp, D. N. (2012, November). *Does group composition matter? Partner characteristics and social contagion*. Poster presented at the annual meeting of the Psychonomic Society, Minneapolis, MN.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, *80*(3, Pt. 2), 1–46. doi:10.1037/h0027577

Bodner, G. E., Musch, E., & Azad, T. (2009). Reevaluating the potency of the memory conformity effect. *Memory & Cognition*, *37*, 1069–1076. doi:10.3758/MC.37.8.1069

Chambers, K. L., & Zaragoza, M. S. (2001). Intended and unintended effects of explicit warnings on eyewitness suggestibility: Evidence from source identification tests. *Memory & Cognition*, *29*, 1120–1129.

Cuc, A., Ozuru, Y., Manier, D., & Hirst, W. (2006). On the formation of collective memories: The role of a dominant narrator. *Memory & Cognition*, *34*, 752–762. doi:10.3758/BF03193423

Davis, S. D., & Meade, M. L. (2013). Both young and older adults discount suggestions from older adults on a social memory test. *Psychonomic Bulletin & Review*, *20*, 760–765. doi:10.3758/s13423-013-0392-5

Dodd, D. H., & Bradshaw, J. T. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning and Verbal Behavior*, *19*, 695–704.

Echterhoff, G., Groll, S., & Hirst, W. (2007). Tainted truth: Overcorrection for misinformation influence on eyewitness memory. *Social Cognition*, *25*, 367–409. doi:10.1521/soco.2007.25.3.367

Echterhoff, G., Higgins, E. T., & Groll, S. (2005a). Audience-tuning effects on memory: The role of shared reality. *Journal of Personality and Social Psychology*, *89*, 257–276.

Echterhoff, G., Hirst, W., & Hussy, W. (2005b). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory & Cognition*, *33*, 770–782. doi:10.3758/BF03193073

French, L., Garry, M., & Mori, K. (2011). Relative—not absolute—judgments of credibility affect susceptibility to misinformation conveyed during discussion. *Acta Psychologica*, *136*, 119–128. doi:10.1016/j.actpsy.2010.10.009

Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of social encountered misinformation. *Legal and Criminological Psychology*, *9*, 215–227.

Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other’s memories for an event? *Applied Cognitive Psychology*, *17*, 533–543.

Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps toward influence during a discussion. *Psychonomic Bulletin & Review*, *13*, 480–485.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*, 309–313. doi:10.3758/BF03197041

Greene, E., Flynn, M., & Loftus, E. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior*, *21*, 207–219.

Harris, C. B., Paterson, H. M., & Kemp, R. I. (2008). Collaborative recall and collective memory: What happens when we remember together? *Memory*, *16*, 213–230. doi:10.1080/09658210701811862

Highhouse, S., & Bottrill, K. V. (1995). The influence of social (mis)information on memory for behavior in an employment interview. *Organizational Behavior and Human Decision Processes*, *62*, 220–229.

Hirst, W., & Echterhoff, G. (2012). Remembering in conversations: The social sharing and reshaping of memories. *Annual Review of Psychology*, *63*, 55–79. doi:10.1146/annurev-psych-120710-100340

Hoffman, H. G., Granhag, P., Kwong See, S. T., & Loftus, E. F. (2001). Social influences on reality monitoring decisions. *Memory & Cognition*, *29*, 394–404. doi:10.3758/BF03196390

Huff, M. J., Davis, S. D., & Meade, M. L. (2013). The effects of initial testing on false recall and false recognition in the social contagion of memory paradigm. *Memory & Cognition*, *41*, 820–831. doi:10.3758/s13421-013-0299-4

Jaeger, A., Lauris, P., Selmeczy, D., & Dobbins, I. G. (2012). The costs and benefits of memory conformity. *Memory & Cognition*, *40*, 101–112. doi:10.3758/s13421-011-0130-z

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3–28. doi:10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. (1998). False memories and confabulations. *Trends in Cognitive Sciences*, *2*, 137–145.

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. doi:10.1037/0278-7393.4.1.19

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*, 679–690. doi:10.3758/s13428-010-0049-5

Meade, M. L., & Roediger, H. L., III. (2002). Explorations in the social contagion of memory. *Memory & Cognition*, *30*, 995–1009. doi:10.3758/BF03194318

- Muller, F., & Hirst, W. (2010). Resistance to the influences of others: Limits to the formation of collective memory through conversational remembering. *Applied Cognitive Psychology, 24*, 608–625.
- Multhaup, K. S. (1995). Aging, source, and decision criteria: When false fame errors do and do not occur. *Psychology and Aging, 10*, 492–497. doi:10.1037/0882-7974.10.3.492
- Paterson, H. M., & Kemp, R. I. (2006). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime & Law, 12*, 181–191. doi:10.1080/10683160512331316334
- Paterson, H. M., Kemp, R. I., & Ng, J. R. (2011). Combating co-witness contamination: Attempting to decrease the negative effects of discussion on eyewitness memory. *Applied Cognitive Psychology, 25*, 43–52.
- Raftery, A. E. (1995). Bayesian model selection in social research. In P. V. Marsden (Ed.), *Sociological methodology 1995* (pp. 111–196). Cambridge, UK: Blackwell.
- Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21*, 89–102. doi:10.3758/BF03211168
- Rajaram, S. (2011). Collaboration both hurts and helps memory: A cognitive perspective. *Current Directions in Psychological Science, 20*, 76–81. doi:10.1177/0963721411403251
- Roediger, H. L., III, Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review, 8*, 365–371. doi:10.3758/BF03196174
- Skagerberg, E. M., & Wright, D. B. (2008). The co-witness misinformation effect: Memory blends or memory compliance? *Memory, 16*, 436–442.
- Smith, V. L., & Ellsworth, P. C. (1987). The social psychology of eyewitness accuracy: Leading questions and communicator expertise. *Journal of Applied Psychology, 72*, 292–300.
- Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading post-event information. *Memory & Cognition, 14*, 329–338. doi:10.3758/BF03202511
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26*, 1–12.
- Underwood, J., & Pezdek, K. (1998). Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review, 5*, 449–453. doi:10.3758/BF03208820
- Wright, D. B., Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology, 91*, 189–202. doi:10.1348/000712600161781