Cognition 200 (2020) 104254

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Reviews

Social interactions can simultaneously enhance and distort memories: Evidence from a collaborative recognition task

Magdalena Abel*, Karl-Heinz T. Bäuml

Department of Experimental Psychology, Regensburg University, Germany

ARTICLE INFO

Keywords:
Social contagion
Collaboration
False memory
Retrieval practice
Testing effect
Collective memory

ABSTRACT

Social interactions can strengthen memories, but they can also contaminate them, for instance, when individuals integrate misinformation from social sources into their own memories. In two experiments, we examined such social contagion in a novel recognition-based collaboration task. Groups of three subjects studied lists of words in isolation, but unbeknownst to subjects, only half of these words were shared by the whole group. The remaining words were shared by two individuals within the group, or they were unshared, studied by single individuals only. Subjects completed a first recognition test collaboratively, and we examined how this group recognition task affected subsequent individual recognition, both for originally studied and unstudied items. Moreover, for some subjects an initial individual test was placed before collaboration, to examine if prior individual retrieval influences social contagion. Three main findings emerged: First, collaboration enhanced individual memory for initially studied information, but mostly for information shared within the group. Second, collaboration entailed robust social contagion effects, which were more pronounced when the misinformation was shared by both other group members. Third, an individual memory test before collaboration reduced social contagion. We discuss the powerful effects of memory retrieval and how social interactions can simultaneously enhance and distort memories.

1. Introduction

People are social, and they frequently engage in conversations with each other (Hirst & Echterhoff, 2012). Consider how often, during a regular day, you interact with other people, and how frequently this requires recalling memories, either in the form of details about previous experiences or events, or in the form of simple facts. Such conversations can have profound influences on memory. Discussed contents may be strengthened and enhanced in memory (e.g., Blumen & Stern, 2011; Blumen, Young, & Rajaram, 2014; Weldon & Bellinger, 1997), whereas unmentioned details may be forgotten (e.g., Cuc, Koppel, & Hirst, 2007). In addition, human memory is also susceptible to misinformation (e.g., Loftus, Miller, & Burns, 1978). Such misinformation can be consumed in many ways, but the term social contagion refers to instances in which the misinformation's source is another human being (e.g., Roediger, Meade, & Bergman, 2001). Even though implications have mostly been discussed for eyewitness memory, social contagion and information transmission across individuals are of potential relevance for many situations in daily life. For instance, they may also operate when students discuss the contents of a recent lecture, when friends chat about the events surrounding last weekend's party, or when coworkers jointly remember a conference or another work-related event.

Researchers have used several approaches to examine social contagion, differing in how misinformation is introduced after original information was encountered (for an overview, see Maswood & Rajaram, 2018). One approach is to pair subjects with a confederate, who “remembers” specific misinformation, thereby contaminating the subjects' later memory of the original study phase (e.g., Numbers, Meade, & Perga, 2014; Roediger et al., 2001). Another frequently used approach is to ask subjects to review a fictitious recall protocol after initial study. This recall protocol supposedly comes from another subject and also contains certain misinformation (e.g., Gabbert, Memon, Allan, & Wright, 2004; Meade & Roediger, 2002). Finally, a third approach uses real social interactions. To introduce misinformation, the initially studied information is manipulated and only partly the same for the two members of a dyad. In particular, each individual encodes single details that are not shared by the other individual. During joint recall of the encoded information, the unshared details can come to light and spread from one group member to the other (e.g., Gabbert, Memon, & Allen, 2003; Garry, French, Kinzett, & Mori, 2008; Mori, 2003; Wright, Self, & Justice, 2000). Notably, much of the work relying on this third approach followed the eyewitness tradition, and applied complex and real-life events in the form of videos as study materials. Doing so, social contagion for single, yet critical details demonstrates the

* Corresponding author at: Department of Experimental Psychology, Regensburg University, 93040 Regensburg, Germany.
E-mail address: magdalena.abel@ur.de (M. Abel).

https://doi.org/10.1016/j.cognition.2020.104254
Received 24 June 2019; Received in revised form 18 February 2020; Accepted 2 March 2020

relevance of mnemonic distortions for legal contexts, but, at the same time, it is hard to examine basic properties of social contagion in the lab on the basis of these tasks.

The aim of the present study was to introduce an experimental paradigm that overcomes this difficulty and provides further benefits as well. Our primary goal was to create a task that allows capturing different aspects of social remembering at once - negative effects, such as social contagion errors, but also positive effects, such as collaborative enhancement or reexposure effects. A secondary goal was that the task should leave room for investigating advanced questions on the mnemonic consequences of social interactions, as, for instance, the question if prior individual retrieval practice plays a modulating role. If successful, the task would go beyond prior approaches by allowing for systematic investigation of real social interactions and their multiple influences on memory. All this holds while, of course, the task was heavily inspired by prior, so far mostly unconnected work in the literature on social memory, and we describe these influences plus some further considerations next.

First, our task follows prior studies that manipulated information distribution during encoding (e.g., Gabbert et al., 2003; Garry et al., 2008). As in these studies, subjects completing our task studied partly overlapping information and partly unshared information, which they could later on carry into a group collaboration phase, thus potentially distorting each others' memories. We chose word lists as study materials, though it should be possible to adapt the task to use other types of materials. Second, our task used the same basic group set-up as prior work on collaborative retrieval (e.g., Basden, Basden, Bryner, & Thomas, 1997; Weldon & Bellinger, 1997). Subjects were tested in groups of three, and engaged in individual or collaborative retrieval after encoding. At the end of the experiment, all subjects completed a final individual test. On this test, social contagion errors should arise for subjects who previously engaged in collaborative recall with the rest of their group. In addition, collaboration with others can also enhance memory for information studied by the whole group, namely by reexposing individuals to information that they would otherwise have forgotten (e.g., Abel & Bäuml, 2017; Bärthel, Wessel, Huntjens, & Verwoerd, 2017; Basden, Basden, & Henry, 2000; Blumen et al., 2014; Blumen & Stern, 2011; Weldon & Bellinger, 1997). Because reexposure with initially studied information should be more likely for information shared across group members, such benefits might be expected to be less pronounced for unshared items, encoded by single group members only. In any case, combining the experimental approach of collaborative retrieval with a manipulation of information distribution during encoding enables simultaneous examination of social contagion and facilitation (for similar approaches, see Choi, Kensinger, & Rajaram, 2017; Muller & Hirst, 2014). Third, grounded in practical considerations, we based the whole task on recognition tests, because using such tests ensures that all initially encoded information (shared and unshared within groups) is encountered by all participants in all conditions.

The present study is novel in that it used a recognition-based collaboration task to examine facilitation and social contagion arising in real social interactions. This holds while a recent study by Choi et al. (2017) also manipulated information distribution, albeit in triads collaborating in a recall task. Ultimately, this prior study focused on different issues than the present study. In particular, it examined how repeated collaboration in the same vs. in reconfigured triads affected individual recall of negative, neutral, and positive information as well as the build-up of collective memories across individuals. Additionally, Choi et al. (2017) also manipulated encoding such that some information was fully shared, some partially shared, and some unshared by single individuals within the triads. Because the specific implementation of this manipulation differed from the present study and implementation affects the analyses that can be carried out, the present study is able to address other research questions than the prior work, thus leading to novel results (for a more detailed discussion on how the results of the present study complement those reported by Choi et al., 2017, see General discussion).

Apart from capturing several aspects of social interactions, our recognition-based task may additionally allow examining advanced questions on social memory, like the question how individual retrieval prior to collaborative retrieval affects social contagion effects. Some previous work indicates that initial individual retrieval can protect from social contagion (e.g., Huff, Davis, & Meade, 2013; Huff, Weinsheimer, & Bodner, 2016). In these studies, subjects were confronted with fictitious recall protocols of other participants as sources of misinformation. The likelihood of subjects integrating misinformation into their own memories was reduced if an individual retrieval test had been completed before subjects reviewed the recall protocols. Overall, this seems well in line with the testing effect literature, for instance showing that initial retrieval can protect from retroactive interference (e.g., Abel & Roediger, 2017; Bäuml, Holterman, & Abel, 2014; Halamish & Bjork, 2011; Potts & Shanks, 2012).

Yet, contrasting with these studies, there is other evidence showing that retrieval practice can increase susceptibility to misinformation (e.g., Chan & LaPaglia, 2011, 2013; Chan, Thomas, & Bulevich, 2009; Gordon & Thomas, 2014, 2017). In corresponding studies, misinformation was presented via summarizing narratives, either in written or audio form (with no emphasis of a potential social source). Individual retrieval practice prior to encountering these narratives increased the likelihood of subjects integrating the misinformation into their own memories, which is consistent with related work showing that retrieval practice can potentiate new learning (e.g., Arnold & McDermott, 2013; Szpunar, McDermott, & Roediger, 2008). Overall, there is no agreement in the literature on how individual retrieval affects the adoption of misinformation. In the present study, we wanted to apply our collaborative recognition task to address the question, and included a third condition in our study design in which subjects engaged in collaborative retrieval as well, but additionally completed a prior individual test.

Experiment 1 was the first study to apply our newly developed task. Subjects were tested in groups of three. We manipulated i) information distribution across groups during encoding, with only some information being studied by all group members, and ii) collaborative vs. individual retrieval after encoding, with only some participants engaging in joint remembering with the other group members. All participants completed a final individual test, and our main objective was to investigate if the task can be used to experimentally assess both social contagion and collaborative facilitation arising on this final test. Finally, in order to explore if the task can be used to examine influence factors on social contagion, we additionally explored the role of prior individual retrieval for social memory effects. In an additional condition, subjects took an individual test before engaging in collaborative group recall. Experiment 2 was run to replicate the basic findings of Experiment 1 while also excluding a potential confound. In particular, Experiment 1 showed that the overall duration of collaborative vs. individual recognition tests can differ greatly, and we excluded this issue as a potential confound by more adequately matching test durations in Experiment 2.

2. Experiment 1

2.1. Method

2.1.1. Participants

Sample sizes were determined on the basis of a priori power analyses for within-between interactions in repeated-measures ANOVAs. Power was set to 0.80, alpha to 0.05, and correlations among repeated measures to 0.50. Assuming three between-subject groups, a total sample of 87 subjects (for two repeated measures) or 72 subjects (for three repeated measures) would be necessary in order to detect a small- to medium-sized interaction effect of $f = 0.17$. Thus, we recruited 90 subjects for each experiment. Participants in all experiments were undergraduates at Regensburg University, who received partial course credit for completing the experiment. Subjects participated in triads,

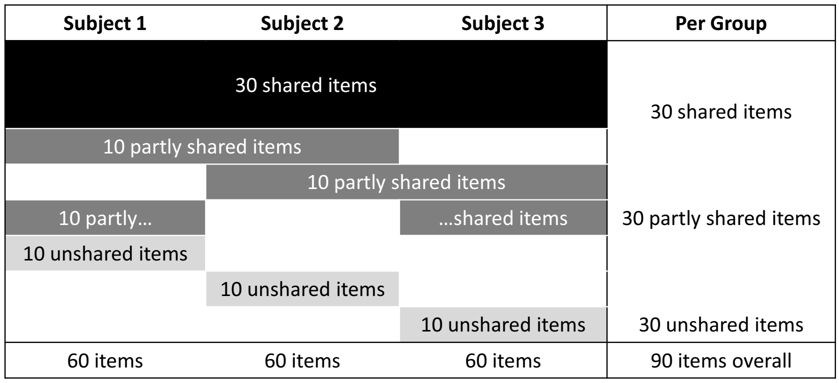

Fig. 1. Distribution of shared, partly shared, and unshared items across the three subjects in each group. Shared items were initially studied by all three group members, partly shared items were studied by two of the three group members, and unshared items were studied by single group members only.

with 10 triads ($n = 30$) being tested in each of three conditions. All subjects knew that they would complete the experiment together with two others when signing up. Mean age in Experiment 1 was 20.8 years ($SD=1.9$); 32 subjects were male, 58 female. Participants in 2 triads reported not knowing each other at all before the experiment; in 12 triads, two of the three participants knew each other beforehand, and in 16 triads all three participants were acquainted.

2.1.2. *Material*

180 concrete nouns were selected from the CELEX-database (Duyck, Desmet, Verbeke, & Brysbaert, 2004) and divided into two sets of 90 words each. Across subjects, each set was equally often used as study and distractor material. The two item sets did not differ with regard to mean number of letters ($M=5.2$, $SD=1.1$), mean number of syllables ($M=1.8$, $SD=0.6$), or mean word frequency (per million; $M=20.7$, $SD=63.8$), $ts(178) < 1.00$, $ps \geq .447$, $ds \leq 0.11$.

To be used as study material, each set was randomly grouped into 30 items that would be studied by all three subjects, 30 items that would be studied by two subjects, and 30 items that would be studied by single individuals within the group only. Similarly, when used as distractor material, the same three groups of 30 items were applied as lures on the three different memory tests (the critical collaborative vs. individual test; the final test for assessing effects of prior collaboration on individual memory; and the initial individual test before the collaboration phase, applied in only one of the three conditions).

2.1.3. *Design*

The experiment had a 3 × 8 mixed-factorial design. The first factor retrieval task (individual, collaborative, collaborative plus prior test) was manipulated between-subjects. In the individual retrieval condition, subjects completed two individual recognition tests (II) in the course of the experiment, and never engaged in social interactions during these tests. In the collaborative retrieval condition, subjects engaged in the same two recognition tests, but completed the first one in a collaborative manner. Subsequently, they completed the same final individual recognition test as subjects from the individual retrieval condition (CI). In the collaborative plus prior test condition, subjects completed the same collaborative and individual recognition tests as subjects in the regular collaborative retrieval condition; prior to these tests, however, subjects in this condition were asked to complete an initital individual test (ICI), to see if it would protect them from or enhance their susceptibility to potential social contagion effects.

The second factor item type was manipulated within-subjects by differentiating between items that were initially studied (three different item types), initially unstudied (two different item types), or used as lures on the conducted recognition tests (three different item types). For each subject, items that were initially studied could be sorted into items shared with both, one or none of the other subjects in a group (which we refer to as shared, partly shared, and unshared items). These items were used to examine collaborative facilitation effects. In addition, for each subject, there were items that they themselves had not studied initially, but that had been studied by both or one of the other group members (which we refer to as strong and weak contagion items). These items were used to examine social contagion. Finally, because all memory tests in Experiment 1 were recognition tests, we also differentiated between the three sets of lures introduced on different tests throughout the experiment to examine regular false alarms: Lures first used during the initial individual test in the ICI-condition (lures 1), the collaborative vs. individual recognition test (lures 2), and the final individual recognition test (lures 3). Our main analyses focused on potential differences on the final individual recognition test as a function of retrieval task (II, CI, ICI). For complenetess, hit and false alarm rates for the individual vs. collaborative tests are reported in appendices (see Appendix A for data from Experiment 1 and Appendix B for data from Experiment 2).

2.1.4. *Procedure*

2.1.4.1. Study phase. Each subject was seated in front of a different computer. After providing informed consent, subjects were instructed that they would be seeing a relatively high number of items presented on the screen, and that they should try to memorize them as best as possible. Subjects received no further information about the study materials, i.e., they were neither warned that they would study partly different items, nor were they – falsely – told that they would all study the same information. During study, items were presented one at a time, in random sequences, and for 2 s each, centrally on the computer screens. Each subject studied 60 items: 30 items shared with both other group members, 20 items shared with one other group member, and 10 unshared items. Thus, 90 items were studied across each triad (see Fig. 1 for an overview).

2.1.4.2. Distractor phase and first individual test. After study, all subjects were distracted for roughly 2 min and asked to provide some demographic information about themselves. Subjects in the individual and collaborative retrieval conditions (II, CI) immediately went on to complete further distractor tasks, namely the connect-the-numbers test (Oswald & Roth, 1987), standard progressive matrices (Raven, 2000), and decision problems (Tversky & Kahneman, 1974). In sum, the distractor phase comprised 20 min. Several different distractor tasks were used to keep participants engaged, but the same tasks were applied in all conditions, with all participants completing the same identical tasks (e.g., Abel & Bäuml, 2019).

Subjects in the collaborative plus prior test condition (ICI) completed a first individual test right after answering the demographic questions. In particular, still seated at their individual computers, each subject completed a recognition test that comprised only the 60 items studied earlier and 30 new items not previously studied (lures 1), intermixed randomly. Subjects were asked to decide for each single item if it was an old item from the study phase or a new item. The test was completely subject paced, and subjects entered their responses via their computer keyboards. On average, completing this first test took 109.5 s (SD = 32.9). Subsequently, subjects were asked to work on the same distractor tasks as subjects in the other two retrieval conditions. The interval spent with the individual test and distractor tasks filled 20 min in this condition, too, so that delay between initial encoding and the critical collaboration phase was the same across all three conditions.

2.1.4.3. Collaborative vs. individual recognition test. Next, subjects in all conditions were asked to complete a recognition test. Because subjects had only partly studied the same information, we expected that collaboration would result in social contagion. In the two collaborative conditions (CI, ICI), triads completed this test together and were asked to decide as a group whether items were old items from the study phase or new items not studied earlier. For this purpose, subjects were seated in front of the same computer. Before the test began, subjects were simply asked to complete the test as a group. They received no further instructions on how to determine group responses. In particular, they were never instructed to try to convince each other or to reach decisions that should be adopted by each individual. All 90 items from the study phase (see Fig. 1) were presented intermixed with 60 lures: the 30 lures already used for the individual recognition test in the ICI condition (lures 1) plus 30 new lures (lures 2). In order to avoid one subject taking over responses, the experimenter entered responses for all collaborative groups via the computer keyboard, thus also ensuring group discussions. Moreover, collaborative tests were audio recorded. Subjects in the individual retrieval conditions completed the same test, but did so at their individual computers, without any social interaction. Afterwards, subjects in all conditions were asked to work on another distractor task for roughly 5 min (the d2 test of attention; Brickenkamp & Zillmer, 1998).

2.1.4.4. Final individual recognition test. The final phase of the experiment comprised another recognition test which was completed individually by subjects in all conditions. On this test, 100 items had to be judged with regard to whether they were old items from the initial study phase or new items, not encountered during the study phase. Not all items were included on this final test because i) we wanted to create similar numbers of observations for to-be-contrasted item types (e.g., for 10 fully shared, 10 partly shared, and 10 unshared items), and ii) we worried about the experiment's duration and participants' concentration across several memory tests on the same materials. Therefore, the final test was limited to a subset of items (see Table 1 for a breakdown of item types and their use across tests).

For each subject, the final test included 30 items taken from the original study phase: The 10 unshared items, as well as 10 of the partly shared items (drawn randomly for each subject out of the overall 20 items per subject), and 10 of the shared items (again, drawn randomly out of the overall 30 items). Critically, the test also included social contagion items, namely the 10 strong contagion items (studied by both other subjects) and 10 of the weak contagion items (drawn randomly out of the overall 20 items studied by one other individual). These 50 items were presented randomly intermixed with 50 lure items: 16 lures were randomly drawn from the set previously used on the initial individual recognition test (lures 1), 17 lures were randomly drawn from the set previously used on the collaborative vs. individual test (lures 2), and 17 lures were randomly drawn from a fresh set of 30 items (lures 3). The test was again completely subject paced, and subjects were asked to enter their responses via their computer keyboards. When all subjects had completed the final test, they were debriefed and thanked for their participation (for a schematic overview of the whole procedure, see Fig. 2).

2.2. Results

2.2.1. Facilitation

We first analyzed mean hit rates for items that were initially studied (see Fig. 3a). A 3 × 3 ANOVA with the factors of ITEM TYPE (shared, partly shared, unshared) and retrieval condition (II, CI, ICI) showed no significant main effect of retrieval condition, $F(2,87) = 1.69$, $MSE = 0.08$, $p = .191$, $\eta_p^2 = 0.04$, but a significant main effect of item type, $F(2,174) = 8.89$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.09$, which was accompanied by a significant interaction of the two factors, $F(4,174) = 2.65$, $MSE = 0.02$, $p = .035$, $\eta_p^2 = 0.06$. To better understand this pattern, we conducted further analyses.

An ANOVA contrasting the individual and collaborative retrieval conditions (II vs. CI) revealed the same pattern, with no significant main effect of retrieval condition, $F(1,58) = 2.98$, $MSE = 0.09$, $p = .089$, $\eta_p^2 = 0.05$, but a significant main effect of item type, $F(2,116) = 3.67$, $MSE = 0.02$, $p = .029$, $\eta_p^2 = 0.06$, and a significant interaction of the two factors, $F(2,116) = 4.48$, $MSE = 0.02$, $p = .013$, $\eta_p^2 = 0.07$. Hit rates were enhanced in the collaborative relative to the individual retrieval condition, but only for shared items, studied by all three subjects in a group (83.0% vs. 67.3%), $t(58) = 2.83$, $p = .007$, $d = 0.73$. There was no corresponding collaborative facilitation for partly shared items (78.0% vs. 71.0%), $t(58) = 1.35$, $p = .182$, $d = 0.35$, or unshared items (69.0% vs. 68.7%), $t(58) < 1.00$, $p = .950$, $d = 0.02$.

A further ANOVA contrasting the two collaborative conditions with and without a prior individual test (CI vs. ICI) only showed a significant main effect of item type, $F(2,116) = 12.86$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.18$, but no other significant effects, $Fs < 1.00$, $MSEs \leq 0.05$, $ps \geq .366$, $\eta_p^2 \leq 0.01$. Unshared items were generally remembered worse than partly and completely shared items (67.7% vs. 77.2% vs. 80.5%), $ts(59) \geq 3.46$, $ps \leq .001$, $ds \geq 0.45$; memory for these items did not differ, $t(59) = 1.52$, $p = .135$, $d = 0.20$.

2.2.2. Social contagion

Next, we examined false alarms for items that were initially unstudied by each subject, but studied by one if not both other group member(s) (see Fig. 3b).[1] A 2 × 3 ANOVA with the factors of item type (weak contagion items, strong contagion items) and retrieval condition (II, CI, ICI) showed a significant main effect of retrieval condition, $F(2,87) = 32.83$, $MSE = 0.05$, $p < .001$, $\eta_p^2 = 0.43$, a significant main effect of item type, $F(1,87) = 21.30$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.20$, and a significant interaction of the two factors, $F(2,87) = 4.55$, $MSE = 0.02$, $p = .013$, $\eta_p^2 = 0.10$. To clarify this pattern, we again ran further analyses.

An ANOVA contrasting the individual and collaborative retrieval conditions (II vs. CI) showed the same pattern, with two significant main effects and a significant interaction effect, $Fs \geq 4.58$, $MSEs \leq 0.05$, $ps \leq .037$, $\eta_p^2 \geq 0.07$. Follow-up tests on false alarms in the individual retrieval condition alone showed no differences between strong and weak contagion items (25.3% vs. 24.3%), $t(29) < 1.00$, $p = .779$, $d = 0.01$, reflecting that subjects in this condition did not engage in joint recall with the other group members who had studied these items. In the collaborative retrieval condition, strong contagion items (studied by both other group members) evoked higher false alarm rates than weak contagion items (studied by one other group member; 64.0% vs. 51.3%), $t(29) = 3.05$, $p = .005$, $d = 0.56$. False alarm rates for both item types were enhanced after collaborative vs. individual

[1] Analyses on social contagion errors included all corresponding items and not only items previously endorsed on the collaborative recognition test. Previous group endorsement may not be necessary for social contagion to arise, because group discussions and disagreements can still affect individual memory later on, even in cases in which an item was not classified as old during collaboration. In addition, false alarms in the individual retrieval condition offer an adequate baseline for evaluation, because increases in false alarms in collaborative relative to individual retrieval conditions are likely due to collaboration and social input.

Table 1
Breakdown of item types and their use across tests in Experiment 1.

Item type	Number of items per subject	Initial individual test (ICI condition only)	Collaborative vs. individual test	Final individual test	
Initially studied items					
Fully shared items	30 items	X	X	X	Only 10 items tested
Partly shared items	20 items	X	X	X	Only 10 items tested
Unshared items	10 items	X	X	X	All 10 items tested
Social contagion items					
Strong contagion items	10 items	X	X	X	All 10 items tested
Weak contagion items	20 items	X	X	X	Only 10 items tested
Lures					
Lures 1	30 items	X	X	X	Only 16 items tested
Lures 2	30 items		X	X	Only 17 items tested
Lures 3	30 items			X	Only 17 items tested

Note. An X indicates that an item type was included on the respective memory test. Usually, all items belonging to a specific item type were tested; the only exception is the final individual test, with the last column additionally showing how many items belonging to each item type were tested. See the main text for further details.

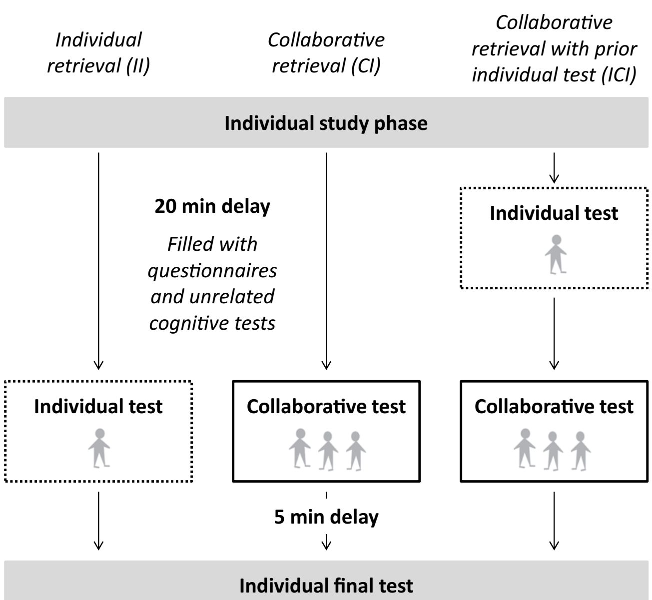

Fig. 2. Schematic depiction of experimental procedures in the three retrieval conditions.

retrieval, but social contagion was more pronounced for strong contagion items (64.0% vs. 25.3%), $t(58) = 8.35$, $p < .001$, $d = 2.16$, than for weak contagion items (51.3% vs. 24.3%), $t(58) = 5.56$, $p < .001$, $d = 1.44$.

To examine if individual retrieval prior to collaboration influenced social contagion, we conducted another ANOVA contrasting the two collaboration conditions. The ANOVA revealed a significant main effect of item type, $F(1,58) = 28.27$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.33$, and a significant main effect of retrieval condition, $F(1,58) = 18.01$, $MSE = 0.04$, $p < .001$, $\eta_p^2 = 0.24$, but no significant interaction between the two factors, $F(1,58) < 1.00$, $MSE = 0.02$, $p = .504$, $\eta_p^2 = 0.01$. False alarm rates were also higher for strong compared to weak items after collaborative retrieval with a prior individual test (50.0% vs. 33.7%), $t(29) = 4.62$, $p < .001$, $d = 0.84$. Relative to the condition without prior individual retrieval, however, false alarm rates were significantly reduced, and this was the case for both weak contagion items (33.7% vs. 51.3%), $t(58) = 3.88$, $p < .001$, $d = 1.00$, and strong contagion items (50.0% vs. 64.0%), $t(58) = 2.99$, $p = .004$, $d = 0.77$.

2.2.3. False alarms for lures

Finally, we also examined false alarms for lures introduced at different times throughout the experiment (see Fig. 3c). A 3 × 3 ANOVA with the factors of ITEM TYPE (lures 1, lures 2, lures 3) and RETRIEVAL CONDITION (II, CI, ICI) showed a significant main effect of retrieval condition, $F(2,87) = 7.08$, $MSE = 0.05$, $p = .001$, $\eta_p^2 = 0.14$, a significant main effect of item type, $F(2,174) = 88.09$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.50$, and a significant interaction of the two factors, $F(4,174) = 3.60$, $MSE = 0.02$, $p = .008$, $\eta_p^2 = 0.08$.

An ANOVA contrasting the individual and collaborative retrieval conditions (II vs. CI) showed two significant main effects for item type and retrieval condition, $Fs \geq 9.97$, $MSEs \leq 0.06$, $ps \leq .003$, $\eta_p^2 \geq 0.15$, but no significant interaction between the two factors, $F(2,116) < 1.00$, $MSE = 0.02$, $p = .619$, $\eta_p^2 = 0.01$. False alarm rates did not differ for lures 1 and lures 2 (35.4% vs. 34.3%), $t(59) < 1.00$, $p = .606$, $d = 0.01$. This makes sense, because in these two retrieval conditions, the two lure sets were both first introduced on the same recognition test, namely the collaborative vs. individual test. In comparison, false alarm rates for lures 3, introduced only on the final recognition test, were reduced (16.7%), $ts(59) \geq 8.17$, $ps < .001$, $ds \geq 1.05$. Further follow-up tests contrasting false alarms across the two retrieval conditions showed that prior collaborative relative to individual retrieval enhanced false alarm rates for all three sets of lures (lures 1: 42.0% vs. 28.9%, $t(58) = 2.61$, $p = .012$, $d = 0.67$; lures 2: 40.2% vs. 28.4%, $t(58) = 2.59$, $p = .012$, $d = 0.67$; lures 3: 21.1% vs. 12.2%, $t(58) = 2.61$, $p = .012$, $d = 0.67$).

We next contrasted the two collaboration conditions to examine if individual retrieval prior to collaboration influenced false alarms for lures, too. The ANOVA revealed a significant main effect of item type, $F(2,116) = 72.94$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.56$, a significant main effect of retrieval condition, $F(1,58) = 14.70$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = 0.20$, and a significant interaction between the two factors, $F(2,118) = 3.99$, $MSE = 0.02$, $p = .025$, $\eta_p^2 = 0.06$. Prior individual retrieval reduced false alarms for lures 2 (26.3% vs. 40.2%), $t(58) = 3.84$, $p < .001$, $d = 0.99$, and lures 3 (8.3% vs. 21.1%), $t(58) = 4.53$, $p < .001$, $d = 1.17$, but not for lures 1 (39.5% vs. 42.0%), $t(58) < 1.00$, $p = .552$, $d = 0.15$. Again, this latter finding makes sense, given that the first set of lures was introduced earlier in the ICI condition, whereas lures 2 and 3 were introduced at the same time in the experiment across the two conditions.[2]

[2] Gender was not controlled in the present study, but upon reviewer's request we conducted additional analyses for participants in the collaborative conditions to see if gender distribution within groups affected the general patterns observed for all item types. Separate analyses were again carried out for facilitation, social contagion, and lures. Comparing purely female/male groups (7 triads, n = 21) and mixed groups (13 triads, n = 39), no significant main effects or interaction effects involving the factor gender distribution were observed (all $ps \geq .117$, all $\eta_p^2 s \leq 0.04$).

Similarly, we were asked to conduct additional analyses to examine if facilitation and social contagion were affected by whether subjects within triads knew each other or not. Comparing completely unacquainted triads (2 triads, n = 6) to triads in which at least two (12 triads, n = 36) or all three participants knew each other (16 triads, n = 48), no significant main or interaction effects involving the factor acquaintance were found (all $ps \geq .106$, all $\eta_p^2 s \leq 0.05$).

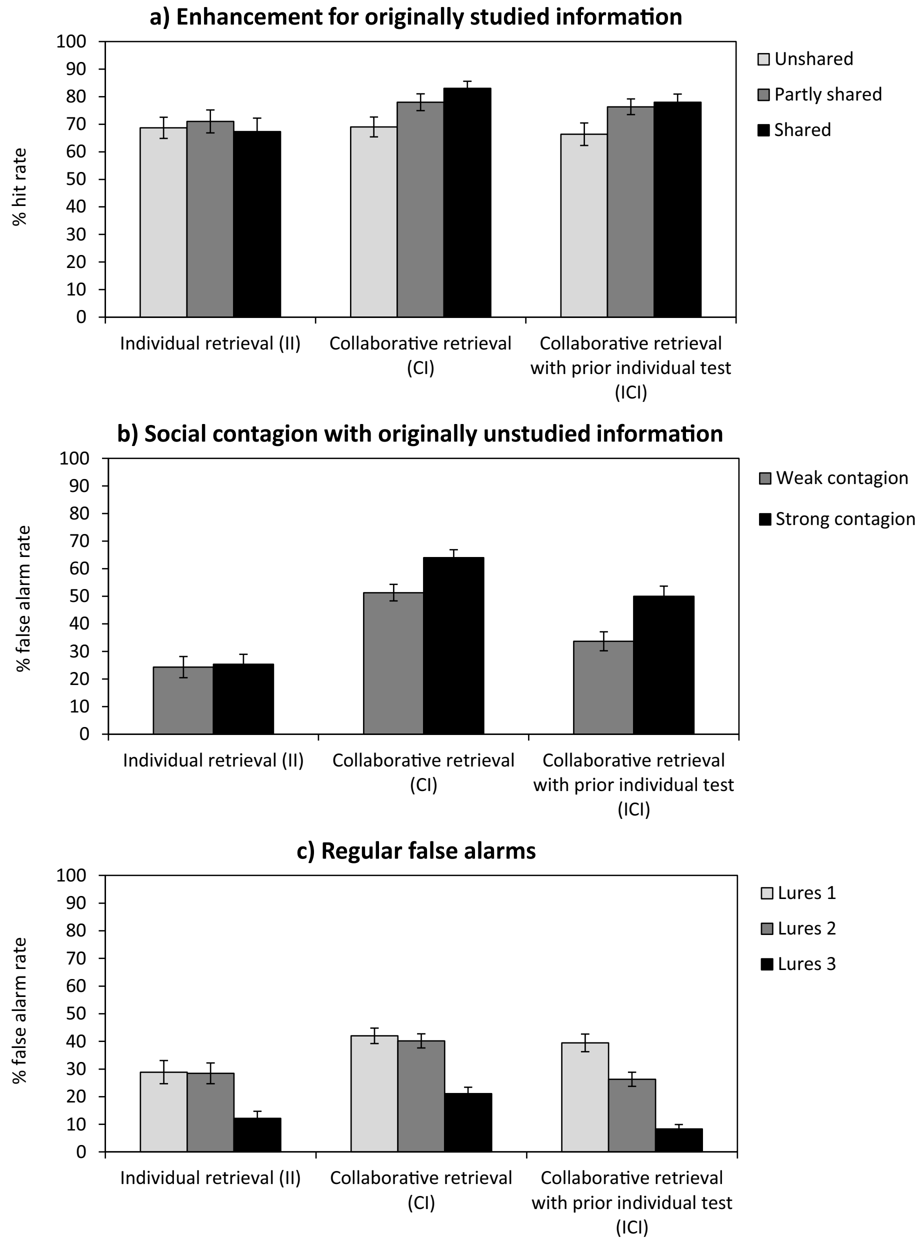

Fig. 3. Results of Experiment 1. Panel a) shows hit rates for originally studied information; panel b) shows social contagion with originally unstudied information; panel c) shows regular false alarm rates.

2.2.4. Additional analysis: performance on the initial individual test (ICI condition only)

Mean hit rate for studied items was at 73.9% ($SD = 12.8$). Hit rates did not differ for shared, partly shared, and unshared items (72.9% vs. 76.5% vs. 72.3%) on this test, $F(2,58) = 1.80$, $MSE = 0.10$, $p = .181$, $\eta_p^2 = 0.06$. The false alarm rate for lures was at 19.6% ($SD = 18.5$).

2.2.5. Additional analysis: test durations of individual vs. collaborative tests

Before running the experiment, it was unclear how long the critical individual vs. collaborative retrieval phase used to trigger social memory effects would take. Therefore, we analyzed duration for the 10 triads in each of the two collaborative conditions, and randomly selected 10 out of the 30 inviduals in the individual condition and included their data in the analysis. A one-way ANOVA indicated a significant difference in mean duration between conditions, $F(2,27) = 68.32$, $MSE = 21.70$, $p < .001$, $\eta_p^2 = 0.84$. There was no significant difference between the two collaborative retrieval conditions (CI: $M = 26.0$ min, $SD = 5.4$; ICI: $M = 21.5$ min, $SD = 5.9$), $t(18) = 1.76$, $p = .096$, $d = 0.79$. Mean test duration in both collaborative conditions was however much longer compared to the individual retrieval condition (II: $M = 3.0$ min, $SD = 0.7$), $ts(18) \geq 9.82$, $ps < .001$, $ds \geq 4.39$. The same test took collaborating groups about 7–8 times longer than individual subjects.

2.3. Discussion

The primary aim of the present study was to introduce a task that allows capturing several different aspects of social remembering at once. Consistently, Experiment 1 shows that collaboration with others can simultaneously enhance and distort memories. On the one hand, subjects who had previously collaborated with others showed enhanced memory for initially studied information. This facilitation was most pronounced for information that was initially studied by all three group members. On the other hand, collaborating with others also triggered social contagion and the integration of misinformation into one's own memory of the study phase. Subjects in the collaborative retrieval condition were generally susceptible to misinformation by others, but showed even higher integration of misinformation when it was represented by both other group members (rather than just by one). Interestingly, collaboration also enhanced regular false alarm rates for neutral lures, indicating that information transmission during joint remembering in groups does not only have to comprise accurate information encoded by some or all group members, but can also involve plainly incorrect information misremembered by single individuals. Clearly, the applied task can be used to examine multiple (positive and negative) aspects of social memory.

The second aim was to design a task that allows examining advanced questions on mnemonic consequences of social interactions, like the potentially modulating role of prior individual retrieval practice. In line with previous findings on social contagion from studies that applied a confederate to introduce misinformation (e.g., Huff et al., 2013, 2016), subjects participating in our recognition-based collaboration task also showed reduced susceptibility to social contagion if they had beforehand completed an individual recognition test on the initially studied information. A novel finding was that individual retrieval reduced social contagion to about the same degree irrespective of whether the misinformation was represented by one or two other group members. The pattern was very similar for lures, with prior individual retrieval reducing false alarm rates for lures introduced during collaboration and during the final individual test. In contrast, prior individual retrieval did not meaningfully alter the collaborative facilitation effects observed for initially studied (shared) information (for related work, see Congleton & Rajaram, 2011).

One caveat raised by additional analyses concerns the overall duration of the critical individual vs. collaborative recognition test, used to elicit social memory effects. Our aim was to create a task that allows for unrestricted social interaction, but paired with a recognition test, this also means that the duration of the collaboration phase cannot be strictly controlled. Experiment 1 shows that such collaborative tests can in fact take 7–8 times longer than the same test taken by individuals. This difference in task durations creates a confound, raising the question if mnemonic differences between conditions will persist when individual and collaborative tests are more adequately matched in overall duration.

We pursued two goals in Experiment 2. Our first goal was to check if the results observed for collaborative vs. individual retrieval conditions (II vs. CI) in Experiment 1 could be replicated when duration of the collaborative vs. individual recognition task is better controlled, so that participants are exposed to stimulus materials for roughly the same time. Instead of restricting social interactions during collaboration, we addressed the issue by changing the individual recognition test from subject-paced to experimenter-timed, thereby increasing its overall duration. Our second goal was to examine if the finding of an initial individual recognition test protecting against social contagion in Experiment 1 hinged on the format of the test, or if it would generalize to a free-recall test format. A positive side effect of using free recall as initial test is that subjects in the corresponding condition are no longer exposed to an early set of lures. Yet, because free recall typically results in lower memory scores and may reactivate fewer study contents compared to recognition, free recall might be expected not to provide as much protection from contamination with misinformation.

3. Experiment 2

3.1. Method

3.1.1. Participants

The sample again comprised 90 subjects, tested in 30 triads. Mean age was 21.0 years (*SD* = 2.3); 13 subjects were male, 77 female. Participants in 7 triads reported not knowing each other at all before the experiment; in 9 triads, two of the three participants knew each other beforehand, and in 14 triads all three participants were acquainted.

3.1.2. Material

The same two sets of 90 items were applied as in Experiment 1 and their use as study and distractor materials was again counterbalanced across participants. When used as distractor material, 60 of the 90 items now served as lures 1, to be introduced on the critical collaborative vs. individual recognition test (which corresponds to lures 1 and lures 2 in Experiment 1). The remaining 30 items were used as fresh lures on the final individual recognition test and are now referred to as lures 2 (but correspond to lures 3 in Experiment 1).

3.1.3. Design

The experiment had a 3 × 7 mixed-factorial design. The first factor retrieval task (II, CI, ICI) was identical to Experiment 1 and again manipulated between-subjects. The second factor item type was very similar to Experiment 1 and manipulated within-subjects. Here, we again differentiated between items that were initially studied (fully shared, partly shared, or unshared within the group) and contagion items that were initially unstudied (but studied by both or one of the other group members). In contrast to Experiment 1, there were only two sets of lures in Experiment 2. One set of lures was introduced during the collaboration phase (lures 1), another one on the final individual test (lures 2).

3.1.4. Procedure

The general procedure was largely identical to the procedure applied in Experiment 1, with three differences. First, the individual test before the collaboration phase in the ICI condition was now realized as a free recall task. Subjects were handed a piece of paper and asked to write down all words from the study phase that they could remember. They had 3 min to complete this test.

Second, in order to roughly equate time spent on the individual vs. collaborative test across II and CI conditions, the first individual test in the II condition was no longer subject-paced. Test items were presented for 8 s each, and subjects were asked to provide their responses within this 8-second window. In order to avoid missed trials, a countdown was presented on the screen for the last 4 s of item presentation. All other details for this test were identical to Experiment 1.

Third, the make-up of the final individual recognition test changed slightly. As in Experiment 1, the test included 30 initially studied items (fully shared, partly shared, or unshared) as well as 20 initially unstudied items (weak vs. strong contagion items). These 50 items were again presented randomly intermixed with 50 lures, but composition of lures differed. In Experiment 2, 25 lures were randomly drawn from the set used during the collaboration phase (lures 1), and 25 lures were randomly drawn from a fresh set of lures (lures 2). The test was subject paced, and subjects entered responses via their computer keyboards.

3.2. Results

3.2.1. Facilitation

Fig. 4a shows mean hit rates for initially studied items. A 3 × 3 ANOVA with the factors of ITEM TYPE (shared, partly shared, unshared) and retrieval condition (II, CI, ICI) showed a significant main effect of item type, $F(2,174) = 6.18$, $MSE = 0.02$, $p = .003$, $\eta_p^2 = 0.07$, as well as a significant main effect of retrieval condition, $F(2,87) = 3.40$, $MSE = 0.07$, $p = .038$, $\eta_p^2 = 0.07$, but no significant interaction of the two factors, $F(4,174) = 1.06$, $MSE = 0.02$, $p = .378$, $\eta_p^2 = 0.02$. We ran further analyses to characterize this pattern.

An ANOVA contrasting the individual and collaborative retrieval conditions (II vs. CI) revealed the same pattern, with a significant main effect of item type, $F(2,116) = 3.89$, $MSE = 0.02$, $p = .023$, $\eta_p^2 = 0.06$, as well as a significant main effect of retrieval condition, $F(1,58) = 8.11$,

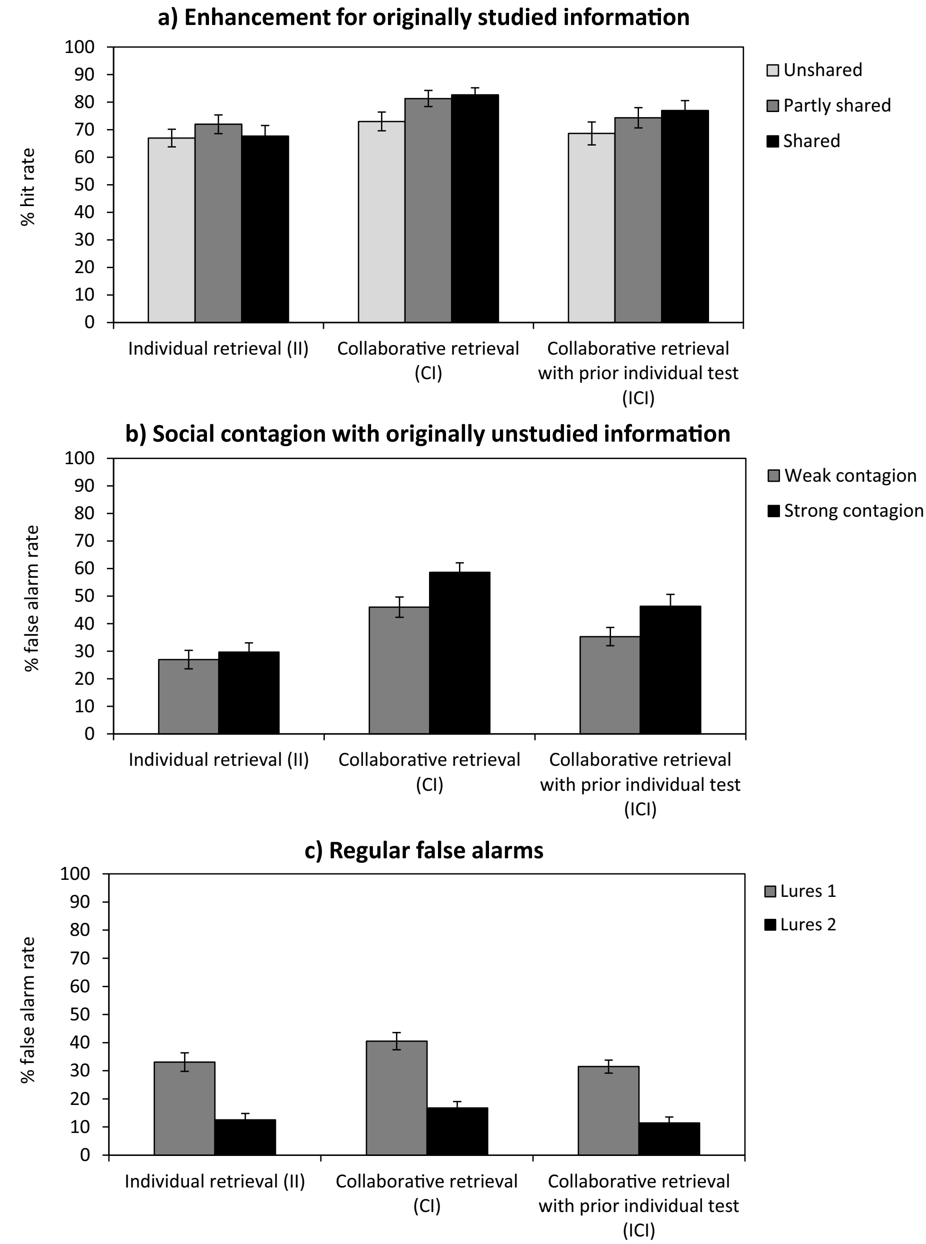

Fig. 4. Results of Experiment 2. Panel a) shows hit rates for originally studied information; panel b) shows social contagion with originally unstudied information; panel c) shows regular false alarm rates.

$MSE = 0.06$, $p = .006$, $\eta_p^2 = 0.12$, but no significant interaction of the two factors, $F(2,116) = 1.64$, $MSE = 0.02$, $p = .198$, $\eta_p^2 = 0.03$. Hit rates were enhanced in the collaborative relative to the individual retrieval condition, but only for fully shared items (82.7% vs. 67.7%), $t(58) = 3.25$, $p = .002$, $d = 0.83$, and partly shared items (81.3% vs. 72.0%), $t(58) = 2.09$, $p = .041$, $d = 0.54$. There was no corresponding facilitation for unshared items (73.0% vs. 67.0%), $t(58) = 1.29$, $p = .203$, $d = 0.33$.[3]

A further ANOVA contrasting the two collaborative conditions with and without a prior individual test (CI vs. ICI) only showed a significant main effect of ITEM TYPE, $F(2,116) = 7.88$, $MSE = 0.02$, $p = .001$, $\eta_p^2 = 0.12$, but no other significant effects, $Fs \leq 2.05$, $MSEs \leq 0.07$, $ps \geq .158$, $\eta_p^2 \leq 0.03$. Unshared items were generally remembered worse than partly and completely shared items (70.8% vs. 77.8% vs. 79.8%), $ts(59) \geq 2.76$, $ps \leq .008$, $ds \geq 0.36$; memory for these items did not differ, $t(59) < 1.00$, $p = .325$, $d = 0.13$.

3.2.2. Social contagion

Fig. 4b shows mean false alarms for items that were initially unstudied by each subject, but studied by one if not both other group member(s). A 2 × 3 ANOVA with the factors of ITEM TYPE (weak contagion items, strong contagion items) and retrieval condition (II, CI, ICI) showed a significant main effect of retrieval condition, $F(2,87) = 15.01$, $MSE = 0.06$, $p < .001$, $\eta_p^2 = 0.26$, a significant main effect of item type, $F(1,87) = 16.83$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.16$, but no significant interaction of the two factors, $F(2,87) = 2.09$, $MSE = 0.02$, $p = .130$, $\eta_p^2 = 0.05$. To clarify this pattern, we again ran further analyses.

An ANOVA contrasting the individual and collaborative retrieval conditions (II vs. CI) showed a significant main effect of retrieval condition, $F(1,58) = 30.93$, $MSE = 0.06$, $p < .001$, $\eta_p^2 = 0.35$, a significant

[3] Please note that none of the reported tests were corrected for multiple comparisons, because all main results relied on ANOVA outcomes. If Bonferroni corrections were applied to the just reported t-tests on facilitation effects (CI vs II), the comparison for partly shared items would no longer be significant.

main effect of item type, $F(1,58) = 10.90$, $MSE = 0.02$, $p = .002$, $\eta_p^2 = 0.16$, and a significant interaction of the two factors, $F(1,58) = 4.63$, $MSE = 0.02$, $p = .036$, $\eta_p^2 = 0.07$. Follow-up tests on false alarms in the individual retrieval condition alone again showed no differences between strong and weak contagion items (29.7% vs. 27.0%), $t(29) = 1.03$, $p = .310$, $d = 0.19$. In the collaborative retrieval condition, however, strong contagion items evoked higher false alarm rates than weak contagion items (58.7% vs. 46.0%), $t(29) = 3.28$, $p = .003$, $d = 0.60$. False alarm rates for both item types were enhanced in the collaborative vs. the individual retrieval condition, but social contagion was again larger for strong (58.7% vs. 29.7%), $t(58) = 6.01$, $p < .001$, $d = 1.55$, than for weak contagion items (46.0% vs. 27.0%), $t(58) = 3.82$, $p < .001$, $d = 0.99$.

To examine if individual retrieval prior to collaboration influenced social contagion, we conducted another ANOVA contrasting the two collaboration conditions. The ANOVA revealed a significant main effect of item type, $F(1,58) = 16.21$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = 0.22$, and a significant main effect of retrieval condition, $F(1,58) = 6.91$, $MSE = 0.06$, $p = .011$, $\eta_p^2 = 0.11$, but no significant interaction between the two factors, $F(1,58) < 1.00$, $MSE = 0.03$, $p = .778$, $\eta_p^2 = 0.001$. False alarm rates were also higher for strong compared to weak items in the collaborative retrieval condition with a prior individual test (46.3% vs. 35.3%), $t(29) = 2.48$, $p = .019$, $d = 0.45$. Relative to the condition without prior individual retrieval, false alarm rates were however significantly reduced, and this was the case for both strong contagion items (46.3% vs. 58.7%), $t(58) = 2.22$, $p = .030$, $d = 0.57$, and weak contagion items, (35.3% vs. 46.0%), $t(58) = 2.15$, $p = .036$, $d = 0.55$.[4]

3.2.3. False alarms for lures

Fig. 4c shows mean false alarms for lures. A 2 × 3 ANOVA with the factors of ITEM TYPE (lures 1, lures 2) and RETRIEVAL CONDITION (II, CI, ICI) showed a significant main effect of ITEM TYPE, $F(1,87) = 216.92$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = 0.71$, but no significant main effect of RETRIEVAL CONDITION, $F(2,87) = 2.89$, $MSE = 0.03$, $p = .061$, $\eta_p^2 = 0.06$, and no significant interaction of the two factors, $F(2,87) < 1.00$, $MSE = 0.01$, $p = .529$, $\eta_p^2 = 0.02$. False alarm rates were generally higher for lures 1 (introduced earlier in the experiment) than lures 2 (35.0% vs. 13.6%). In contrast to Experiment 1, collaborating with others did not significantly enhance false alarms for lures in Experiment 2.[5]

3.2.4. Additional analysis: performance on the initial individual test (ICI condition only)

Mean free recall was at 22.7%. Subjects recalled similar percentages of shared items (22.7%), partly shared items (23.5%), and unshared items (21.3%), $F(2,58) < 1.00$, $MSE = 0.01$, $p = .758$, $\eta_p^2 = 0.01$. On average, subjects generated 1.5 intrusions ($SD = 1.7$; range: 0–6), but intrusions never included contagion items that had been studied by other group members.

3.2.5. Additional analysis: test durations of individual vs. collaborative tests

With 8 s per item, the first individual test in the II condition always took 20 min. Test duration did not differ in the two collaborative retrieval conditions (CI: $M = 19.2$ min, $SD = 7.6$; ICI: $M = 16.3$ min, $SD = 6.6$), $t(18) < 1.00$, $p = .364$, $d = 0.42$. One-sample t-tests showed that test durations in the collaborative conditions did not significantly deviate from the fixed 20 min in the II condition, $ts(9) \leq 1.80$, $ps \geq .105$.

3.2.6. Additional analysis across Experiments 1 and 2: contrasting social contagion errors and regular false alarms for lures

In addition to the analyses above, conducted separately for social contagion items and lures, we followed a previous study by Kensinger, Choi, Murray, and Rajaram (2016) and compared the magnitude of these memory errors in the collaborative retrieval condition (CI) only. Such analysis can indicate if social contagion errors were similar in size as or even exceeded regular false alarms. For brevity, we collapsed the data of Experiments 1 and 2. Because social contagion arises during collaboration, we used false alarms to those lure sets as comparisons that were introduced during the collaboration phase, too. For Experiment 2, this corresponds to lures 1; for Experiment 1, false alarms were collapsed across lures 1 and lures 2.

A 3 × 2 ANOVA with the factors of item type (weak contagion items, strong contagion items, lures) and Experiment (Experiment 1, Experiment 2) showed a significant main effect of item type, $F(2,116) = 31.73$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.35$, but no significant main effect of experiment, $F(1,58) = 1.40$, $MSE = 0.05$, $p = .241$, $\eta_p^2 = 0.02$, and also no significant interaction of the two factors, $F(2,116) < 1.00$, $MSE = 0.02$, $p = .572$, $\eta_p^2 = 0.01$. False alarms were indeed enhanced for contagion items relative to lures, for both weak contagion items (48.7% vs. 40.8%), $t(59) = 3.56$, $p = .001$, $d = 0.46$, and strong contagion items (61.3% vs. 40.8%), $t(59) = 7.57$, $p < .001$, $d = 0.98$. Thus, social contagion errors exceeded regular false alarms for lures in the present experiments.

3.3. Discussion

Experiment 2 replicated most results of Experiment 1. Collaboration again enhanced individual memory, but only for information shared by group members. Simultaneously, collaboration induced social contagion with misinformation, which was again more pronounced when the misinformation was represented by both other group members. Most importantly, these parallels between experiments emerged even though test durations between conditions differed largely in Experiment 1, but were fairly well matched in Experiment 2, indicating that the effects observed in Experiment 1 were not due to the initial confound in task duration.

Similarly, individual retrieval prior to collaboration was again beneficial and reduced susceptibility to social contagion, replicating the results of Experiment 1. This parallel in findings arose even though the free-recall test applied in Experiment 2 resulted in relatively low recall rates. The finding may therefore generalize across different test formats and not depend critically on whether test formats match between the previous individual and all subsequent tests. In the present study, the finding arose in a collaborative retrieval task with real social interactions, but more broadly it is well in line with previous work, which showed that individual retrieval can reduce social contagion from fictitious social sources (e.g., Huff et al., 2013, 2016).

Even though the numerical pattern observed for false alarms for lures was similar between Experiments 1 and 2, the statistical results differed. In particular, prior collaboration had significantly enhanced false alarms in Experiment 1, but the corresponding comparisons were not significant in Experiment 2. The difference in results could have to do with the controlled (longer) duration of the critical individual recognition test in Experiment 2. At least potentially, this could have resulted in slightly higher false alarm rates in the II condition, even

[4] If Bonferroni corrections were applied to these t-tests on retrieval practice protecting from social contagion (CI vs. ICI), then the two single comparisons for strong and weak social contagion items would no longer be significant.

[5] Upon reviewer's request we conducted additional analyses to see if gender distribution in collaborative groups affected the results. Separate analyses were again carried out for facilitation effects, social contagion, and false alarms for lures. Comparing purely female/male groups (15 triads, n = 45) and mixed groups (5 triads, n = 15), no significant main effects or interaction effects involving the factor gender distribution were observed (all $ps \geq .092$, all $\eta_p^2s \leq 0.05$.)

Similar analyses were conducted to examine if facilitation and social contagion were affected by whether subjects within triads knew each other. Comparing completely unacquainted triads (7 triads, n = 21) to triads in which at least two (9 triads, n = 27) or all three participants knew each other (14 triads, n = 42), no significant main or interaction effects involving the factor acquaintance were found (all $ps \geq .211$, all $\eta_p^2s \leq 0.07$).

though hit rates should have shown a similar increase (which they did not). Additionally, the format of the initial individual test in the ICI condition was changed in Experiment 2. We only differentiated between two different types of lures, which also changed the overall make-up of the lure materials in the indvidual vs. collaborative retrieval condition. Although these procedural changes may have affected effect sizes and statistical outcomes, it seems important to stress that the general numerical pattern for lures was similar across Experiment 1 and 2.

An issue not addressed by analyzing hits and false alarms alone is what exactly happened when we asked participants in collaborative conditions to complete a recognition test together. For instance, one might wonder how group members interacted with each other in case of disagreements about old/new judgements. Going through the audio files that were recorded during collaboration, it seems that subjects mostly discussed how sure or unsure they were that a specific item had (or had not) been studied initially. If there was no agreement, participants frequently mentioned details that they remembered in order to demonstrate just how sure they were about a specific item (e.g., that the item had been presented early during the study phase, or right after some other word; or they even explained how they had processed and tried to memorize the item during study). The audio recordings of both experiments consistently demonstrated that the joint memory task was not only enjoyable for participants, but could be completed without any issues.

4. General discussion

The present study had the primary goal to introduce an experimental paradigm that allows simultaneous investigation of different aspects of real social interactions in the lab, and that also allows to examine advanced questions, as for instance the potentially modulating role of retrieval practice. Both experiments showed the same three main findings: First, collaboration enhanced memory for initially studied information, and this enhancement was mostly present for information shared across group members. Second, collaboration simultaneously distorted individual memory of the study phase, because subjects integrated initially unstudied information into their own memory of the study phase. Such social contagion was strongest when the information was represented by both other group members (relative to just one other group member). Third, individual retrieval prior to collaboration reduced social contagion and the distorting effect of collaboration, without reducing collaborative facilitation and its beneficial effect for initially studied (and shared) information. Additionally, collaboration also enhanced false alarms for lures in Experiment 1, but the corresponding contrast was not significant in Experiment 2.

Our first main finding connects very well with previous work on social memory. Enhancement of individual memory after collaboration has rather consistently been observed in studies applying the regular collaborative retrieval paradigm, in which subjects typically study the same (fully shared) information (e.g., Abel & Bäuml, 2017; Bärthel et al., 2017; Blumen et al., 2014; Blumen & Stern, 2011; Weldon & Bellinger, 1997). In general, such collaborative enhancement is ascribed to reexposure effects, with joint group recall reexposing single group members to information that they would have forgotten if recalling by themselves. The present study adds to this literature by showing that enhancement effects are indeed limited to information initially shared by group members, which is in good agreement with the assumed reexposure mechanism. If information was initially studied by all group members, at least one of them may remember it during the collaborative test, thus providing a chance for reexposure to the other group members who may have forgotten it. In contrast, if information was initially only studied by one group member, but not remembered, this miss cannot be corrected by input from the other group members, who never studied the word. Collaboration should therefore provide little chance for reexposure with unshared information, and post-collaborative facilitation should be reduced or even eliminated, which is is exactly what the present data show (for related findings, see also Choi et al., 2017). In addition, the present experiments revealed another interesting aspect of collaborative remembering. Across collaborative and individual retrieval conditions, memory for initially studied, but unshared information did not differ. This indicates that, during collaboration, subjects were not "talked out of" their individually held memories by the other two group members. Such memories were simply not affected by collaboration, showing that collaboration did not result in "censoring" of unshared information.

Our second main finding also connects well to other prior work. Social contagion effects have previously been examined by means of different tasks, with many studies focusing on dyads and applying confederates or fictitious recall protocols as social sources of misinformation (e.g., Gabbert et al., 2004; Numbers et al., 2014; Reysen, 2005; Roediger et al., 2001). In the present study, we adopted a different approach and manipulated the distribution of information that was initially encoded (thus following previous work on social contagion in eyewitness memory, e.g. Gabbert et al., 2003; Garry et al., 2008; Mori, 2003; see also Schneider & Watkins, 1996). In line with this prior work, the present study found that collaboration resulted in social transmission of information, with single subjects integrating information that they themselves had not initially encoded into their own memories of the study phase. Social contagion is not limited to dyads, but can be observed in freely interacting triads as well (see also Choi et al., 2017). Moreover, the present data go beyond the prior work by showing that social contagion is stronger when potentially distorting information is represented by two rather than by only one other group member.

One issue that may be of particular relevance for social contagion and that the present experiments can not address is the question of how social pressure influences the effects. Our approach to collaborative recognition tests was based upon previous work on collaborative recognition (e.g., Blumen & Rajaram, 2009; Thorley & Dewhurst, 2009), in which collaborating groups were asked to provide one group response only. Arguably, this approach may have increased social pressure on participants to develop and adopt group consensus, and such social pressure might increase susceptibility to social contagion (e.g., Reysen, 2007). Other previous studies applied a slightly different type of collaborative recognition test, with participants discussing each item collaboratively, but then providing individual recognition responses (Kensinger et al., 2016; Rajaram & Pereira-Pasarin, 2007). Because such procedure does not require any type of consensus, it might be argued that it could reduce social contagion rates. To the best of our knowledge, no study so far has directly compared different types of collaborative recognition tests, which should be a high priority in future work on the issue.

Our third main finding on the influence of prior individual retrieval also connects to previous studies, even though results in the literature are more mixed with regard to this issue. In general, our finding of beneficial effects of prior individual retrieval fits well with literature on the testing effect, indicating that retrieval practice can have multiple benefits (e.g., Halamish & Bjork, 2011; for an overview, see Roediger, Putnam, & Smith, 2011). In particular, the finding that individual retrieval prior to collaboration can reduce susceptibility to social contagion errors mirrors previous findings by Huff et al. (2013, 2016). Because these prior studies relied on recall protocols of fictitious subjects to introduce misinformation, the present data might indicate that the influence of prior individual retrieval generalizes across different variants of social contagion paradigms and that prior individual retrieval continues to be beneficial when real social interactions are examined.

There are however further reports in the literature that call into question whether prior individual retrieval has a uniform protective influence against misinformation. Several other previous studies on misinformation effects (without a focus on social contagion) indeed observed the opposite pattern, with prior individual retrieval increasing susceptibility to misinformation (e.g., Chan et al., 2009; Chan & LaPaglia, 2011; Gordon & Thomas, 2014; for a review, see Chan, Manley, & Lang, 2017). A potentially critical factor may be how misinformation relates to initially studied information. In particular, in studies reporting that individual retrieval decreased susceptibility to misinformation (as in the present study), the misinformation contained additive information, which was not initially studied by specific individuals, but could well have been present at study in addition to what participants actually encoded. In contrast, in studies reporting that individual retrieval increased susceptibility to misinformation, the misinformation contained contradicting information that conflicted with what participants had initally studied (e.g., different traffic signs than actually encountered ones, or diverging characteristics of a perpetrator). Considering differences in outcomes between studies, Huff et al. (2013) noted that retrieval practice might trigger different types of processing of the misinformation depending on how it is connected to the initial information, thereby sometimes increasing and sometimes decreasing susceptibility to the misinformation (for further discussion, see Chan et al., 2017). We think the novel task introduced in this manuscript is perfectly suited to test this proposal in future studies.

Although pursuing different research questions, a previous study by Choi et al. (2017) also manipulated information distribution in collaborating triads. The manipulation was implemented such that only a single participant in each triad studied information that was unshared with the other two participants, plus some information that was fully shared across all three participants (the other two participants shared all initially studied information). In contrast, in the present study, each individual in a triad initially studied some information that was fully shared across the whole group, some information that was only partially shared with one other group member, and some information that was not shared with any of the group members. This enabled us to directly compare facilitation and contagion effects as a function of information distribution and, thus, to examine questions not addressed in Choi et al. (2017). In particular, the present study suggests that collaborative facilitation effects for subsequent individual memory are mostly present for information that was shared within the group, and that social contagion effects are larger for information represented by two group members (rather than just one). The present study thus goes beyond the previous work and adds to our understanding of effects of information distribution in social groups.

Social interactions have been proposed to result in the emergence of collective memories, i.e., of memories that are shared by members of (larger) groups and that often have a bearing on the group's social identity (e.g., Hirst & Manier, 2008; Wertsch & Roediger, 2008). In particular, previous work on collaborative memory retrieval has shown that individuals who previously engaged in joint recall with other group members show a higher degree of mnemonic overlap, such that they end up remembering similar contents (rather than unique ones; e.g., Blumen & Rajaram, 2008; Cuc, Ozuru, Manier, & Hirst, 2006). Together with prior studies (Choi et al., 2017; Cuc et al., 2006), the present data may suggest that this development of shared memories may not only include information available to all group members during encoding, but also information that is initially only known by some or even single individuals within the group. Corresponding statistical analyses on the degree of overlap could not be run for the present data due to the format of our final individual recognition test (with tested items being randomly drawn out of larger pools of items for each participant), but a least numerically, data reported by Choi et al. (2017) indicate that different sources of information are likely pieced together during collaboration, including information that is encountered for the first time in this context. In this manner, the emergence of shared memories may rely on both collaborative facilitation effects for initially encountered information and social transmission effects for information not encountered earlier. Examining different types of misinformation might also be interesting from this perspective, because the creation of shared memories across all group members should be easier when single pieces of information remembered by individuals are additive rather than contradictory. Again, the task introduced here seems perfectly suited to pursue this question in future studies.

In particular, our hope is that the newly introduced recognition-based collaboration task may facilitate investigations into social memory more generally, because it allows for systematic variations that are relatively easy to implement. For instance, the task allows for variations of encoding factors (e.g., amount and type of initially studied information, encoding duration and practice opportunities). It also allows for variations on the group level (e.g., group size, how well group members know each other, how they are asked to interact with each other, etc.) and for variations at test (e.g., the type of final test, remember/know or source monitoring judgements, etc.). In addition, future studies could also add further tasks to the protocol, for instance to examine how prior collaboration and the development of "shared" (true and false) memories may affect performance on subsequent tasks that depend on the information in one way or another.

5. Conclusions

The present study introduced a novel recognition-based task that allows examining multiple mnemonic consequences of real social interactions in the lab. The results of two experiments show robust collaborative facilitation and social contagion effects that vary with the degree to which the information is shared across group members. In addition, the task also allows investigating advanced questions on social memory. Both experiments explored the role of prior individual retrieval practice for social contagion, showing that retrieval practice can reduce susceptibility to social contagion. The task opens up new avenues for systematic research into social memory effects, because it not only allows examining real social interactions in the lab, but also provides the stage for relatively straightforward experimental manipulations.

Authors' note

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The data and materials for all experiments are available on the Open Science Framework (https://osf.io/s74vp/).

Declaration of competing interest

The authors have no competing interests to declare.

Appendix A. Individual vs. collaborative recognition test in Experiment 1

Mean hits and false alarms on this test are shown in Table A1. Note that numbers of observations differ for individual ($n = 30$ subjects) and collaborative retrieval conditions ($n = 10$ triads), which precludes direct comparisons.

For the two collaborative conditions, an ANOVA for initially studied information (shared, partly shared, unshared) revealed significant main effects for ITEM TYPE and RETRIEVAL CONDITION, $Fs \geq 7.01$, $MSEs \leq 0.02$, $ps \leq .016$, $\eta_p^2 \geq 0.28$, but no significant interaction between the two factors, $F(2,36) = 1.62$, $MSE = 0.01$, $p = .211$, $\eta_p^2 = 0.08$. In general, groups were more critical in the ICI condition and accepted fewer items as old compared to the CI condition, but in single comparisons this difference only reached significance for unshared items, $t(18) = 2.53$, $p = .021$, $d = 1.13$ (other comparisons: $ts(18) \leq 1.59$, $ps \geq .129$, $ds \leq 0.71$).

A corresponding ANOVA for lures revealed a significant main effect for ITEM TYPE, $F(1,18)=12.07$, $MSE=0.01$, $p=.003$, $\eta_p^2=0.40$, no significant main effect for RETRIEVAL CONDITION, $F(1,18)=3.06$, $MSE=0.01$, $p=.098$, $\eta_p^2=0.15$, and a significant interaction between the two factors, $F(1,18)=16.01$, $MSE=0.01$, $p=.001$, $\eta_p^2=0.47$. False alarm rates for lures 1 were comparable across the two collaborative conditions, $t(18)=1.16$, $p=.263$, $d=0.52$, but groups were better at rejecting items from the second set of lures after completing a prior individual test, $t(18)=5.39$, $p < .001$, $d=2.41$.

In the individual retrieval condition, false alarms for lures 1 and lures 2 were comparable, $t(29) < 1.00$, $p=.959$, $d=0.01$. Some other means are hard to interpret in this condition. In particular, partly shared and unshared items on this test comprise all such items that were studied across the whole group - but each subject within a group only studied two thirds of the partially shared and one third of the unshared items. Recall rates therefore differ markedly for these items types in the individual retrieval condition, but these differences are only due to item types reflecting information distribution across the whole group (not for every single subject). For reasons of completeness, we report mean hits in Table A1, but will not conduct statistical tests comparing hits for the three item types.

Table A1
Mean hits (unshared, partly shared, shared items) and false alarm rates (lures 1 and 2) on the individual vs. collaborative recognition test, separately for the three retrieval conditions in Experiment 1. Standard deviations are provided in parentheses.

	Unshared	Partly shared	Shared	Lures 1	Lures 2
Individual retrieval (II)	34.77	54.83	71.30	21.87	21.97
	(12.38)	(13.61)	(18.10)	(16.06)	(13.05)
Collaborative retrieval (CI)	59.20	84.10	92.20	26.40	28.00
	(11.08)	(6.94)	(6.55)	(9.79)	(8.89)
Collaborative retrieval with prior individual test (ICI)	44.70	79.00	85.70	32.70	10.00
	(14.31)	(9.90)	(11.15)	(14.18)	(5.70)

Appendix B. Individual vs. collaborative recognition test in Experiment 2

Mean hits and false alarms on this test are shown in Table B1. Numbers of observations were again different for individual (n = 30 subjects) and collaborative retrieval conditions (n = 10 triads).

For the two collaborative conditions, an ANOVA for initially studied information (shared, partly shared, unshared) only revealed a significant main effect for ITEM TYPE, $F(2,36)=195.85$, $MSE=0.004$, $p < .001$, $\eta_p^2=0.92$, but no other significant effects, $Fs < 1.00$, $ps \geq .484$, $\eta_p^2 s=0.03$. Hit rates were higher for shared vs. partly shared items and for partly shared vs. unshared items, and this was the case in both collaborative retrieval conditions, $ts(9) \geq 2.77$, $ps \leq .022$, $ds \geq 0.88$. False alarms for lures did not differ between the two collaborative retrieval conditions, $t(18)=1.14$, $p=.269$, $d=0.51$.

For completenes, mean hits and false alarms in the individual retrieval condition are also shown in Table B1. As in Experiment 1, performance should be evaluated with caution for this condition, because it most likely reflects information distribution across groups (see Appendix A for details).

Table B1
Mean hits (unshared, partly shared, shared items) and false alarm rates (lures 1) on the individual vs. collaborative recognition test, separately for the three retrieval conditions in Experiment 2. Standard deviations are provided in parentheses.

	Unshared	Partly shared	Shared	Lures 1
Individual retrieval (II)	37.70	55.43	71.27	28.13
	(9.23)	(11.43)	(15.03)	(14.20)
Collaborative retrieval (CI)	55.20	85.80	92.70	24.70
	(10.02)	(8.54)	(5.27)	(11.31)
Collaborative retrieval with prior individual test (ICI)	52.40	81.00	92.40	19.40
	(11.24)	(14.39)	(6.31)	(9.36)

References

Abel, M., & Bäuml, K.-H. T. (2017). Collaborative remembering revisited: Study context access modulates collaborative inhibition and later benefits for individual memory. *Memory & Cognition, 45*, 1319–1334.
Abel, M., & Bäuml, K.-H. T. (2019). List-method directed forgetting after prolonged retention interval: Further challenges to contemporary accounts. *Journal of Memory and Language, 106*, 18–28.
Abel, M., & Roediger, H. L., III (2017). Comparing the testing effect under blocked and mixed practice: The mnemonic benefits of retrieval practice are not affected by practice format. *Memory & Cognition, 45*, 81–92.
Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 940–945.
Bärthel, G. A., Wessel, I., Huntjens, R. J. 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*, 1176–1189.
Basden, B. H., Basden, D. R., & Henry, S. (2000). Costs and benefits of collaborative remembering. *Applied Cognitive Psychology, 14*, 497–507.
Bäuml, K.-H. T., Holterman, C., & Abel, M. (2014). Sleep can reduce the testing effect - It enhances recall of restudied items but can leave recall of retrieved items unaffected. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 1568–1581.
Blumen, H. M., & Rajaram, S. (2008). Influence of re-exposure and retrieval disruption during group collaboration on later individual recall. *Memory, 16*, 231–244.
Blumen, H. M., & Rajaram, S. (2009). Effects of repeated collaborative retrieval on individual memory vary as a function of recall versus recognition tasks. *Memory, 17*, 840–846.
Blumen, H. M., & Stern, Y. (2011). Short-term and long-term collaboration benefits on individual recall in younger and older adults. *Memory & Cognition, 39*, 147–154.
Blumen, H. M., Young, K. E., & Rajaram, S. (2014). Optimizing group collaboration to improve later retention. *Journal of Applied Research in Memory and Cognition, 3*, 244–251.
Brickenkamp, R., & Zillmer, E. A. (1998). *d2 test of attention.* Göttingen, Germany: Hogrefe & Huber.
Chan, J. C. K., & LaPaglia, J. A. (2011). The dark side of testing memory: Repeated retrieval can enhance eyewitness suggestibility. *Journal of Experimental Psychology: Applied, 17*, 418–432.
Chan, J. C. K., & LaPaglia, J. A. (2013). Impairing existing declarative memory in humans by disrupting reconsolidation. *Proceedings of the National Academy of Sciences of the*

United States of America, 110, 9309–9313.
Chan, J. C. K., Manley, K. D., & Lang, K. (2017). Retrieval-enhanced suggestibility: A retrospective and a new investigation. *Journal of Applied Research in Memory and Cognition, 6*, 213–229.
Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science, 20*, 66–73.
Choi, H.-Y., Kensinger, E. A., & Rajaram, S. (2017). Mnemonic transmission, social contagion, and emergence of collective memory: Influence of emotional valence, group structure, and information distribution. *Journal of Experimental Psychology: General, 146*, 1247–1265.
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*, 535–551.
Cuc, A., Koppel, J., & Hirst, W. (2007). Silence is not golden: A case for socially shared retrieval-induced forgetting. *Psychological Science, 18*, 727–733.
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.
Duyck, W., Desmet, T., Verbeke, L. P. C., & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers, 36*, 488–499.
Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology, 9*, 215–227.
Gabbert, F., Memon, A., & Allen, K. (2003). Memory conformity: Can eyewitnesses influence each other's memorie for an event? *Applied Cognitive Psychology, 17*, 533–543.
Garry, M., French, L., Kinzett, T., & Mori, K. (2008). Eyewitness memory following a discussion: Using the MORI technique with a Western sample. *Applied Cognitive Psychology, 22*, 431–439.
Gordon, L. M., & Thomas, A. K. (2014). Testing potentiates learning in the misinformation paradigm. *Memory & Cognition, 42*, 186–197.
Gordon, L. M., & Thomas, A. K. (2017). The forward effects of testing on eyewitness memory: The tension between suggestibility and learning. *Journal of Memory & Language, 95*, 190–199.
Halamish, V., & Bjork, R. A. (2011). When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 801–812.
Hirst, W., & Echterhoff, G. (2012). Remembering in conversations: The social sharing and reshaping of memories. *Annual Review of Psychology, 10*, 55–79.
Hirst, W., & Manier, D. (2008). Towards a psychology of collective memory. *Memory, 16*, 183–200.
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.
Huff, M. J., Weinsheimer, C. C., & Bodner, G. E. (2016). Reducing the misinformation effect through initial testing: Take two tests and recall me in the morning? *Applied Cognitive Psychology, 30*, 61–69.
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*, 706–716.
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.
Maswood, R., & Rajaram, S. (2018). Social transmission of false memory in small groups and large networks. *Topics in Cognitive Science, 2018*, 1–23.
Meade, M. L., & Roediger, H. L., III (2002). Explorations in the social contagion of memory. *Memory & Cognition, 30*, 995–1009.
Mori, K. (2003). Surreptitiously projecting different movies to two subsets of viewers. *Behavior Research Methods, Instruments, & Computers, 35*, 599–604.
Muller, F., & Hirst, W. (2014). Remembering stories together: Social contagion and the moderating influence of disagreements in conversations. *Journal of Applied Research in Memory and Cognition, 3*, 7–11.
Numbers, K. T., Meade, M. L., & Perga, V. A. (2014). The influences of partner accuracy and partner memory ability on social false memories. *Memory & Cognition, 42*, 1225–1238.
Oswald, W. D., & Roth, E. (1987). *Der Zahlenverbindungstest (ZVT) [Connect-the-Numbers Test]*. Göttingen, Germany: Hogrefe.
Potts, R., & Shanks, D. R. (2012). Can testing immunize memories against interference? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1780–1785.
Rajaram, S., & Pereira-Pasarin, L. P. (2007). Collaboration can improve individual recognition memory: Evidence from immediate and delayed tests. *Psychonomic Bulletin & Review, 14*, 95–100.
Raven, J. (2000). The Raven's Progressive Matrices: Change and stability over culture and time. *Cognitive Psychology, 41*, 1–48.
Reysen, M. B. (2005). The effects of conformity on recognition judgements. *Memory, 13*, 87–94.
Reysen, M. B. (2007). The effects of social pressure on false memories. *Memory & Cognition, 35*, 59–65.
Roediger, H. L., III, Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review, 8*, 365–371.
Roediger, H. L., III, Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In J. Mestre, & B. Ross (Eds.). *Psychology of learning and motivation: Cognition in education* (pp. 1–36). Oxford: Elsevier.
Schneider, D. M., & Watkins, M. J. (1996). Response conformity in recognition testing. *Psychonomic Bulletin & Review, 3*, 481–485.
Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1392–1399.
Thorley, C., & Dewhurst, S. A. (2009). False and veridical collaborative recognition. *Memory, 17*, 17–25.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
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.
Wertsch, J. V., & Roediger, H. L., III (2008). Collective memory: Conceptual foundations and theoretical approaches. *Memory, 16*, 318–326.
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.