COGNITIVE SCIENCE
A Multidisciplinary Journal

Check for updates

Cognitive Science 49 (2025) e70060
© 2025 The Author(s). *Cognitive Science* published by Wiley Periodicals LLC on behalf of Cognitive Science Society (CSS).
ISSN: 1551-6709 online
DOI: 10.1111/cogs.70060

Social Network Structure Shapes the Formation of True and False Memories at the Collective Level

Tania Valle,[a,b] Annamaria Krizovenska,[b] Josué García-Arch,[b,c] Maria Teresa Bajo,[d,e] Lluís Fuentemilla[b,c,f]

[a]*Department of Psychology, University of Jaen*
[b]*Department of Cognition, Development and Education Psychology, Faculty of Psychology, University of Barcelona*
[c]*Institute of Neuroscience (UBNeuro), University of Barcelona*
[d]*Mind, Brain and Behaviour Research Center, University of Granada*
[e]*Department of Experimental Psychology, University of Granada*
[f]*Bellvitge Institute for Biomedical Research*

Received 3 December 2024; received in revised form 19 March 2025; accepted 24 March 2025

Abstract

Societal structures and memory organization models share network-like features, offering insights into how information spreads and shapes collective memories. In this study, we manipulated the structure of lab-created community networks during a computer-mediated recall task using the Deese–Roediger–McDermott paradigm to test the spreading activation theory of true and false memory formation. We hypothesized that social network structure, whether clustered or not, would influence memory accuracy. Our results showed that clustered networks reinforced true memories by promoting mnemonic convergence, while non-clustered networks led to more false memories by increasing widespread cross-activation. These findings highlight how social network topology impacts memory dynamics and collective knowledge evolution.

Keywords: Collective memory; True and false memories; Mnemonic convergence; Social network

Correspondence should be sent to Lluís Fuentemilla, Department of Cognition, Development and Educational Psychology, University of Barcelona, Pg Vall Hebrón, 171, 08035 Barcelona, Spain. E-mail: llfuentemilla@ub.edu

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

1. Introduction

Human societies are structured as interconnected networks, enabling the exchange of information between individuals and groups. How groups share and remember their past shapes collective narratives, traditions, and identity (Hirst, Yamashiro, & Coman, 2018; Roediger et al., 2019). Similarly, in cognitive psychology, memory is viewed as a network where the activation of one node spreads to associated nodes, influencing memory and cognition (Anderson, 1983; Collins & Loftus, 1975; McClelland & Rumelhart, 1985). Like individuals in a community, memories are linked to other elements within their network and can be influenced by concurrent activations. Thus, exploring the shared organizational and propagation properties between societal and memory structures is a compelling area of study.

One way to explore the interaction between social networks and memory is by examining how memories are shaped through social communication. Research shows that memories remain malleable after encoding (Schacter, 2012), and conversational recall can selectively reinforce or weaken memories, aligning them among participants (Coman, Manier, & Hirst, 2009; Congleton & Rajaram, 2014). As these influences spread through social interactions, they contribute to collective memories (Coman, Momennejad, Drach, & Geana, 2016; Hirst & Echterhoff, 2012). Clustered networks, with tightly interconnected individuals, tend to develop more convergent memories, as information is reinforced and propagated within clusters. In less clustered networks, information can permeate more sparsely across the entire community network, thereby promoting the coactivation of a broader range of encoded memories at a collective level (Coman et al., 2016; Hirst et al., 2018). Thus, network structure influences how information spreads and how collective memories are constructed.

Similarly, the idea of memories as network-like structures, where activating one node spreads to others, has offered insights into the organization of stored representations and their interaction with new experiences (Anderson, 1983). For instance, in semantic priming, recognizing a word (e.g., *doctor*) is faster if preceded by a related word (e.g., *nurse*), compared to an unrelated one (e.g., *house*) because activation of one memory node directly influences the activation of a closely related node (Neely, 1977, 1991). However, the activation of one memory node can spread more broadly through the network influencing not only directly related nodes but also indirectly related nodes. This can lead to the recall or recognition of items that were never presented (i.e., false memories) but are strongly associated with the encoded items. Perhaps the most widely used tool to investigate false memories is the Deese–Roediger–McDermott (DRM) task (Deese, 1959; Roediger & McDermott, 1995). In this design, participants encode a list of words (e.g., *bed, tired, rest, nap, dream, wake, snooze, blanket, yawn, drowsy*) semantically associated with a non-presented critical word (e.g., *sleep*) that is falsely recalled in a subsequent memory test by the participants. In the DRM task, false memories arise from the cumulative activation of a critical word triggered by encoded words in a semantic cluster. As the number of encoded words increases, the activation of related concepts spreads more widely within the memory network. This broader activation enhances the likelihood that non-studied but semantically related words, known as critical lures, will be mistakenly recalled as if they had been presented during encoding. This effect occurs because the increased activation makes it more difficult for individuals to accurately differentiate between true memories—those based on studied words—and false recollections triggered by related but non-presented words (Robinson & Roediger, 1997). Conversely, when fewer words are encoded, memory for the studied items becomes more precise and robust. With a smaller set of learned words, activation remains more localized, reducing the likelihood of spreading to related but unstudied concepts. As a result, individuals experience fewer false memories, and the recall of critical lures diminishes (Robinson & Roediger, 1997). This reduction in activation spread allows for a clearer distinction between actual memories and intrusions, thereby improving the accuracy of recall.

The interplay between network clustering and memory processes raises distinct but related hypotheses regarding recall rates and mnemonic convergence. At the individual level, memory recall in the context of the DRM paradigm is influenced by the extent to which encoded words activate related concepts, affecting both true and false memories. In highly clustered social networks, communication between individuals selectively reinforces specific memories, potentially leading to higher recall rates of true memories and a decrease in false memories due to more constrained activation spread. Conversely, in less clustered networks, where information propagates more broadly across the entire community, a larger number of words may be reinforced, increasing the activation of false memories.

At the collective level, network clustering also influences mnemonic convergence or the extent to which individuals within a group come to share similar memories after social interaction. In clustered networks, repeated exposure to the same information within tightly connected subgroups should promote stronger alignment in what is remembered, leading to greater convergence on true memories. In contrast, in non-clustered networks, where information is exchanged more diffusely, mnemonic convergence may be weaker, particularly for true memories, as individuals are exposed to a wider range of recalled items, including lures.

To test this hypothesis, we asked 170 healthy individuals to participate in a memory experiment using online recruitment systems from the local institution. Following previous research (Coman et al., 2016), the experiment included four phases, each of them conducted with network community groups of 10 participants each that completed them on separate computers (Fig. 1). In the pre-conversational study phase (Phase I), each participant encoded 100 words presented on a computer. The stimuli included 10 wordlists from different semantic categories, each associated with a non-presented critical lure word (Table S1). Subsequently, during the pre-conversational recall phase (Phase II), each participant was asked to individually recall the studied words by typing them in a textbox on their computer. This was followed by the conversational recall phase (Phase III), wherein participants from the 10-member communities engaged in paired conversations with three partners, collectively recalling the studied content. These interactions occurred in a chat-like computer-mediated environment where participants typed their responses in a turn-taking manner. Last, in the post-conversational recall phase (Phase IV), participants freely recalled the initially studied word lists again.

In the conversational recall phase, each participant completed three conversational free recalls with three different group members within the network community, pre-arranged experimentally. They were tasked with collaboratively recollecting as many words as possible from the studied wordlists. Within the clustered condition (n = 80 participants; eight 10-member networks), interactions followed a network structure with two subclusters.

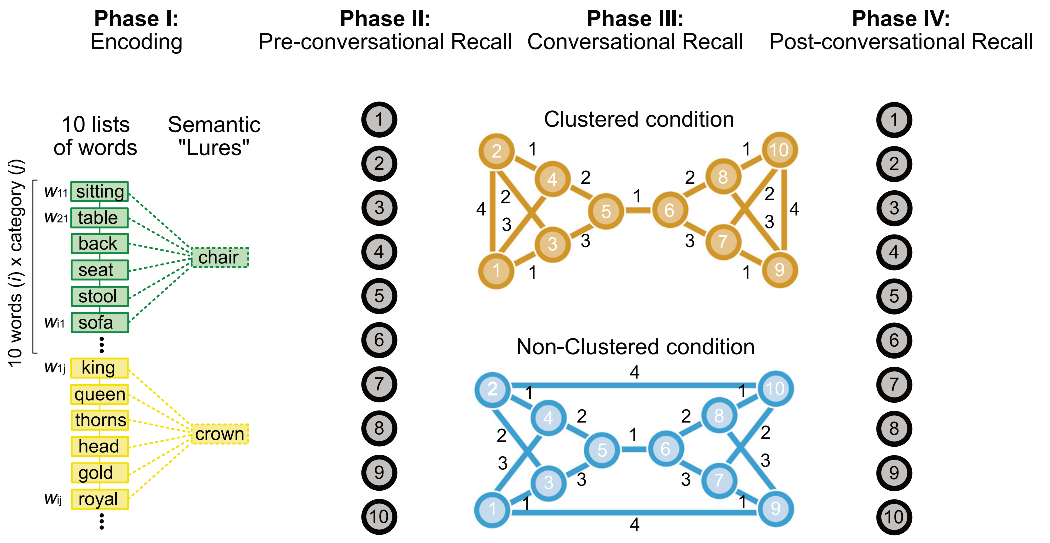

Fig. 1. Experimental design. Phases of the experiment involve participants initially learning 10 lists of semantically related words (Phase I). In the pre-conversational (Phase II) and post-conversational (Phase IV) recall phases, 10 participants individually recollect the learned information. The conversational recall phase (Phase III) includes participants, indicated by white numbers, in either the clustered (top) or non-clustered (bottom) condition. Participants are depicted as circles, and interactions are represented by links. The order of the sequential conversations between paired participants is indicated by numbers in black.

Conversely, in the non-clustered condition ($n = 90$ participants; nine 10-member networks), interactions occurred in a single large cluster. As in Coman et al. (2016) study, the global clustering coefficient, C (Freeman, 1978; Griffiths, Lewandowsky, & Kalish, 2013), contrasted between the clustered condition (C = 0.40) and the non-clustered condition (C = 0.00), thereby setting up an experimental design in which both network conditions were made comparable regarding factors such as the number of participants per network, the sequence of conversational interactions, and each participant's involvement in three conversations within their respective network.

2. Methods

2.1. Participants

Following previous studies using similar experimental designs, we aimed to recruit 10 groups of 10 participants each for each experimental condition. The statistical power afforded by this sample size was deemed adequate given the effect sizes obtained in previous studies using a similar sample size and experimental paradigm (Coman et al., 2016; Vlasceanu, Morais, Duker, & Coman, 2020). However, due to technical problems with the need to run the experiment synchronically within groups of 10 people, the final sample included in the study consisted of 170 healthy participants (78.2% females) with a mean age of 24.7

(*SD* = 7.3). All participants were native or highly proficient Spanish speakers. The study was advertised on a platform for students affiliated with the University of Barcelona or the University of Granada. Additionally, flyers of the study were posted around the university campuses and shared on social media. The participants were self-selected and received either two-course credits or 5 euros as compensation. For each group of 10 participants, an additional prize draw of 20 euros was conducted. Informed consent was obtained upon familiarization with the experimental procedure before the experiment onset. The study was approved by the University of Barcelona Ethics Committee.

2.2. Materials

The DRM paradigm was used to collect data in the present study. Ten DRM word lists used in the paradigm were adopted from Alonso, Fernández, and Díez (2004) and attached in Table S1. Each list contained 10 words semantically related to a critical non-presented word—a lure. For example, (translated from Spanish) *wind, breathe, fresh*, and so forth were the presented words associated with the critical non-presented word *air*. The distractor task consisted of 36 arithmetical problems (e.g., (12 / 4) + 4 = 7) with a "yes" or "no" answer, each presented for 5 s. The experimental task was programmed using the Qualtrics platform (Qualtrics.com), specifically its branch SMARTRIQS (smartriqs.com), which allows interactive online experiments to be programmed.

2.3. Design and procedure

The task (Fig. 1) was completed in synchrony in groups of 10 individuals via a computer. The participants completed it either on-site in the university computer room or remotely. Participants who were physically present at the task were seated at individual computers within a spacious room. They were not provided with information regarding which among the potential others were also engaged in the same task. Participants received full instructions in person or via Google Meets, then they provided the consent form, and then they started the task.

Following Coman et al. (2016) study, we defined two network structures—clustered and non-clustered, each of them including 10 participants. In the non-clustered condition, the participants were equally connected to all the individuals in the network. In the clustered condition, the network was split into two subclusters of five individuals who were connected by only one individual from each cluster. In the non-clustered condition, individuals were connected in an unconstrained manner to other individuals in the network. Individuals performed the task on their computers and interacted with each associated member in the conversational phase via the chat box that appeared during the conversational phase on each of their computers. Participants knew that 10 other individuals were concurrently engaged in the task, but they had no direct interaction with anyone except during the conversational phase when using their computers.

The task started with an encoding phase, followed by a distractor task. In the encoding phase, the participants were asked to memorize words presented on the screen for 2 s each. The order of the DRM word lists was random for each participant, but the order of words within each DRM list was kept constant. In the distractor task, the participants were asked to indicate whether the arithmetical problem solution was correct or incorrect by clicking the corresponding button. After completing the distractor task, the experiment continued with three distinct recall phases: a pre-conversational individual recall, followed by a conversational recall, and a post-conversational individual recall. The participants automatically entered the pre-conversational individual recall phase after finishing the distractor task. They were asked to type down all the words they remembered from the encoding phase for a maximum of 6 min. Subsequently, in the conversational phase, each participant was automatically paired and connected with another participant. Each participant completed three conversational recalls with three different group members. The pairs of participants entered a chat window shown in their individual computers where they were asked to recall words in collaboration by taking turns, as false recall was greater in turn-taking groups, compared to both the free-for-all and nominal groups (Basden, Basden, Bryner, & Thomas, 1997; Meade & Roediger, 2009; Thorley & Dewhurst, 2007) and when individuals are allowed to free-flowing collaboration (Barber, Rajaram, & Aron, 2010). Each pair member recalled one word at a time in their computer and subsequently waited for the other pair member to recall and share their word. In case they could not remember a word, they had the option to skip the turn by writing "pass" in the chat. In this phase, the groups of 10 were organized either into the clustered or non-clustered network. The total time of each conversation was 5 min. After the chat, the participants entered the post-conversational recall phase, which was identical to the first individual recall phase. The total duration of the experiment was approximately 45 min.

3. Results

3.1. Social network structure modulates the recall of true and false memories

We first examined whether participants' recall for studied and non-studied lure words changed after collaborative recall as a function of network type. To assess for this possibility, we calculated the recall rate for studied and critical lures before and after collaborative recall between participants of the clustered and the non-clustered network conditions. This analysis involved quantifying the recall rate for true and false items for each individual at the pre-conversational (Phase II) and at the post-conversational (Phase IV) recall phase. A mixed factorial ANOVA, with recall type (true vs. false) and time (pre-conversation vs. post-conversation) as within-subjects factors and network condition (clustered vs. non-clustered) as a between-subject factor, revealed a significant main effect of recall type ($F(1,168) = 23.08$, $p < .01$, $\eta^2 = 0.12$) and a main effect of time ($F(1,168) = 149.73$, $p < .01$, $\eta^2 = 0.47$) but not a significant interaction recall type $\times$ time ($F(1,168) = 0.14$, $p < .91$, $\eta^2 < 0.01$). The results indicated that participants recalled, overall, a greater number of true than lure words during the experiment but that their recall rate increased for both true and lure words in the post-conversational (Phase IV) when compared to pre-conversational (Phase II) recall phase (Table S2). However, we found a non-significant recall type $\times$ group ($F(1,168) = 0.31$, $p = .58$, $\eta^2 = 0.002$) nor time $\times$ group ($F(1,168) = 0.53$, $p = .47$, $\eta^2 = 0.003$) or a recall type $\times$ time ($F(1,168) < 0.01$, $p = .91$, $\eta^2 < 0.001$) interaction but a significant triple recall

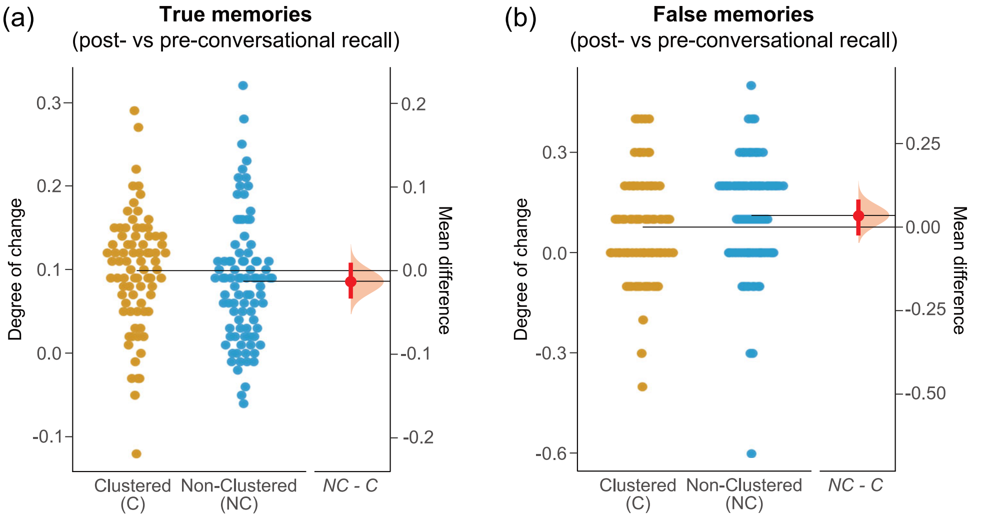

Fig. 2. Memory performance in pre- and post-conversational recall phases. Differences in recall rate of (a) true and (b) lure words (false memories) in post-conversational compared to pre-conversational recall phase. The mean difference between the clustered and the non-clustered group is shown in this Gardner–Altman estimation plot (Ho, Tumkaya, Aryal, Choi, & Claridge-Chang, 2019). Both groups are plotted on the left axes; the mean difference is plotted on a floating axis on the right as a bootstrap sampling distribution. The mean difference is depicted as a dot; the 95% confidence interval is indicated by the ends of the vertical error bar.

type × time × network condition effect ($F(1,168) = 3.93$, $p = .049$, $\eta^2 = 0.02$), indicating that the degree of pre-post conversational recall rate differed for true and false memories as a function of network condition (Fig. 2).

Separate repeated measures ANOVA for true and false memories showed a significant effect of time (true memories: $F(1,168) = 281.44$, $p < .01$, $\eta^2 = 0.63$; false memories: $F(1,168) = 55.83$, $p < .01$, $\eta^2 = 0.25$) but did not show a significant time × group interaction effect either for true ($F(1,168) = 1.34$, $p = .25$, $\eta^2 = 0.01$) or false memories ($F(1,168) = 1.93$, $p = .16$, $\eta^2 = 0.01$). However, upon closer examination of the participants' recall differences between pre- (phase II) and post-conversational (phase IV) recall phases, we detected outliers within the dataset (defined by those data points that exceeded above the third or below the first interquartile range of the data; Fig. 2). Consequently, we implemented a robust linear regression model to assess for differences between recall phases, as this analysis is less sensitive to outliers than ANOVA (Maechler et al., 2023). We found that true word recall increased to a greater extent in the post-conversational phase compared to pre-conversational recall for members in the clustered condition ($\beta = 0.021$, $SE = 0.011$, $t(168) = 1.93$, $p = .053$), whereas the extent of false memories during post-conversational recall increased more in the non-clustered condition than in the clustered condition ($\beta = 0.045$, $SE = 0.023$, $t(168) = 1.918$, $p = .056$). These results underscore the influence of network type on the recall rates of both true and false memories, with clustered networks enhancing true memory recall in the post-conversational phase and non-clustered networks promoting a greater increase in false memories.

To investigate whether these findings were specific to the studied words and the associated lures, we also analyzed memory intrusions of non-related words during the two recall phases (clustered group in phase II: $M = 7.42\%$, $STD = 8.19\%$ and in Phase IV: $M = 7.70\%$, $STD = 7.35\%$; non-clustered group in Phase II: $M = 7.87\%$, $STD = 10.62\%$ and in Phase IV: $M = 9.17\%$; $STD = 12.39\%$). A repeated measures ANOVA including phase (II vs. IV) as within-subjects factor and network condition (clustered vs. non-clustered) as a between-subjects factor, confirmed that memory intrusions for non-related words did not change in either network group before and after the conversational phase (main effect of phase: $F(1,168) = 1.27$, $p = .26$, $\eta^2 = 0.08$; phase $\times$ network condition interaction effect: $F(1,168) = 0.53$, $p = .47$, $\eta^2 = 0.03$). Altogether, these effects were in the hypothesized direction but only marginally significant; thus, we conducted additional analyses to explore the effect of network structure more precisely on true and false memories.

3.2. *Dynamics of memory performance during shared recall*

We next examined whether the repeated conversations modulated the memory accuracy observed at the individual level when comparing pre- and post-conversational recall phases in our previous analysis. To investigate this issue in our data, we first measured the proportion of correctly recalled words, and the proportion of lure words recalled in each participant's conversation iteration (first, second, and third; Table S3). We implemented a mixed factorial ANOVA that included conversation iteration (first, second, and third) and type of memory (true and false) as within-subjects factors and network condition (clustered and non-clustered) as a between-subjects factor to assess for statistical effects. This analysis confirmed that, in overall, participants tended to recall more true than false memories (main effect of type: $F(1,168) = 44.16$, $p < .001$, $\eta^2 = 0.21$) and that the two types of memories changed throughout conversation iteration (main effect of iteration: $F(2,336) = 6.99$, $p = .001$, $\eta^2 = 0.04$). However, we found that the degree of memory change throughout conversation iteration varied for true and false memories as a function of network condition, as indicated by a significant triple interaction type $\times$ iteration $\times$ network condition ($F(2,336) = 3.26$, $p = .04$, $\eta^2 = 0.02$). A separate ANOVA for true memories indicated that memory accuracy increased throughout the conversational phase iterations (main effect of iteration: $F(2,336) = 9.25$, $p < .001$, $\eta^2 = 0.05$), but the increase was similar between members of the two network conditions (network condition $\times$ iteration: $F(2,336) = 1.94$, $p = .15$, $\eta^2 = 0.01$; Fig. 3a). A polynomic contrast confirmed that the memory accuracy increase was linear ($F(1,168) = 14.54$, $p < .001$, $\eta^2 = 0.08$). Conversely, the same analysis on false memories indicated a more pronounced increase in non-clustered than in clustered conditions (main effect of iteration: $F(2,336) = 3.32$, $p = .04$, $\eta^2 = 0.02$; network condition $\times$ iteration: $F(2,336) = 2.83$, $p = .06$, $\eta^2 = 0.02$; Fig. 3b). A polynomic contrast confirmed that the interaction of the effects was linear ($F(1,168) = 5.71$, $p = .02$, $\eta^2 = 0.03$). Post hoc contrasts revealed that the rate of false memories was greater in the second ($t(89) = 2.99$, $p = .004$; Cohen's d = 0.31) and third ($t(89) = 2.98$, $p = .004$; Cohen's d = 0.34) recall iteration compared with the first recall iteration in the non-clustered

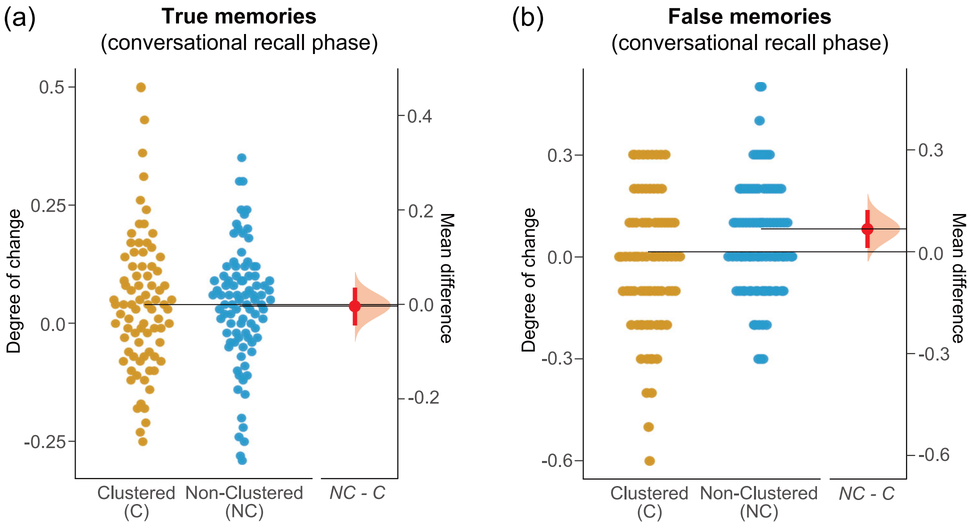

Fig. 3. Memory changes during the conversational Phase III. Differences in recall rate of (a) true and (b) lure words (false memories) between the first and the third conversational recall in Phase III. The mean difference between the clustered and the non-clustered group is shown in this Gardner–Altman estimation plot (Ho et al., 2019). Both groups are plotted on the left axes; the mean difference is plotted on a floating axis on the right as a bootstrap sampling distribution. The mean difference is depicted as a dot; the 95% confidence interval is indicated by the ends of the vertical error bar.

condition, whereas false memories did not differ significantly between recall iterations in the clustered condition (all $t(79) < 1.5$, $p > .1$). Similar trends were found when differences between first and third recall performance were analyzed by means of a robust linear regression model (true memories: $\beta = 0.013$, $SE = 0.019$, $t(168) = .65$, $p = .51$; false memories: $\beta = 0.055$, $SE = 0.029$, $t(168) = 1.91$, $p = .05$). These results support the hypothesis that the conversational network structure influences the emergence of false memories.

3.3. Memory consistency and network mnemonic convergence

Our findings suggest that network structure influences the way memories evolve within a group, producing opposing mnemonic effects depending on the structure of interactions. These effects emerge through an iterative process of shared recall, where memory representations are shaped by the dynamics of conversational exchange. More specifically, we hypothesize that a clustered network, characterized by localized and controlled information exchange, facilitates the preservation of accurate memories at the individual level. This occurs through the reinforcement of studied items in repeated recall conversations among closely connected members of the network. The controlled nature of these interactions ensures that recalled information remains relatively stable, strengthening individual memory traces and reducing the spread of activation that might otherwise lead to false recollections. In contrast, in non-clustered network conditions, recall conversations involve a more diffuse and less constrained flow of information, leading to greater divergence in the items recalled across conversational iterations. This increased variability in shared recall interactions coactivates less convergent memories among network members, making it more difficult to establish a consistent set of remembered items. As a consequence, the likelihood of false memories increases as activation spreads more freely across loosely connected concepts, facilitating the erroneous recollection of critical lures.

To empirically evaluate this hypothesis, we measured the degree of memory consistency across recall iterations. Specifically, we quantified the overlap in correctly recalled items between the first and second recall iterations, as well as between the second and the third iterations for each participant (Fig. 4a). To account for individual differences in initial recall performance, each of these consistency measures was normalized by the total number of correctly recalled items in the first recall stage. This approach allowed us to assess how memory stability evolved over time and whether the network structure influenced the retention and transformation of recalled information across conversational rounds. Confirming our hypothesis, the results of a repeated measures ANOVA including iteration (first/second and second/third) as a within-subjects factor and network condition as a between-subjects factor, revealed a significant iteration × network condition effect ($F(1,168) = 5.98$, $p = .01$, $\eta^2 = 0.034$; Fig. 4b). A post hoc analysis comparing memory consistency scores between the first/second and second/third iterations in the two network groups confirmed that memory preservation increased to a greater extent in the second/third iteration in the clustered network group, compared to the non-clustered group ($t(168) = 2.45$, $p = .02$, Cohen's d = 0.36). The current results indicate that the communication among members in social networks, whether characterized by clustering or lack thereof, plays a pivotal role in shaping the cognitive mechanisms underpinning memory processes.

These findings were specific to the studied words and the associated critical lures as both groups showed similar patterns of memory intrusions of non-related words during the conversational phase (clustered group: $M = 26.04\%$, $STD = 20.05\%$, $M = 28.49\%$, $STD = 20.23\%$ and $M = 25.41\%$, $STD = 17.85\%$, for first, second, and third iterations, respectively; non-clustered group: $M = 28.18\%$, $STD = 20.25\%$, $M = 26.72\%$, $STD = 17.37\%$ and $M = 25.14\%$, $STD = 18.89\%$). This was confirmed via a repeated measures ANOVA including iteration and network condition as within and between-subjects factors, respectively (main effect of iteration: $F(2,336) = 1.09$, $p = .39$, $\eta^2 = 0.005$; interaction iteration × network condition effect: $F(2,336) = 0.70$, $p = .49$, $\eta^2 = 0.004$).

The greater preservation of true memories among members of clustered networks across conversational iterations, compared to non-clustered networks, suggests two related consequences at the network level. First, memories in non-clustered networks are likely to become less similar over time than those shared in clustered networks. This effect should be most pronounced in the third conversation round, when interactions occur across participant clusters. To evaluate this possibility, we measured the similarity of the five memory recalls in each conversational iteration, corresponding to the five pairings of the 10 individuals from a network. This index was obtained by averaging the correlations between memory recall vectors for each pair of recalls within the network and across iterations. Since memory scoring was binary (word recalled or not recalled), we used the Phi coefficient, a measure designed for binary categorical data, to quantify recall similarity, where +1 indicates perfect positive

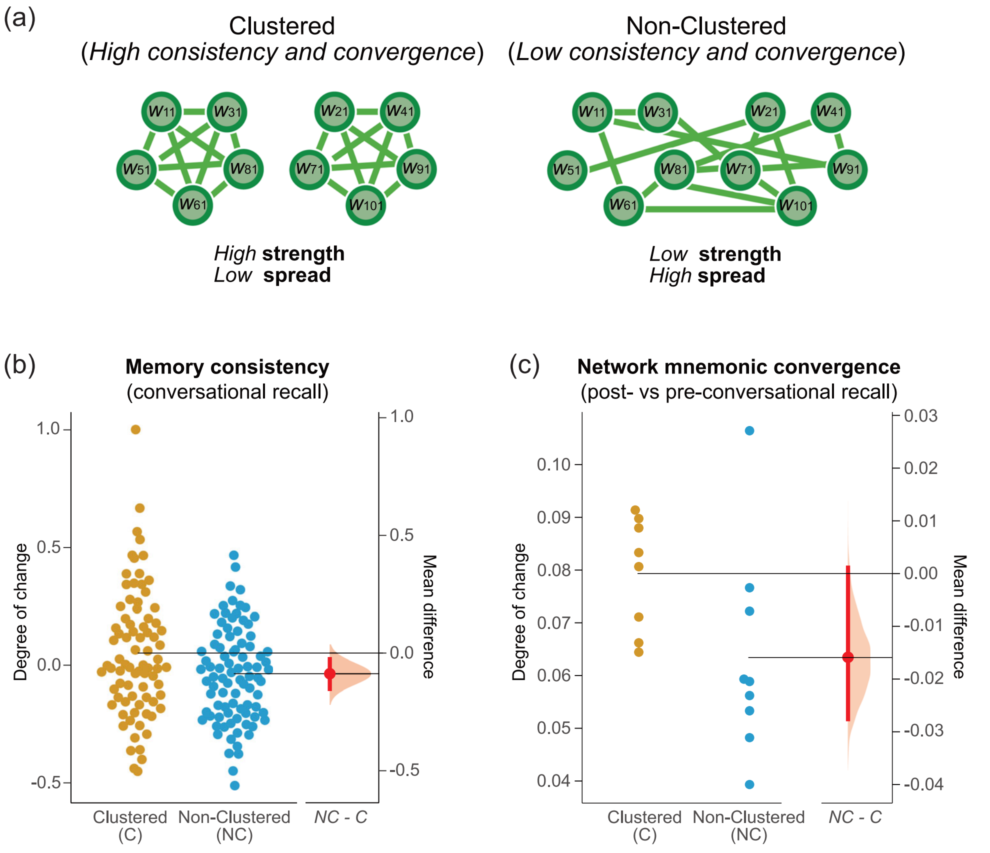

Fig. 4. (a) A graphic summary of the hypothesis of the study. We suggested that in clustered networks, where information exchange is localized, accurate memories are better preserved due to reinforced related memories within communities (left). In contrast, non-clustered networks with widespread information exchange may lead to related false memories. This is because interconnected nodes can coactivate less convergent conceptually related memories, similar to the cognitive dynamics promoting false memories (right). (b) Changes in the degree of memory consistency of true memories (words) between first/second and second/third conversational recall iteration (Phase III) for each member of the two network conditions. (c) Difference of mnemonic convergence scores for the clustered and non-clustered conditions in the post- compared to the pre-conversational recall phases. The mean difference between the clustered and the non-clustered group is shown in this Gardner–Altman estimation plot in (b) and (c) (Ho et al., 2019). Both groups are plotted on the left axes; the mean difference is plotted on a floating axis on the right as a bootstrap sampling distribution. The mean difference is depicted as a dot; the 95% confidence interval is indicated by the ends of the vertical error bar.

association and −1 indicates perfect negative association. The average similarity across all possible recall pairs within a network was then calculated for each iteration. However, this analysis revealed numerically consistent but statistically inconclusive results. More specifically, in the clustered condition, mean values were 0.20 ($SD = 0.06$) in round one, 0.12 ($SD = 0.06$) in round two, and 0.21 ($SD = 0.13$) in round three. In the non-clustered condition, mean values were 0.19 ($SD = 0.07$) in round one, 0.16 ($SD = 0.04$) in round two, and 0.14

($SD = 0.03$) in round three. Nevertheless, a repeated measures ANOVA showed a main effect of conversation round ($F(2,30) = 3.09$, $p = .06$, $\eta^2 = 0.17$), but a non-significant interaction conversation round × network condition ($F(2,30) = 2.20$, $p = .13$, $\eta^2 = 0.13$).

A second plausible consequence of the greater preservation of true memories across conversational rounds in clustered networks is that individuals in this condition tended to exhibit greater convergence at the network level after the conversational phase. To account for this possibility, we calculated a mnemonic similarity score for true and false memories for each pair of participants in the network by adding the number of items remembered in common by both participants and then dividing this sum by the total number of items recalled. Then, following Coman et al. (2016), a network mnemonic convergence score was calculated by averaging the mnemonic similarity scores among all the pairs of participants in the network, separately for the pre- and the post-conversational recalls. We found that, in overall, network mnemonic convergence was higher for true than for false memories (main effect of type: $F(1,15) = 484.54$, $p < .001$, $\eta^2 = 0.97$). We also found that the degree of mnemonic convergence increased after conversational recalls (main effect of recall phase: $F(1,15) = 331.34$, $p < .001$, $\eta^2 = 0.96$) and that this change differed between true than false memories (type × pre effect: $F(1,15) = 213.34$, $p < .01$, $\eta^2 = 0.93$). However, we also found a significant type × recall phase × network condition interaction effect ($F(1,15) = 4.71$, $p = .046$, $\eta^2 = 0.24$), indicating that the change between the pre- and the post-conversational mnemonic convergence for true and false memories differed between network conditions.

A separate repeated measures ANOVA for network mnemonic convergence for true and false memories allowed identifying the source of the triple interaction. More specifically, the ANOVA including recall phase (i.e., pre- and post-conversation) and network condition suggested that mnemonic convergence for true items increased more in the clustered than in the non-clustered network condition (main pre-post effect: $F(1,15) = 331.23$, $p < .01$, $\eta^2 = 0.96$; network condition × pre-post interaction effect: $F(1,15) = 4.12$, $p = .06$, $\eta^2 = 0.22$). Similar results were found when differences in mnemonic convergence between pre- and post-conversational recall phases were analyzed with a robust regression model that controlled for outliers in the data (true memories: $\beta = 0.02$, $SE = 0.01$, $t(168) = 2.35$, $p = .03$; Fig. 4c). Conversely, a similar increase was found in the clustered and the non-clustered groups regarding the mnemonic convergence for lure items (main pre-post effect: $F(1,15) = 10.42$, $p < .01$, $\eta^2 = 0.41$; network condition × pre-post interaction effect: $F(1,15) = 0.82$, $p = .38$, $\eta^2 = 0.05$). The results suggest that conversational recall fosters greater mnemonic convergence for true memories, especially in clustered networks, where shared recall strengthens collective memory consistency.

4. Discussion

In this study, we manipulated the topological structure of lab-created social community networks during a computer-mediated conversational recall task using a DRM word list paradigm to examine how network structure affects the formation of true and false memories.

We found that clustered networks promoted true memories by reinforcing memory convergence among community members, while non-clustered networks led to more false memories due to widespread cross-activation of non-overlapping memories. These findings provide empirical evidence that social network structure influences memory dynamics, highlighting the relationship between network topology and memory processes.

Our findings align with previous studies that showed that conversational remembering is selective (Marsh, 2007; Rajaram & Pereira-Pasarin, 2010), susceptible to errors (Schacter, 2022), capable of altering the memories of the interlocutors (Hirst & Echterhoff, 2012), and shaped by the degree of clustering network of the community structure (Coman et al., 2016). They also align with the notion that these effects can be explained by cognitive processes that take place during collaborative recall, such as memory reinforcement (Roediger, Zaromb, & Butler, 2009) and social contagion (Maswood & Rajaram, 2019; Meade & Roediger, 2009; Roediger, Meade, & Bergman, 2001). What distinguishes the current study from prior research is that it explored the potential that a general information transmission principle, based on the modularity of a network structure, could elucidate the nature of how memories are represented at the cognitive level. Specifically, collective memories are believed to arise from the exchange of information among engaged members of a community, with influence indirectly transmitted through connected peers (Yamashiro & Hirst, 2014). Cognitive models explaining memory representation suggest that a comparable process underlies the interconnection of memories sharing common content, facilitating the development of abstract and semantic-based representations (McClelland, McNaughton, & Lampinen, 2020). Our findings support this idea, showing that members of a clustered community tend to align their memories, strengthening and converging on shared items. In contrast, recall among members of a non-clustered community promotes memory divergence, increasing the likelihood of semantic-based false memories. These findings indicate that social network structures shape the nature of memory representations by inducing cognitive processes that occur at the individual level.

Our results highlight the dynamic nature of memory, where information is actively reconstructed and reorganized by the brain. This is consistent with the notion that memories are stored and retrieved through interconnected neural networks in the brain and false memories occur when these networks overlap during recall, distorting recollections (Kurkela & Dennis, 2016; Wing et al., 2020; Ye et al., 2016). However, while neuroimaging-supported empirical studies (e.g., Chadwick et al., 2016) provide valuable evidence supporting the role of mnemonic neural network activation in the brain, there are persisting limitations in our ability to mechanistically examine and precisely capture the correspondence of memory representations at the neuronal level. Our research contributes to this field by leveraging the analogy that these interconnected memory networks can be likened to a social network, where each node represents a community member engaged in dynamic interactions. This analogy provides a novel perspective, demonstrating empirically that memories are indeed significantly influenced by the cross-coactivation of associated nodes. In doing so, our study advances our understanding of the complex interplay between cognition and social networks, opening new avenues for research in this multidisciplinary field.

It is important to acknowledge that even though our study isolated the effects on true and false memories associated with network structure, these effects assumed that all individual members of the community and their impact at the network-wide level are of equal significance. However, not all individual members possess equal potential to influence the network's collective memory. For instance, individuals who connect between clusters have a significant influence on the network (Derex & Boyd, 2016), especially if their shared recall takes place at the early stages of other dyadic-level conversations within the network community (Momennejad, Duker, & Coman, 2019). Interpersonal factors, such as source credibility (partners vs. strangers; French, Garry, & Mori, 2008), perception of power (influential vs. weak; Skagerberg & Wright, 2009), and confidence (competitive vs. cooperative; Wright, Gabbert, Memon, & London, 2008; see for a review, Maswood & Rajaram, 2019) influences social contagion and likely the emergence of false memories.

Another important consideration in our study is that, unlike naturally occurring clustered networks, where close and trusting relationships may influence memory transmission, our clusters consisted of strangers interacting via computers. This lack of pre-existing social bonds could impact the likelihood of adopting false memories, as trust and familiarity may influence whether individuals accept or reject information introduced by others. Thus, systematic manipulations involving different temporal arrangements of shared conversations within members within the network topology and the inclusion of relationship characteristics of the members of the community will likely reveal meaningful network dynamics involving the formation of true and false collective memories.

While our study provides lure recall frequency, differences in recall structure must be considered when comparing our findings to classic DRM studies, where individuals study and recall independently. The turn-taking recall process introduces social dynamics that may influence false memory formation differently. A key distinction from the classic DRM paradigm is that, in our conversational setup, a lure memory can emerge either because a speaker spontaneously recalls it or because it is introduced by another participant and subsequently adopted. This interactive process contrasts with individual recall settings, where lure memories arise solely from an individual's internal memory reconstruction. The consequences of false memories as a function of whether a specific lure is reinforced by the repeated mentioning by others or by one's recall spontaneously remains unclear. As such, a direct comparison between our DRM-like setup and the classic DRM paradigm should be approached with caution. Future research should further explore how these interactive recall structures affect memory convergence, the transmission of errors, and the mechanisms that drive false memory propagation in social networks.

The impact of information transmission within social communities is an important topic of research as it reaches a large-scale societal impact, from attitudes and beliefs (Hirst & Echterhoff, 2012), with ramifications for political (Bakshy, Messing, & Adamic, 2015; Frenda, Knowles, Saletan, & Loftus, 2013) and health-related attitudes (Centola, 2010). One concerning aspect is the potential for these strategies to inadvertently foster false beliefs. Mitigating the emergence of false beliefs in society is challenging because attitude-congruent false events promote feelings of recognition and familiarity, which in turn interfere with source attributions (Johnson, Hashtroudi, & Lindsay, 1993). Our research suggests that an effective strategy may be to structure dissemination efforts based on social network design, controlling information spread within distinct communities rather than broadly through social networks or media. However, the relevance of this approach to real-world recall scenarios requires further investigation.

While false memories might seem like flaws in memory, many researchers argue they result from adaptive processes, reflecting gist-based thinking that aids generalization and abstraction (Gallo, 2010; Roediger & McDermott, 1995). False memory generation is linked to divergent thinking—generating multiple ideas—and convergent thinking—selecting the best ones—both essential to creativity (Thakral, Devitt, Brashier, & Schacter, 2021; Dewhurst, Thorley, Hammond, & Ormerod, 2011). Creativity, a key aspect of human development, is a major focus in education (Patston, Kaufman, Cropley, & Marrone, 2021), yet training creativity in schools remains challenging (DeHaan, 2011). Investigating the potential of novel educational strategies based on non-clustered information exchange in classrooms to promote creative thinking could be an intriguing research avenue.

Our findings underscore the intricate interplay between network topology, memory dynamics, and the construction of collective memories. This not only enriches our understanding of the cognitive processes underlying memory but also provides a lens through which we can examine the intricate relationships between social network structure, memory representation, and the evolution of collective knowledge.

Acknowledgments

We thank Andras Molnar for his generosity in modifying the SMARTRIQS software to meet the necessary specifications for conducting this study. This work was supported by the Spanish Ministerio de Ciencia, Innovación y Universidades, which is part of Agencia Estatal de Investigación (AEI), through the project PID2019-111199GB-I00 and PID2022-140426NB-I00 and PID2021-127728NB-I00. (Co-funded by the European Regional Development Fund. ERDF, a way to build Europe.) We thank the CERCA Programme/Generalitat de Catalunya for institutional support.

References

Alonso, M. Á., Fernández, Á., & Díez, E. (2004). Índices de producción de falso recuerdo y falso reconocimiento para 55 listas de palabras en castellano. *Psicothema*, *16*(3), 357–362

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261–295. https://doi.org/10.1016/S0022-5371(83)90201-3

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Barber, S. J., Rajaram, S., & Aron, A. (2010). When two is too many: Collaborative encoding impairs memory. *Memory & Cognition*, *38*(3), 255–264. https://doi.org/10.3758/MC.38.3.255

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*(5), 1176–1189. https://doi.org/10.1037/0278-7393.23.5.1176

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, *329*(5996), 1194–1197. https://doi.org/10.1126/science.1185231

Chadwick, M. J., Anjum, R. S., Kumaran, D., Schacter, D. L., Spiers, H. J., & Hassabis, D. (2016). Semantic representations in the temporal pole predict false memories. *Proceedings of the National Academy of Sciences*, *113*(36), 10180–10185. https://doi.org/10.1073/pnas.1610686113

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Coman, A., Manier, D., & Hirst, W. (2009). Forgetting the unforgettable through conversation: Socially shared retrieval-induced forgetting of September 11 memories. *Psychological Science*, *20*(5), 627–633. https://doi.org/10.1111/j.1467-9280.2009.02343.x

Coman, A., Momennejad, I., Drach, R. D., & Geana, A. (2016). Mnemonic convergence in social networks: The emergent properties of cognition at a collective level. *Proceedings of the National Academy of Sciences*, *113*(29), 8171–8176. https://doi.org/10.1073/pnas.1525569113

Congleton, A. R., & Rajaram, S. (2014). Collaboration changes both the content and the structure of memory: Building the architecture of shared representations. *Journal of Experimental Psychology: General*, *143*(4), 1570–1584. https://doi.org/10.1037/a0035974

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17–22. https://doi.org/10.1037/h0046671

DeHaan, R. L. (2011). Teaching creative science thinking. *Science*, *334*(6062), 1499–1500. https://doi.org/10.1126/science.1207918

Derex, M., & Boyd, R. (2016). Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences*, *113*(11), 2982–2987. https://doi.org/10.1073/pnas.1518798113

Dewhurst, S. A., Thorley, C., Hammond, E. R., & Ormerod, T. C. (2011). Convergent, but not divergent, thinking predicts susceptibility to associative memory illusions. *Personality and Individual Differences*, *51*(1), 73–76. https://doi.org/10.1016/j.paid.2011.03.018

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

French, L., Garry, M., & Mori, K. (2008). You say tomato? Collaborative remembering leads to more false memories for intimate couples than for strangers. *Memory*, *16*(3), 262–273. https://doi.org/10.1080/09658210701801491

Frenda, S. J., Knowles, E. D., Saletan, W., & Loftus, E. F. (2013). False memories of fabricated political events. *Journal of Experimental Social Psychology*, *49*(2), 280–286. https://doi.org/10.1016/j.jesp.2012.10.013

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, *38*(7), 833–848. https://doi.org/10.3758/MC.38.7.833

Griffiths, T. L., Lewandowsky, S., & Kalish, M. L. (2013). The effects of cultural transmission are modulated by the amount of information transmitted. *Cognitive Science*, *37*(5), 953–967. https://doi.org/10.1111/cogs.12045

Hirst, W., & Echterhoff, G. (2012). Remembering in conversations: The social sharing and reshaping of memories. *Annual Review of Psychology*, *63*(1), 55–79. https://doi.org/10.1146/annurev-psych-120710-100340

Hirst, W., Yamashiro, J. K., & Coman, A. (2018). Collective memory from a psychological perspective. *Trends in Cognitive Sciences*, *22*(5), 438–451. https://doi.org/10.1016/j.tics.2018.02.010

Ho, J., Tumkaya, T., Aryal, S., Choi, H., & Claridge-Chang, A. (2019). Moving beyond *P* values: Data analysis with estimation graphics. *Nature Methods*, *16*(7), 565–566. https://doi.org/10.1038/s41592-019-0470-3

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Kurkela, K. A., & Dennis, N. A. (2016). Event-related fMRI studies of false memory: An activation likelihood estimation meta-analysis. *Neuropsychologia*, *81*, 149–167. https://doi.org/10.1016/j.neuropsychologia.2015.12.006

Marsh, E. J. (2007). Retelling is not the same as recalling: Implications for memory. *Current Directions in Psychological Science*, *16*(1), 16–20. https://doi.org/10.1111/j.1467-8721.2007.00467.x

Maswood, R., & Rajaram, S. (2019). Social transmission of false memory in small groups and large networks. *Topics in Cognitive Science*, *11*(4), 687–709. https://doi.org/10.1111/tops.12348

McClelland, J. L., McNaughton, B. L., & Lampinen, A. K. (2020). Integration of new information in memory: New insights from a complementary learning systems perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1799), 20190637. https://doi.org/10.1098/rstb.2019.0637

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*(2), 159–188. https://doi.org/10.1037/0096-3445.114.2.159

Meade, M. L., & Roediger, H. L. (2009). Age differences in collaborative memory: The role of retrieval manipulations. *Memory & Cognition*, *37*(7), 962–975. https://doi.org/10.3758/MC.37.7.962

Maechler, M., Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M., Conceicao, E. L., & Anna di Palma, M. (2023). robustbase: Basic robust statistics. R package version 0.99-0. https://robustbase.r-forge.r-project.org/

Momennejad, I., Duker, A., & Coman, A. (2019). Bridge ties bind collective memories. *Nature Communications*, *10*(1), 1578. https://doi.org/10.1038/s41467-019-09452-y

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226–254. https://doi.org/10.1037/0096-3445.106.3.226

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Patston, T. J., Kaufman, J. C., Cropley, A. J., & Marrone, R. (2021). What is creativity in education? A qualitative study of international curricula. *Journal of Advanced Academics*, *32*(2), 207–230. https://doi.org/10.1177/1932202×20978356

Rajaram, S., & Pereira-Pasarin, L. P. (2010). Collaborative memory: Cognitive research and theory. *Perspectives on Psychological Science*, *5*(6), 649–663. https://doi.org/10.1177/1745691610388763

Robinson, K. J., & Roediger, H. L., III. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*(3), 231–237. https://doi.org/10.1111/j.1467-9280.1997.tb00417.x

Roediger, H. L., Abel, M., Umanath, S., Shaffer, R. A., Fairfield, B., Takahashi, M., & Wertsch, J. V. (2019). Competing national memories of World War II. *Proceedings of the National Academy of Sciences*, *116*(34), 16678–16686. https://doi.org/10.1073/pnas.1907992116

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*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Roediger, H. L., Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review*, *8*(2), 365–371. https://doi.org/10.3758/BF03196174

Roediger, H. L., Zaromb, F. M., & Butler, A. C. (2009). The role of repeated retrieval in shaping collective memory. In J. V. Wertsch & P. Boyer (Eds.), *Memory in mind and culture* (pp. 138–170). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511626999.009

Schacter, D. L. (2012). Constructive memory: Past and future. *Dialogues in Clinical Neuroscience*, *14*, 7–18. https://doi.org/10.31887/DCNS.2012.14.1/dschacter

Schacter, D. L. (2022). The seven sins of memory: An update. *Memory*, *30*(1), 37–42. https://doi.org/10.1080/09658211.2021.1873391

Skagerberg, E. M., & Wright, D. B. (2009). Sibling differentials in power and memory conformity. *Scandinavian Journal of Psychology*, *50*(2), 101–107. https://doi.org/10.1111/j.1467-9450.2008.00693.x

Thakral, P. P., Devitt, A. L., Brashier, N. M., & Schacter, D. L. (2021). Linking creativity and false memory: Common consequences of a flexible memory system. *Cognition*, *217*, 104905. https://doi.org/10.1016/j.cognition.2021.104905

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. https://doi.org/10.1080/09541440600872068

Vlasceanu, M., Morais, M. J., Duker, A., & Coman, A. (2020). The synchronization of collective beliefs: From dyadic interactions to network convergence. *Journal of Experimental Psychology: Applied*, *26*, 453–464. https://doi.org/10.1037/xap0000265

Wing, E. A., Geib, B. R., Wang, W.-C., Monge, Z., Davis, S. W., & Cabeza, R. (2020). Cortical overlap and cortical-hippocampal interactions predict subsequent true and false memory. *Journal of Neuroscience*, *40*(9), 1920–1930. https://doi.org/10.1523/JNEUROSCI.1766-19.2020

Wright, D. B., Gabbert, F., Memon, A., & London, K. (2008). Changing the criterion for memory conformity in free recall and recognition. *Memory*, *16*(2), 137–148. https://doi.org/10.1080/09658210701836174

Yamashiro, J. K., & Hirst, W. (2014). Mnemonic convergence in a social network: Collective memory and extended influence. *Journal of Applied Research in Memory and Cognition*, *3*(4), 272–279. https://doi.org/10.1016/j.jarmac.2014.08.001

Ye, Z., Zhu, B., Zhuang, L., Lu, Z., Chen, C., & Xue, G. (2016). Neural global pattern similarity underlies true and false memories. *Journal of Neuroscience*, *36*(25), 6792–6802. https://doi.org/10.1523/JNEUROSCI.0425-16.2016

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Lists of the words and the associated lure words included in the study

Table S2. Percentage of true and lure words recalled in the pre-conversational recall phase and in the post-conversational recall phase for each network group. Values indicate means and standard deviation in parenthesis.

Table S3. Percentage of true and lure words recalled in each conversational iteration in conversational recall phase for each network group. Values indicate means and standard deviation in parenthesis.