Acta Psychologica xxx (xxxx) xxx-xxx

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: http://ees.elsevier.com

Long-lasting positive effects of collaborative remembering on false assents to misleading questions

Clelia Rossi-Arnaud [a,*], Pietro Spataro [b], Divya Bhatia [a], Fabrizio Doricchi [a,c], Serena Mastroberardino [a], Vincenzo Cestari [a]

[a] Sapienza University, Department of Psychology, Via dei Marsi 78, Rome, Italy
[b] Mercatorum University, Department of Economy, Piazza Mattei 10, Rome, Italy
[c] Fondazione Santa Lucia IRCCS, Rome, Italy

ARTICLE INFO

Keywords
Collaboration
Eyewitness
Misleading questions
Misinformation
False memory
Long-term recognition

ABSTRACT

Previous studies showed that collaborative remembering can reduce false memories through a process of mutual error checking, although conclusions were limited by the nature of the memory tasks (very few errors). The present experiments extend these findings to eyewitness memory by using a paradigm designed to increase the frequency of memory errors. Collaborative and nominal pairs viewed a video-clip illustrating a bank robbery, provided an immediate free recall, were forced to confabulate answers to false-event questions, and, after a short- (1 h: Experiment 1) or a long-term delay (1 week: Experiment 2), were administered a yes/no recognition task in which the misleading statements either matched the questions presented in the confabulation phase (answered questions) or not (control questions). Collaborative pairs recalled fewer correct details in the immediate free recall task, replicating the negative effects of collaborative inhibition. Most importantly, in the final recognition test, collaborative pairs were less likely to provide false assents to misleading statements, regardless of whether they had provided a response to the related false-event questions 1 h or 1 week earlier. Our results suggest that collaboration can increase the eyewitnesses' tendency to check the accuracy of others' responses and reject false memories through discussion.

1. Introduction

Starting from the publication of the seminal studies by Weldon and Bellinger (1997) and Basden, Basden, Bryner, and Thomas (1997), data on the effects of collaborative remembering on episodic memory has rapidly accumulated. Most experiments in this field have used a standard paradigm in which participants were individually presented with to-be-remembered stimuli and were later asked to collaborate during the retrieval phase by working with other members, in dyads or triads (Rajaram, 2011; Rajaram & Pereira-Pasarin, 2010). Memory accuracy of collaborative groups was later compared with the mean performance of individuals working alone, to determine whether group interaction benefitted memory over and above the statistical facilitation resulting from aggregating independent correct responses together (see Brennan & Enns, 2015a, 2015b, 2016, for discussion). The typical result was that the recall of collaborative groups exceeded the average recall of individuals working independently (e.g., Clark, Hori, Putnam, & Martin, 2000; Clark, Stephenson, & Kniveton, 1990; Lorge & Solomon, 1961; Stephenson, Brandstätter, & Wagner, 1983; Weldon & Bellinger, 1997).

A different question is whether collaboration may result in a process loss, such that two (or three) individuals working in group recall less items than two (or three) individuals who do not collaborate (*nominal groups*). According to Basden et al. (1997), the fairest test to examine this issue would be to compare the total recall of collaborative groups with the total nonredundant output of nominal groups. For example, suppose that Participant 1 recalls three items (A, B, and C), Participant 2 recalls four items (B, C, D, and E), and Participant 3 recalls five items (C, D, E, F, and G). If participants work alone, their mean recall would be 4 items – $(3 + 4 + 5)/3$. However, if participants collaborate during retrieval, then, according to the additive model (Lorge & Solomon, 1955), their expected recall should be 7 items (A, B, C, D, E, F, G). This method of scoring the performance of the nominal group has the advantage of providing a direct test for the presence of inhibition: if collaborative groups recall <7 items, then it can be inferred that something about collaboration prevented the group from reaching the maximum expected performance. In agreement, previous stud-

* Corresponding author at: Department of Psychology, Sapienza University, Via deiMarsi 78, 00185 Rome, Italy.
E-mail addresses: clelia.rossi-arnaud@uniroma1.it (C. Rossi-Arnaud); pietro.spataro@unimercatorum.it (P. Spataro); divya.bhatia@uniroma1.it (D. Bhatia); fabrizio.doricchi@uniroma1.it (F. Doricchi); serena.mastroberardino@uniroma1.it (S. Mastroberardino); vincenzo.cestari@uniroma1.it (V. Cestari)

https://doi.org/10.1016/j.actpsy.2019.102986
Received 20 June 2019; Received in revised form 20 November 2019; Accepted 18 December 2019
Available online xxx
0001-6918/

ies using this approach have repeatedly confirmed that collaborative groups remember significantly less information than nominal groups (see Rajaram, 2011, and Rajaram & Pereira-Pasarin, 2010, for reviews). This robust negative effect on productivity, thenceforth called *collaborative inhibition* (Weldon & Roediger, 1997), is thought to arise because collaboration undermines the use of individual retrieval strategies (Basden et al., 1997; Wright & Klumpp, 2004). More specifically, the *Retrieval Strategy Disruption* (RSD) hypothesis assumes that each participant encodes the studied material in an idiosyncratic order and that his/her memory performance will benefit to the extent that he/she can use the same organizational structure to guide retrieval. In a collaborative context, however, each participant is simultaneously exposed to the material retrieved by collaborators, which likely follows different orders (Barber, Harris, & Rajaram, 2015). This misalignment would induce participants to adopt a different, less optimal retrieval strategy, thereby reducing recall accuracy. A quantitative meta-analysis has recently summarized this body of research (Marion & Thorley, 2016): overall, the mean effect size across 64 studies was −0.78, confirming that collaborative groups remembered significantly less information than nominal groups.

In addition to negative consequences, an increasing body of research has begun to show that collaboration can also have positive effects on later memory performance. Specifically, it has been found that the recall of incorrect information tends to be lower in collaborative than in nominal groups – a finding referred to as *error pruning* (Ross, Spencer, Blatz, & Restorick, 2008; Ross, Spencer, Linardatos, Lam, & Perunovic, 2004; Rossi-Arnaud, Pieroni, Spataro, & Cestari, 2011; Takahashi, 2007). The prevailing view is that collaboration helps participants curtail errors because they receive corrective feedback from collaborators (Clark, Abbe, & Larson, 2006; Clark et al., 2000; Rajaram, 2011). Consider an associative recognition task in which participants study triples of words (denoted ABC, DEF, GHI, etc.) and are later asked to discriminate intact old triples (e.g., ABC) from newly rearranged ones (e.g., ABF or AEI). Clark, Abbe and Larson (2006, see also Clark, Hori, Putnam, & Martin, 2000) reasoned that collaboration can produce a decrease in false alarm rates if interacting members base their decisions to reject new triples on the retrieval of affirmative evidence. That is, the rearranged triple ABF can be rejected by the collaborative group if at least one member recalls that A and B were studied with C, not F, and shares this evidence with collaborators. According to this hypothesis, a decrease in the false alarm rates of collaborative groups can occur if: (a) the group members disagree on the old/new status of a rearranged triple, (b) one of the group members recalls the correct triple, and (c) the recalled information is sufficiently convincing to induce other members to revise their original responses. Clark et al. (2006) found that such a recall-to-reject strategy was effectively used in three-person collaborative groups; however, they also showed that positive evidence was difficult to obtain in standard recognition tasks because the optimal combination of outcomes – specifically, disagreement and recall of correct pairings – was very infrequent (in their first experiment, it occurred only 27 times, out of 400 rearranged trials).

Later studies have tried to increase the likelihood of disagreements in various ways – for example, by presenting participants with incorrect post-event information, by requiring them to answer misleading questions or by using paradigms specifically designed to generate memory errors (like the Deese-Roediger-McDermott paradigm: Takahashi, 2007). Regarding misinformation, Karns, Irvin, Suranic, and Rivardo (2009) employed a three-phase procedure in which participants first observed a video depicting a car accident, then read a description that included accurate or inaccurate details and, finally, answered a series of questions, some of which addressed the false details reported in the inaccurate description. The latter phase was performed either individually or in collaboration with a partner. As expected, participants exposed to the inaccurate description incorporated the false details into their memory traces and thus produced significantly less correct responses, compared to participants exposed to the accurate description; importantly, the negative effect of misinformation was smaller in the collaborative than in the individual condition, suggesting that collaboration led to a decrease in the acceptance of questions including incorrect details. However, a later study by the same research group (Rivardo et al., 2013) reached a very different conclusion, showing no difference in the misinformation effect between nominal and collaborative pairs. More clear-cut evidence has been recently reported by Rossi-Arnaud, Spataro, Bhatia, and Cestari (2019), who used the Gudjonsson Suggestibility Scale to measure *interrogative suggestibility* (Gudjonsson, 1997). Pairs of participants listened to a short story, provided immediate and delayed free recalls (after a 30-min interval), and then answered a series of misleading questions before and after having received a negative feedback in which they were told that some of their responses were wrong. The classical effect of collaborative inhibition was not observed in the immediate and delayed free recall tasks. However, collaborative pairs produced a lower number of confabulated elements in the recall tasks, were less likely to accept leading questions (both before and after receiving the negative feedback), and exhibited lower levels of total suggestibility, compared to nominal dyads. Rossi-Arnaud et al. (2019) concluded that the members of collaborative dyads were more likely to cross-check the accuracy of their partners' responses and provide corrective feedback after the recall of inaccurate details (Ross et al., 2008).

The present study was aimed at determining whether the elimination of false memories through mutual error checking could reduce the suggestibility of collaborative dyads in a condition in which the probability of memory errors was increased by asking participants to confabulate responses to false-event questions during an initial forced fabrication phase. To this purpose, we modified a paradigm previously described by Memon, Zaragoza, Clifford, and Kidd (2010), in which college students witnessed a filmed event and were later forced to fabricate answers to misleading questions. One week later, they were given a yes/no recognition test designed to assess false assents to statements reporting either the same details that participants had confabulated during the previous session or novel, non-confabulated details (called "yoked control items", because they provided a base rate of spontaneous false assents). Our experiments adopted the same paradigm, with one relevant exception that was justified by the use of a group setting (see the Method section for a more detailed explanation). In particular, in the final yes/no recognition task the specific details produced by participants during the forced fabrication phase were not presented: rather, participants answered misleading statements asking them to determine whether the related false events had occurred or not in the videoclip. We reasoned that the confabulation of incorrect details during the forced fabrication phase should increase participants' tendency to incorrectly assent to later misleading statements and this should result in a greater incidence of disagreements in collaborative dyads – two assumptions that are well supported by previous studies (Ackil & Zaragoza, 1998; Gombos, Pezdek, & Haymond, 2012; Pezdek, Lam, & Sperry, 2009; Pezdek, Sperry, & Owens, 2007; Stolzenberg & Pezdek, 2013; Zaragoza, Payment, Ackil, Drivdahl, & Beck, 2001). The question of interest was whether this increased suggestibility could be reduced by the use of a recall-to-reject strategy in collaborative dyads.

Besides including a forced fabrication phase, we also manipulated the delay between the presentation of the video clip and the yes/no recognition task. Previous studies showed that the effects of collaboration, whether negative or positive, are reduced by long study-test delays (e.g., Congleton & Rajaram, 2011; Takahashi & Saito, 2004). It is thus important to ascertain whether the positive effects of collaborative remembering observed by Rossi-Arnaud et al. (2019) after a 30-min interval persist when responses to misleading questions are collected 1 h (Experiment 1) or 1 week (Experiment 2) after the presentation of the original event. In the framework outlined above, a decrease in the positive effects of collaboration might occur because, after a 1-week delay, participants might be less able to recall the correct source of the inaccurate details suggested in the misleading statements and thus less likely to provide corrective feedback.

2. Experiment 1

Experiment 1 addressed the question of whether collaboration could reduce the proportions of false assents to misleading statements in a yes/no recognition task when the delay from the presentation of the video-clip was relatively short (1 h) and participants self-generated inaccurate responses to false-event questions in a previous forced fabrication phase. Following previous studies in this field, the performance of collaborative dyads was compared with the performance of nominal dyads – the latter representing the sum of the non-redundant responses provided by two individuals that work in a group setting but do not collaborate. As mentioned above, this method of scoring the performance of non-interacting individuals is based on the Lorge and Solomon (1955) additive model and has been previously adopted to test for the presence of collaborative inhibition in the recall of correct details (Basden et al., 1997; Weldon & Bellinger, 1997). In addition, it has been also used to examine the error pruning advantages associated with collaboration: here, the key question was whether the number of incorrect details recalled by collaborative groups was significantly lower than that expected by pooling together the non-redundant outputs of individuals working in nominal groups (Ross et al., 2004, 2008; Takahashi, 2007).

2.1. Method

2.1.1. Participants

A total of 30 pairs of graduate psychology students volunteered to participate in Experiment 1 ($N = 60$), of which 15 were randomly assigned to the nominal condition and the other 15 were assigned to the collaborative condition. Overall, the mean age and education (number of years of formal instruction) of the participants were 24.3 (range: 22–31) and 17.4 (range: 16–21) years, respectively. The two groups did not differ in terms of age and education. As concerns gender, there were 7 mixed pairs (female/male) in both the nominal and collaborative conditions; all other pairs were composed by two female students. This study was conducted in accordance with the recommendations of the International Review Board, Department of Psychology, Sapienza University, and in accordance with the guidelines reported in the Declaration of Helsinki. Each participant provided a written informed consent to participate in this study.

2.1.2. Materials

We selected an 11 minute video clip, taken from the movie 'Dog Day After-noon' (Sidney Lumet, 1975) and dubbed in Italian. The video clip depicts an attempted bank robbery by three men (Sonny, Stevie and Sal) and has been used in several previous studies (Memon et al., 2010; Natali, Marucci, & Mastroberardino, 2012; Rossi-Arnaud, Spataro, & Geraci, 2018). Besides increasing ecological validity (Ihlebæk, Løve, Eilertsen, & Magnussen, 2003), the choice of the material contributed to enhance the frequency of errors, since recalling a high number of details from a longer video-clip can be considered much more difficult than recalling a limited number of details from a short video, especially after a 1-week delay. Furthermore, the recall of the witnessed events from a long video should be largely based on the use of story schema (Kintsch, 1978); this is important, because the false events introduced in the misleading questions were constructed to be consistent with the norms that specify what is to be expected in certain situations. Thus, a higher reliance on story schema should increase the participants' tendency to accept misleading questions in the yes/no recognition task.

For the forced fabrication phase, we translated in Italian the eight false-event questions reported by Memon et al. (2010). The eight questions were divided into two sets called 'Version A' and 'Version B', to counterbalance their use in the forced fabrication phase and in the final yes/no recognition task. All other questions presented either in the forced fabrication phase (9) or in the final recognition task (9) were taken from Natali et al. (2012) and Memon et al. (2010) and have been previously used by Rossi-Arnaud et al. (2018).

2.1.3. Procedure

The entire procedure was adapted from Memon et al. (2010) and included an encoding phase, an immediate free recall test, a forced fabrication phase and a 1-hour delayed recognition task (see Fig. 1).

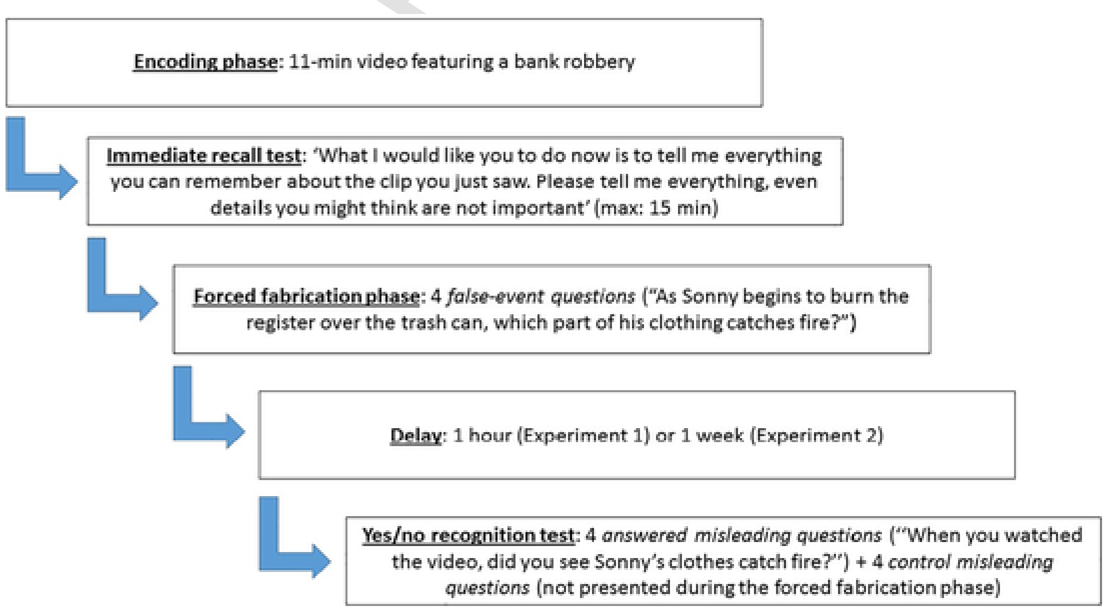

Fig. 1. Schematic illustration of the procedure followed in Experiments 1 and 2.

Upon arrival at the laboratory, the members of each pair were accompanied in a quiet testing room and asked to sign two independent consent forms in which they were informed that the general aim of the experiment was to examine their memory for a bank robbery. At this point, they were shown the video clip. In both the nominal and collaborative conditions, the participants were forewarned not to interact in any way during the 11-min presentation (the experimenter remained in the room to guarantee adherence to the instructions).

Immediately after the encoding phase, participants were given an immediate free recall test. The instructions, taken from Mastroberardino, Natali, and Candel (2012) and visually presented on a computer screen, were as follows: 'What I would like you to do now is to tell me everything you can remember about the clip you just saw. Please tell me everything, even details you might think are not important'. The members of collaborative pairs were given a single response sheet and were required to collaborate throughout the recall task (one participant was randomly chosen to write the output). Following Weldon and Bellinger (1997), no specific instructions were given regarding the modalities of collaboration and turn-taking was not explicitly required, favouring free-flowing discussion. In contrast, the members of nominal pairs were given separate response sheets and instructed to write two independent summaries. They were not allowed to interact (again, the experimenter remained in the room to guarantee that instructions were correctly followed). In both conditions, the maximum time allowed to complete the task was 15 min.

Next, participants engaged in the forced fabrication phase. They were requested to answer 13 questions, of which 9 were 'true-event' questions (referring to details that participants had seen in the video clip) and 4 were 'false-event' questions (referring to details that, although plausible, did not appear in the video clip). For example, one false-event question asked: "*As Sonny begins to burn the register over the trash can, which part of his clothing catches fire?*", although Sonny's clothes never caught fire in the video. These questions were shown on the computer screen and participants provided written responses. The instructions specified that participants had to provide a response to every question, even if they had to guess; thus, they were not given the option to use a "don't know" answer (Pezdek et al., 2007, 2009). As expected, participants were sometimes reluctant to answer 'false-event' questions: in these cases, they were prompted to "just provide their best guesses" (Memon et al., 2010, p.109). Before collecting the response sheets, the experimenter checked that all questions were given a valid response. As in the free recall phase, collaborative groups were given a single response sheet, were instructed to work together to provide joint answers and the instructions did not specify how to resolve disagreements (Weldon & Bellinger, 1997); in contrast, the members of nominal groups were given separate response sheets and told to provide independent answers.

The forced fabrication phase was followed by a 1-hour delay, after which participants were asked to come back to the laboratory and complete a yes/no recognition test assessing their memory for the video clip. As in the study by Memon et al. (2010), the questionnaire included a total of 17 questions, that were displayed one by one on the computer screen. They were divided as follows: (a) 4misleadingstatementsrelated to the false-event questions to which participants had provided a response during the previous forced fabrication phase (hereafter referred to as *answered misleading questions*: "*When you watched the video, did you see Sonny's clothes catch fire?*"), (b) 4 misleading statements for which the related false-event questions were not presented during the forced fabrication phase (hereafter referred to as *control misleading questions*), (3) 5 true-event statements to which participant had provided a response during the forced fabrication phase, (4) 2true-event statements referred to details that were in the video but were not mentioned during the forced fabrication phase, and (5) 2 additional misleading statements referred to details that never occurred in the video clip. Collaborative pairs were again requested to provide joint yes/no responses, whereas the members of nominal pairs gave independent answers.

As mentioned above, our procedure in the latter phase differed from that described by Memon et al. (2010) in one relevant aspect: namely, the specific responses confabulated by participants during the forced fabrication phase were not represented in the misleading questions included in the yes/no recognition test. This choice was motivated by the fact that we worked with dyads, rather than with single participants: the implication was that different members might have provided different responses to the false-event questions presented in the forced fabrication phase. For example, in a nominal group, the two participants working independently could potentially confabulate different items when answering a false-event question like "*As Sonny begins to burn the register over the trash can, which part of his clothing catches fire?*": participant A could write "tie", whereas participant B could write "jacket". In this case, considering that all the misleading questions in the yes/no recognition task were visually displayed on the computer screen, the presentation of one of the two confabulated details would have probably resulted in unwanted differences in the probability to provide false assents – i.e., a misleading question like "*When you watched the video, did you see Sonny's tie catch fire?*" would have triggered a false assent from participant A, but not from participant B. Similar problems applied to collaborative dyads, since the reported answer to a misleading question could correspond to the answer provided by only one of the two members, which was later chosen as the group response. To avoid these confounds, we opted for asking general misleading questions that did not include the specific details confabulated in the forced fabrication phase ("*When you watched the video, did you see Sonny's clothes catch fire?*").

2.1.4. Data coding

The written reports obtained in the free recall task were coded by a trained rater who was blind to the study aims, following a scheme developed by Memon et al. (1997; see also Rossi-Arnaud et al., 2018), which distinguishes between information pertaining to people (*person details*), actions (*action details*), objects (*object details*), and the environment in which the scene took place (*location details*). For example, in the sentence 'The man goes into the bank', the raters coded 'The man' as a person detail, 'goes' as an action detail, and 'into the bank' as a location detail (Natali et al., 2012). *The choice to assess the recall of different categories of details was based on* previous studies showing that the central, thematically important details of emotional events are remembered more accurately than peripheral details (Christianson, 1992; Christianson & Safer, 1996). Wessel, Zandstra, Hengeveld, and Moulds (2015) proposed that, because of these differences in relevance, central details might be relatively impervious to the inhibitory effects of collaboration. Along the same direction, a study by Dalton and Daneman (2006) showed that the participants' susceptibility to misleading suggestions was higher for peripheral than for central features.

Each recalled detail was then classified as *correct*, *incorrect* or *confabulated*. Incorrect information refers to details presented in the video clip but wrongly described by the participants (i.e., distorted), whereas confabulated information refers to details never presented in the video (Mastroberardino et al., 2012). A randomly selected subset (25%) of transcribed protocols was coded by a second trained rater, who was also blind to the study aims. The mean kappa values for inter-rater agreement, averaged across Experiments 1 and 2, were $\kappa = 0.76$ for correct details, $\kappa = 0.82$ for incorrect details, and $\kappa = 0.79$ for confabulated details. The nominal performance was computed by summing, for each category, the correct, incorrect and confabulated details recalled by the two members, but counting redundant items only once (Basden et al., 1997; Weldon & Bellinger, 1997).

During the forced fabrication phase, participants were requested to answer a total of nine true-event questions – i.e., questions concerning details that were originally presented in the video. Correct responses to these questions were scored in the same way as the details of the free recall task. In particular, the nominal performance was computed by summing the correct responses of the two collaborating members but counting redundant responses only once.

The primary interest in our study was whether collaborative remembering reduced the proportions of false assents in the yes/no recognition task. Following Memon et al. (2010), statistical analyses took into account both the proportions of incorrect assents to the four *answered misleading questions* (for which the corresponding false-event questions had been responded by participants during the forced fabrication phase) and the proportions of incorrect assents to the four *control misleading questions* (for which the corresponding false-event questions had not been presented during the forced fabrication phase). The latter measure represented the participants' base rate of false assents. The performance of nominal groups was again computed by summing the false assents provided by the members of each pair and counting redundant responses only once (this was done separately for answered and control misleading questions).

2.2. Results

2.2.1. Free recall performance

The raw number of correct person, action, object, and location details recalled by nominal and collaborative pairs is reported in Fig. 2 (data for Experiments 1 and 2 can be freely accessed at the following link: https://osf.io/23ygf/?view_only=2c6629cfd86f495aab8c393a5d050d13).These scores were analyzed with a MANOVA followed by univariate ANOVAs, considering Condition (collaborative vs. nominal pairs) as the between-subjects factor. The results showed a significant multivariate effect of Condition, $\lambda = 0.39$, $F(4, 25) = 9.91$, $p < 0.001$, $\eta_p^2 = 0.61$.The follow-up univariate analyses confirmed that the difference between collaborative and nominal dyads was significant for person details, $F(1, 28) = 35.01$, $MSE = 46.92$, $p < 0.001$, $\eta_p^2 = 0.55$, object details, $F(1, 28) = 17.97$, $MSE = 58.09$, $p < 0.001$, $\eta_p^2 = 0.39$, action details, $F(1, 28) = 20.61$, $MSE = 461.15$, $p < 0.001$, $\eta_p^2 = 0.42$, and location details, $F(1, 28) = 21.75$, $MSE = 28.76$, $p < 0.001$, $\eta_p^2 = 0.43$. In line with the results typically reported in previous studies (Rajaram, 2011; Rajaram & Pereira-Pasarin, 2010), the performance of collaborative pairs was always lower than that of nominal pairs, thereby replicating the standard negative effect of collaborative inhibition.

Two similar MANOVAs were conducted on the mean number of incorrect and confabulated details reported by collaborative and nominal pairs. As concerns incorrect details, the multivariate effect of Condition was not significant, $\lambda = 0.79$, $F(4, 25) = 1.65$, $p = 0.19$, $\eta_p^2 = 0.20$. However, the follow-up univariate analyses revealed significant differences for action details, $F(1, 28) = 5.21$, MSE = 1.84, $p = 0.030$, $\eta_p^2 = 0.15$, and location details, $F(1, 28) = 4.26$, MSE = 0.19, $p = 0.048$, $\eta_p^2 = 0.13$, with collaborative pairs reporting a lower number of incorrect details than nominal pairs in both cases – M(nominal) = 2.13 vs. M(collaborative) = 1.00 for action details, and M(nominal) = 0.47 vs. M(collaborative) = 0.13 for location details. Regarding confabulated details,[1] neither the multivariate effect of Condition, $\lambda = 0.92$, $F(3, 26) = 0.67$, $p = 0.57$, $\eta_p^2 = 0.07$, nor the follow-up univariate analyses, $F(1, 28) < 2.15$, MSE = 0.062, $p > 0.15$, $\eta_p^2 < 0.07$, reached the significance level.

[1] Note that confabulated details for the 'person' category were produced neither by collaborative nor by nominal pairs: as a consequence, this category was eliminated from statistical analyses.

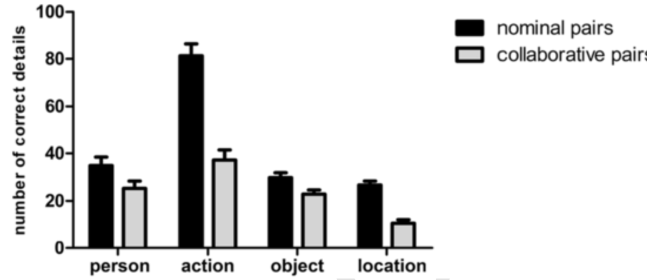

Fig. 2. Mean number of correct person, action, object and location details recalled by nominal and collaborative pairs in Experiment 1 (1-hour delay). Bars represent standard errors.

2.2.2. Forced fabrication phase

To determine whether collaborative inhibition persisted in the cued-recall task, the responses of collaborative and nominal groups were analyzed with a *t*-test for independent samples. As predicted by the RSD hypothesis (Finlay, Hitch, & Meudell, 2000), the magnitude of collaborative inhibition was substantially reduced, with the proportions of correct responses for nominal dyads ($M = 0.90$) being numerically, but not significantly, higher than those of collaborative dyads ($M = 0.83$), $t(28) = 1.72$, $p = 0.096$.

2.2.3. Yes/no recognition test

The raw proportions of false assents to answered and control misleading questions provided by nominal and collaborative dyads are illustrated in Table 1. They were analyzed via a mixed ANOVA, considering Condition (collaborative vs. nominal pairs) as the between-subjects factor and Question Type (answered vs. control) as the within-subjects factor. The results showed a significant main effect of Question Type, $F(1, 28) = 35.45$, MSE = 0.032, $p < 0.001$, $\eta_p^2 = 0.56$, indicating that false assents to answered misleading questions ($M = 0.53$) exceed the base rate of false assents to control misleading questions ($M = 0.25$). The main effect of Condition was also significant, $F(1, 28) = 22.89$, MSE = 0.050, $p < 0.001$, $\eta_p^2 = 0.45$, confirming that collaborative groups ($M = 0.25$) produced a lower number of false assents, as compared to nominal groups ($M = 0.52$). Interestingly, the two-way interaction was not significant, $F(1, 28) = 0.03$, MSE = 0.032, $p = 0.85$, $\eta_p^2 = 0.001$, suggesting that retrieval collaboration reduced false assents to both answered and control questions.

In summary, Experiment 1 revealed that working in a collaborative dyad significantly reduced the participant's ability to freely recall the details of a video clip depicting a bank robbery. This finding is in line with previous studies showing that collaboration inhibits recalling the details of emotional events (Bärthel, Wessel, Huntjens, & Verwoerd, 2017; Kensinger, Choi, Murray, & Rajaram, 2016; Vredeveldt, Hildebrandt, & van Koppen, 2016; Wessel et al., 2015; Yaron-Antar & Nachson, 2006). Therefore, when it comes to obtaining the maximum amount of information about a crime in free re-

Table 1
Proportions of false assents to answered and control misleading questions provided by nominal and collaborative pairs in Experiment 1 (1-hour delay) and Experiment 2 (1-week delay).

Type of dyad	Nominal	Collaborative
Experiment 1 (1-hour delay)		
Answered misleading questions	0.66 (0.20)	0.38 (0.22)
Control misleading questions	0.38 (0.20)	0.11 (0.15)
Experiment 2 (1-week delay)		
Answered misleading questions	0.79 (0.26)	0.61 (0.24)
Control misleading questions	0.45 (0.21)	0.30 (0.23)

call tasks, it would be preferable for police officers and investigators to avoid collaboration between eyewitnesses.

At the same time, the results obtained in the yes/no recognition task indicate that collaborative dyads were less likely to provide false assents to misleading statements reporting information that was not presented in the original video clip. Thus, when it comes to the tendency to resist to misleading questions suggesting inaccurate details, collaboration appears to offer a substantial advantage over individual retrieval. The present data confirm the conclusions reported by Rossi-Arnaud et al. (2019) and extend them to a slightly longer study-test interval (1 h instead of 30 min). A second important novelty with respect to previous literature is that this advantage was observed even when participants had generated inaccurate responses to related false-event questions during the forced fabrication phase. Thus, collaborative remembering can exert a positive effect in all the conditions in which eyewitnesses are pressed to answer misleading questions.

3. Experiment 2

Experiment 2 aimed at determining whether the results of Experiment 1 could be extended to a longer delay between the presentation of the video clip and the yes/no recognition task. To this purpose, we used a 1-week delay, like in the Memon et al. (2010) study.

3.1. Method

3.1.1. Participants

A total of 26 pairs of psychology students volunteered to participate in Experiment 2 ($N = 52$), of which 11 were assigned to the nominal condition and the other 15 were assigned to the collaborative condition. Overall, the mean age and education (number of years of formal instruction) of participants were 23.7 (range: 18–33) and 18.1 (range: 16–23) years, respectively. The two groups did not differ in terms of age and education. As concerns gender, there was 1mixedpair (female-male) in the nominal condition and 9 in the collaborative condition; all other pairs comprised two female students, except for one dyad in the collaborative condition, which included two males. As in Experiment 1, each participant signed an informed consent before participating in the study.

3.1.2. Materials

Study materials (the 11-min video clip and the questionnaires used during the forced fabrication phase and the yes/no recognition task) were the same as in Experiment 1.

3.1.3. Procedure

The procedure was in all respects equal to that described in Experiment 1, except for the fact that the delay between the forced fabrication phase and the yes/no recognition task was extended to 1 week (Memon et al., 2010).

3.2. Results

3.2.1. Free recall performance

The raw number of correct person, action, object, and location details recalled by nominal and collaborative pairs is reported in Fig. 3.

These scores were analyzed with a MANOVA followed by univariate ANOVAs, considering Condition (collaborative vs. nominal pairs) as the between-subjects factor. The results showed a significant multivariate effect of Condition, $\lambda = 0.26$, $F(4, 21) = 14.45$, $p < 0.001$, $\eta_p^2 = 0.73$.The follow-up univariate analyses confirmed that collaborative pairs performed worse than nominal pairs on person details, $F(1, 24) = 4.10$, $MSE = 141.84$, $p = 0.054$, $\eta_p^2 = 0.15$, object details, $F(1, 24) = 6.56$, $MSE = 46.72$, $p = 0.017$, $\eta_p^2 = 0.22$, action details, $F(1, 24) = 44.78$, $MSE = 275.27$, $p < 0.001$, $\eta_p^2 = 0.65$, and location details, $F(1, 24) = 52.83$, $MSE = 30.98$, $p < 0.001$, $\eta_p^2 = 0.69$. Thus, the standard effect of collaborative inhibition was replicated after a 1-week delay.

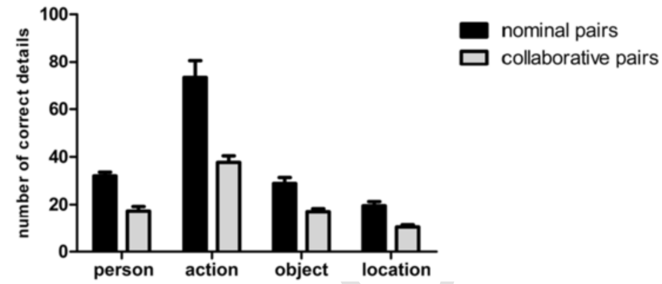

Fig. 3. Mean number of correct person, action, object and location details recalled by nominal and collaborative pairs in Experiment 2 (1-week delay). Bars represent standard errors.

Two similar MANOVAs were conducted on the number of incorrect and confabulated details. For incorrect details, the multivariate effect of Condition did not reach statistical significance, $\lambda = 0.82$, $F(4, 21) = 1.10$, $p = 0.38$, $\eta_p^2 = 0.17$, although the follow-up univariate analyses revealed that collaborative pairs ($M = 0.13$) produced a lower number of incorrect action details, as compared to nominal pairs ($M = 0.64$), $F(1, 24) = 4.65$, $MSE = 0.35$, $p = 0.041$, $\eta_p^2 = 0.16$. For confabulated details, neither the multivariate effect of Condition, $\lambda = 0.79$, $F(4, 21) = 1.34$, $p = 0.28$, $\eta_p^2 = 0.20$, nor the follow-up univariate analyses, $F(1, 24) < 1.56$, MSE $= 0.072$, $p > 0.22$, $\eta_p^2 < 0.06$, turned out to be significant.

3.2.2. Forced fabrication phase

Regarding the true-event questions answered during the forced fabrication phase, the responses of collaborative and nominal pairs were analyzed with a *t*-test for independent samples. As in Experiment 1, there was no significant difference between the proportions of correct responses for nominal ($M = 0.77$) and collaborative dyads ($M = 0.70$), $t(24) = 1.31$, $p = 0.20$.

3.2.3. Yes/no recognition test

The raw proportions of false assents to answered and control misleading questions are illustrated in Table 1.A mixed ANOVA, considering Condition (collaborative vs. nominal pairs) as the between-subjects factor and Question Type (answered vs. control) as the within-subjects factor, found a pattern of results similar to that reported in Experiment 1. The main effects of Condition and Question Type were both significant, $F(1, 24) = 5.72$, $MSE = 0.062$, $p = 0.025$, $\eta_p^2 = 0.19$ and $F(1, 24) = 24.23$, $MSE = 0.057$, $p < 0.001$, $\eta_p^2 = 0.50$, indicating that the mean proportions of false assents were lower for collaborative than for nominal pairs ($M = 0.46$ vs. $M = 0.63$), and greater for answered than for control misleading questions ($M = 0.71$ vs. $M = 0.38$). The two-way interaction was not significant, $F(1, 24) = 0.03$, $MSE = 0.057$, $p = 0.86$, $\eta_p^2 = 0.001$, again suggesting that collaborative remembering reduced false assents to answered and control misleading questions to the same extent.

To summarize, Experiment 2 provided a replication of the collaborative inhibition found in the immediate free recall task; more importantly, it also showed that collaborative remembering reduced the false assents to answered and control misleading questions even when the delay between the encoding phase and the yes/no recognition test was extended from 1 h to 1 week. This is important from a forensic point of view because witnesses are often questioned after long delays from the original event (Pezdek et al., 2007) and their testimonies can undergo quantitative and qualitative changes (Poole & White, 1993).

4. General discussion

To recap, the present study examined the effects of collaborative remembering on the recall of the details of a video clip illustrating a bank robbery and critically assessed the question of whether collaboration could reduce eyewitnesses' suggestibility when the misleading questions were presented after a long delay (1 week) and participants were forced to fabricate wrong details in a previous confabulation phase. The results were clear-cut and can be summarized as follows. First, collaboration had a clear negative effect on free recall performance in both experiments, confirming that collaborative inhibition represents a robust phenomenon in eyewitness memory tasks (Vredeveldt et al., 2016; Wessel et al., 2015; Yaron-Antar & Nachson, 2006). Second, and most central to our purposes, collaboration reduced the participants' tendency to give in to misleading questions, irrespective of whether participants had generated a response to the corresponding false-event questions during a previous forced fabrication phase and irrespective of the delay between the encoding phase and the yes/no recognition task.

As mentioned above, previous studies using associative recognition tasks showed that participants working in collaborative groups of two or three persons use a recall-to-reject strategy to avoid judging false alarms as 'old' (studied) words (Clark et al., 2000, 2006). Since memory errors are often unique to an individual, the group partners can exercise a quality control by inspecting the accuracy of each response and providing corrective feedback (Ross et al., 2004, 2008). So, for example, if a participant incorrectly recalls that the rearranged triple ABF was studied at encoding, the collaborator is then able to modify her/his response by recalling that the correct combination was ABC (not ABF). The results reported by Clark et al. (2006) supported the use of this mutual error checking strategy in collaborative groups; however, they also showed that obtaining confirmative evidence was difficult in the standard recognition task, because the optimal combination of events (disagreements, plus recall of the correct triple) was rather infrequent. Later studies have therefore relied on different paradigms, that were expected to increase the overall number of errors and disagreements. Takahashi (2007), for example, investigated the effects of collaboration in the Deese-Roediger-McDermott (DRM) paradigm (Roediger & McDermott, 1995; see also Thorley & Dewhurst, 2007, 2009). In the DRM paradigm, participants independently study lists of words that are each semantically associated to a non-studied critical lure which is not presented. The typical finding is that false memory for these critical lures is similar to the veridical recall of studied words, and participants often claim to vividly remember the critical lures they falsely recognize. Takahashi (2007) reported that collaborative pairs of friends and non-friends recalled fewer correct words and fewer critical non-presented lures than individuals working in nominal pairs. The author argued that collaboration induced the dyad's members to focus more attention on their memories and thus to evaluate them more carefully.

In the present experiments, we tried to increase the probability of errors in two different ways. First, we asked participants to generate wrong response to false-event questions during a confabulation phase that preceded the final yes/no recognition task. In line with our expectations, this procedure significantly increased suggestibility, as shown by the fact that the members of both collaborative and nominal pairs were more likely to incorrectly assent to misleading questions for which they had confabulated a false detail in the previous phase (as compared to control misleading questions). Second, in Experiment 2 we imposed a relatively long delay between the confabulation phase and the recognition task (1 week). We reasoned that participants should be less able to recall the source of confabulated details, and therefore more likely to provide wrong responses to misleading questions, after a 1-week delay. The data reported in Table 1 are consistent with our predictions, since the proportions of false assents increased from Experiment 1 (which used a 1-hour delay) to Experiment 2. By adopting these two methodological modifications, our study had a greater chance to highlight the error pruning benefits of collaborative remembering (Ross et al., 2004, 2008; Takahashi, 2007). In agreement, the analysis of the performance in the yes/no recognition task revealed that the proportions of false assents were consistently lower in collaborative than in nominal dyads, even after a 1-week delay from the encoding phase. Such a reduction aligns well with earlier evidence indicating that collaboration diminished the negative effects of misinformation (Karns et al., 2009) and lowered interrogative suggestibility (Rossi-Arnaud et al., 2019).

Taken together with the results of previous studies, our data provide further evidence in support of the hypothesis that working in a collaborative group increased the participants' tendency to cross-check, and eventually correct, the responses provided by other members (Harris, Barnier, & Sutton, 2012; Ross et al., 2008). In agreement with this conclusion, Harris et al. (2012) analyzed recall completeness (the percentage of studied items correctly recalled) and inaccurate recall (the percentage of intrusions) in three different groups: consensus collaborative groups (in which the members had to reach consensus on each response), turn-taking collaborative groups (in which the members alternated in producing their responses) and nominal groups. Regarding recall completeness, the results replicated the standard collaborative inhibition, showing that both consensus and turn-taking groups recalled less studied items than nominal groups. For inaccurate recall, it was found that consensus groups made fewer intrusions than turn-taking and nominal groups, which did not differ between them. Importantly, Harris et al. (2012) recorded the conversations between the members of collaborative groups and were therefore able to score the inaccurate stimuli that were mentioned during the discussion but were subsequently discounted. When these "inclusive scores" were analyzed, the between-group differences disappeared, suggesting that the members of consensus groups mentioned a similar number of incorrect items as did the members of nominal groups; however, these items were later rejected by the group and were not included in the written output. A post-hoc analysis of the recall strategies demonstrated that this occurred because a higher number of participants in consensus groups (than in turn-taking groups) reported that they were checking the accuracy of the words retrieved by their collaborators. A similar conclusion has been reached by Ross et al. (2008), who asked younger and older participants to recall items from six briefly exposed household scenes, either alone or in collaboration with their spouses. The analysis of the conversations occurring in the collaborative condition showed that the younger and older collaborating couples eliminated about 50% and 30% of their errors while discussing potential answers with their spouses. Although we did not record the conversations occurring in the collaborative dyads, the results illustrated by Harris et al. (2012) and Ross et al. (2008) provide a plausible explanation for the present findings, by suggesting that, in the yes/no recognition task, the members of collaborative dyads produced the same number of false assents as did the members of nominal dyads, and that collaboration increased their ability to reject errors through discussion (Harris et al., 2012; Ross et al., 2008).

Regarding the recall of correct details in the free recall task, our findings showing a strong collaborative inhibition in both experiments are apparently in contrast with those reported by Takahashi and Saito (2004). In this study, participants read a story composed of 30 sentences and took two successive free recall tests. The first test was individual for all participants; in the second test, some participants were assigned to recall the material in collaborative dyads, whereas all other participants continued to recall on their own and their outputs were later pooled to estimate the nominal performance. The results showed that, in the second test, collaborative pairs recalled significantly less information than nominal pairs when the delay between the encoding and test phases was short (some minutes), but not when it was long (1 week). The authors proposed that, when the delay was short, the participants' recall was strongly dependent on the use of individual retrieval strategies; in this case, hearing the output of the other member had a disruptive effect because the participant was forced to adopt a different, less effective, retrieval strategy (Basden et al., 1997). However, when the delay was long, the stored retrieval strategies were likely to be degraded, meaning that participants could not use them to optimize individual performance and were therefore less negatively affected by hearing the collaborator's output. Our findings do not support this explanation, because collaborative inhibition in the free recall task was robust and significant in Experiment 2, which used a 1-week delay (the same as in Takahashi & Saito, 2004). Furthermore, the effects size associated to the Condition manipulation was numerically higher in Experiment 2 ($\eta_p^2 = 0.73$) than in Experiment 1 ($\eta_p^2 = 0.61$), providing no evidence for a decrease in the strength of the collaborative inhibition effect. Clearly, there are several methodological differences between the two studies that can account for the discrepant results. Unlike Takahashi and Saito (2004), participants in our study did not take an individual free recall test before collaborating in the second recall test. The use of this procedure might have promoted a better idiosyncratic organization of the study material, a factor which in turn is known to protect against the negative effects of collaborative inhibition (Congleton & Rajaram, 2011). Another potential factor might be represented by the nature of the material used during the study phase. We presented participants with a relatively long video (11-min) describing a series of events, whereas Takahashi and Saito (2004) employed a short, written story. The meta-analysis by Marion and Thorley (2016) showed that the size of collaborative inhibition was moderated by material type, with studies using uncategorized stimuli yielding a stronger inhibition than studies using story-like formats. Regardless of the correct explanation, additional research is needed to further elucidate the question of whether collaborative retrieval produces a persistent inhibition of memory performance after long study-test delays.

Three other points should be noted about the participants' performance in free- and cued-recall tasks. First, in contrast to our predictions, collaborative inhibition was not moderated by the saliency of the recalled details in both experiments. That is, collaborative dyads suffered not only when recalling peripheral details (for example, location details), but also when recalling central details (namely, person and action details), which were expected to be more resistant to collaborative inhibition (Wessel et al., 2015). Similar results have been however reported by Bärthel et al. (2017). Second, collaborative dyads produced a lower number of incorrect action and location details, confirming the idea that individuals make fewer errors when they remember together than when they remember alone (Ross et al., 2004, 2008; Rossi-Arnaud et al., 2019; Vredeveldt, van Deuren, & van Koppen, 2019; Wessel et al., 2015). Third, the negative effects of collaborative inhibition were eliminated when participants answered the true-event questions assessing the recall of correct information during the forced fabrication phase. This is consistent with the notion that the disrupting impact of collaborative retrieval is less apparent in memory tasks having a low reliance upon self-initiated strategies (Finlay et al., 2000).

5. Conclusions

In summary, the present study showed having two or more eyewitnesses collaborating in the recall of a witnessed crime can have beneficial effects on memory (Vredeveldt et al., 2016; Vredeveldt, Groen, Ampt, & van Koppen, 2017). Specifically, collaborative dyads were less likely to provide false assents to misleading statements, regardless of whether participants had produced inaccurate responses to false-event questions in a previous forced fabrication phase occurring 1 h or 1 week earlier. This is especially important in those contexts in which the interview procedures might suggestively influence eyewitness memory. For example, if an interviewer believes that an eyewitness observed a given event, then he/she may press the eyewitness to confabulate inaccurate answers to misleading questions. In this way, the interviewer might induce the eyewitness to self-generate information that were not included in the original episode (Gombos et al., 2012; Pezdek et al., 2007). Our data indicate that collaborative remembering can be effectively used to reduce this negative effect. The present findings are also relevant in those contexts in which the same event is experienced by a group of witnesses that may discuss it before being interviewed by the police. For example, if one of the witnesses misremembers an event detail, the collaborative discussion with other witnesses may reduce the probability to incorporate the erroneous information. Moreover, these effects hold true even when the delay from the original event is relatively long.

CRediT authorship contribution statement

Clelia Rossi-Arnaud: Conceptualization, Writing - original draft, Methodology, Data curation, Formal analysis, Writing - review & editing. **Pietro Spataro:** Conceptualization, Writing - original draft, Methodology, Data curation, Formal analysis, Writing - review & editing. **Divya Bhatia:** Methodology, Data curation, Formal analysis, Writing - review & editing. **Fabrizio Doricchi:** Methodology, Data curation, Formal analysis, Writing - review & editing. **Serena Mastroberardino:** Writing - review & editing. **Vincenzo Cestari:** Conceptualization, Writing - original draft, Methodology, Data curation, Formal analysis, Writing - review & editing.

Declaration of competing interest

None.

References

Ackil, J. K., & Zaragoza, M. S. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. Developmental Psychology, 34, 1358–1372.

Barber, S. J., Harris, C. B., & Rajaram, S. (2015). Why two heads apart are better than two heads together: Multiple mechanisms underlie the collaborative inhibition effect in memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 41, 559–566.

Bärthel, G. A., Wessel, I., Huntjens, R. C., & Verwoerd, J. (2017). Collaboration enhances later individual memory for emotional material. Memory, 25, 636–646.

Basden, B. H., Basden, D. R., Bryner, S., & Thomas, R. L., III (1997). A comparison of group and individual remembering: Does collaboration disrupt retrieval strategies? Journal of Experimental Psychology. Learning, Memory, and Cognition, 23, 1176–1191.

Brennan, A. A., & Enns, J. T. (2015). What's in a friendship? Partner visibility supports cognitive collaboration between friends. PLoS One, 10(11), e0143469.

Brennan, A. A., & Enns, J. T. (2015). When two heads are better than one: Interactive versus independent benefits of collaborative cognition. Psychonomic Bulletin & Review, 22, 1076–1082.

Brennan, A. A., & Enns, J. T. (2016). But is it social? How to tell when groups are more than the sum of their members. Behavioral and Brain Sciences, 39, 17–18.

Clark, N. K., Stephenson, G. M., & Kniveton, B. (1990). Social remembering: Quantitative aspects of individual and collaborative remembering by police officers and students. British Journal of Psychology, 81, 73–94.

Clark, S., Hori, A., Putnam, A., & Martin, T. P. (2000). Group collaboration in recognition memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 1578–1588.

Clark, S. E., Abbe, A., & Larson, R. P. (2006). Collaboration in associative recognition memory: Using recalled information to defend "new" judgments. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(6), 1266.

Congleton, A. R., & Rajaram, S. (2011). The influence of learning methods on collaboration: Prior repeated retrieval enhances retrieval organization, abolishes collaborative inhibition, and promotes post-collaborative memory. Journal of Experimental Psychology: General, 140(4), 535–551.

Dalton, A. L., & Daneman, M. (2006). Social suggestibility to central and peripheral misinformation. Memory, 14(4), 486–501.

Finlay, F., Hitch, G. J., & Meudell, P. R. (2000). Mutual inhibition in collaborative recall: Evidence for a retrieval-based account. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1556–1567.

Gombos, V., Pezdek, K., & Haymond, K. (2012). Forced confabulation affects memory sensitivity as well as response bias. Memory & Cognition, 40, 127–134.

Gudjonsson, G. H. (1997). The Gudjonsson Suggestibility Scales manual. Hove, East Sussex: Psychology Press.
Harris, C. B., Barnier, A. J., & Sutton, J. (2012). Consensus collaboration enhances group and individual recall accuracy. The Quarterly Journal of Experimental Psychology, 65, 179–194.
Karns, T. E., Irvin, S. J., Suranic, S. L., & Rivardo, M. G. (2009). Collaborative recall reduces the effect of a misleading post event narrative. North American Journal of Psychology, 11, 17–28.
Kensinger, E. A., Choi, H. Y., Murray, B. D., & Rajaram, S. (2016). How social interactions affect emotional memory accuracy: Evidence from collaborative retrieval and social contagion paradigms. Memory & Cognition, 44(5), 706–716.
Lorge, I., & Solomon, H. (1961). Group and individual behavior in free recall. In Criswell, J. H., Solomon, H., & Suppes, P. (Eds.), Mathematical methods in small group processes (pp. 221–231). Stanford, CA: Stanford University Press.
Lorge, J., & Solomon, H. (1955). Two models of group behavior in the solution of eureka-type problems. Psychometrika, 20, 139–148.
Marion, S. B., & Thorley, C. (2016). A meta-analytic review of collaborative inhibition and postcollaborative memory: Testing the predictions of the retrieval strategy disruption hypothesis. Psychological Bulletin, 142, 1141–1164.
Mastroberardino, S., Natali, V., & Candel, I. (2012). The effect of eye closure on children's eyewitness testimonies. Psychology, Crime & Law, 18(3), 245–257.
Memon, A., Zaragoza, M., Clifford, B. R., & Kidd, L. (2010). Inoculation or antidote? The effects of cognitive interview timing on false memory for forcibly fabricated events. Law and Human Behavior, 34(2), 105–117.
Natali, V., Marucci, F. S., & Mastroberardino, S. (2012). Long-term memory effects of eye closure on children eyewitness testimonies. Applied Cognitive Psychology, 26(5), 730–736.
Pezdek, K., Lam, S. T., & Sperry, K. (2009). Forced confabulation more strongly influences event memory if suggestions are other generated than self-generated. Legal and Criminological Psychology, 14, 241–252.
Pezdek, K., Sperry, K., & Owens, S. M. (2007). Interviewing witnesses: The effect of forced confabulation on event memory. Law and Human Behavior, 31, 463–478.
Poole, D. A., & White, L. T. (1993). Two years later: Effect of question repetition and retention interval on the eyewitness testimony of children and adults. Developmental Psychology, 29(5), 844.
Rajaram, S. (2011). Collaboration both hurts and helps memory a cognitive perspective. Current Directions in Psychological Science, 20, 76–81.
Rajaram, S., & Pereira-Pasarin, L. P. (2010). Collaborative memory: Cognitive research and theory. Perspectives on Psychological Science, 5(6), 649–663.
Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 803–814.
Ross, M., Spencer, S. J., Blatz, C. W., & Restorick, E. (2008). Collaboration reduces the frequency of false memories in older and younger adults. Psychology and Aging, 23, 85–92.
Ross, M., Spencer, S. J., Linardatos, L., Lam, K. C., & Perunovic, M. (2004). Going shopping and identifying landmarks: Does collaboration improve older people's memory? Applied Cognitive Psychology, 18(6), 683–696.
Rossi-Arnaud, C., Pieroni, L., Spataro, P., & Cestari, V. (2011). Effects of pair collaboration and word-frequency in recognition memory: A study with the remember-know procedure. Scandinavian Journal of Psychology, 52, 516–523.
Rossi-Arnaud, C., Spataro, P., Bhatia, D., & Cestari, V. (2019). Collaborative remembering reduces suggestibility: A study with the Gudjonsson Suggestibility Scale. Memory, 27, 603–611.
Rossi-Arnaud, C., Spataro, P., & Geraci, L. (2018). Effects of stereotype threat and prior task success on older adults' eyewitness memory. Journal of Applied Research in Memory and Cognition, 7(3), 422–431.
Stephenson, G. M., Brandstätter, H., & Wagner, W. (1983). An experimental study of social performance and delay on the testimonial validity of story recall. European Journal of Social Psychology, 13, 175–191.
Stolzenberg, S., & Pezdek, K. (2013). Interviewing child witnesses: The effect of forced confabulation on event memory. Journal of Experimental Child Psychology, 114(1), 77–88.
Takahashi, M. (2007). Does collaborative remembering reduce false memories? British Journal of Psychology, 98(1), 1–13.
Takahashi, M., & Saito, S. (2004). Does test delay eliminate collaborative inhibition? Memory, 12, 722–731.
Thorley, C., & Dewhurst, S. A. (2007). Collaborative false recall in the DRM procedure: Effects of group size and group pressure. European Journal of Cognitive Psychology, 19(6), 867–881.
Thorley, C., & Dewhurst, S. A. (2009). False and veridical collaborative recognition. Memory, 17(1), 17–25.
Vredeveldt, A., Groen, R. N., Ampt, J. E., & van Koppen, P. J. (2017). When discussion between eyewitnesses helps memory. Legal and Criminological Psychology, 22(2), 242–259.
Vredeveldt, A., Hildebrandt, A., & van Koppen, P. J. (2016). Acknowledge, repeat, rephrase, elaborate: Witnesses can help each other remember more. Memory, 24, 669–682.
Vredeveldt, A., van Deuren, S., & van Koppen, P. J. (2019). Remembering with a friend or a stranger: Comparing acquainted and unacquainted pairs in collaborative eyewitness interviews. Memory. https://doi.org/10.1080/09658211.2019.1662052.
Weldon, M. S., & Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, 1160–1175.
Wessel, I., Zandstra, A. E., Hengeveld, H. E., & Moulds, M. L. (2015). Collaborative recall of details of an emotional film. Memory, 23, 437–444.
Wright, D. B., & Klumpp, A. (2004). Collaborative inhibition is due to the product, not the process, of recalling in groups. Psychonomic Bulletin & Review, 11, 1080–1083.
Yaron-Antar, A., & Nachson, I. (2006). Collaborative remembering of emotional events: The case of Rabin's assassination. Memory, 14, 46–56.
Zaragoza, M. S., Payment, K. E., Ackil, J. K., Drivdahl, S. B., & Beck, M. (2001). Interviewing witnesses: Forced confabulation and confirmatory feedback increase false memories. Psychological Science, 12, 473–477.