Ditta, A.S., & Steyvers, M. (2013). Collaborative Memory in a Serial Combination Procedure. *Memory, 21(6), 668-674.*

Collaborative Memory in a Serial Combination Procedure

Annie Stanfield Ditta

Mark Steyvers

Department of Cognitive Sciences

University of California, Irvine

Address for correspondence:
Mark Steyvers
University of California, Irvine
Department of Cognitive Sciences
2316 Social & Behavioral Sciences Gateway Building
Irvine, CA 92697-5100
E-mail: mark.steyvers@uci.edu Phone: (949) 824-7642 Fax: (949) 824-2307

Word count: 3572 (excluding abstract, references, tables and legends)

Abstract

This article describes a new approach for studying collaborative memory that examines people's editing processes for naturally occurring memory errors. In this approach, memories of individuals are combined via a chaining method in which each participant indirectly receives information from the previous participant. Participants were asked to individually study word lists and recall as many words as possible in an online setting. Once a participant completed the recall task, his/her answers were provided for the next participant as suggested answers for their own recall. However, that participant was allowed to add or subtract words from the provided list of suggested answers. The final answer of the group was an aggregate of recalled words based on the answer given by the last participant in the chain. Results showed that participants displayed a very high accuracy of recall throughout the chain, though they were not able to replicate the entire study list or eliminate all errors by the end of the chain. This procedure has the advantage that it allows examination of the memory-editing processes individuals utilize when they communicate information indirectly, independent from social factors that arise in face-to-face group memory settings.

In collaborative memory settings, people work together to retrieve information from memory. Many recent studies have investigated the cognitive and social effects of collaboration on memory (Roediger & McDermott, 2011; see Rajaram & Pereira-Pasarin, 2010 for an overview). In some cases, memory improves when information is recalled in groups. For example, Edwards and Middleton (1986) found that individuals in groups use information provided by others as cues for their own memories, thus adding to the overall number of items recalled. However, groups remember fewer items than when the same people recall separately and their recall is subsequently combined (Thorley & Dewhurst, 2007), a phenomenon called collaborative inhibition (Weldon & Bellinger, 1997). This could be because when people remember in groups, they are subject to social variables that may influence their performance (e.g. social loafing; Latané, Williams, & Harkins, 1979), or because of cognitive factors such as retrieval interference (Weldon, Blair, & Huebsch, 2000).

Another problematic effect of collaboration is the creation and persistence of false memories. A study by Gabbert, Memon, Allan, and Wright (2004) found that when people in groups encounter misinformation about an event, they are more likely to falsely recall this information later (when they would not otherwise have made these mistakes). In some situations, false memories are easily passed from one group member to another, leading to the “social contagion of memory” (Roediger, Meade, & Bergman 2001). For example, Roediger et. al. (2001) showed false memories implanted by a confederate are more likely to be accepted as truth (when compared to a control group with no false memory plant) when they are more consistent with the scene being studied (e.g. a toaster in a kitchen, as opposed to oven mitts) and when participants had a limited amount of time to study the scene.

It is important to note that typical memory studies of social contagion involve only two people: a confederate and a participant, or two subjects (see Gabbert, Memon, & Wright, 2006 and Meade & Roediger, 2002). Such studies that involve only two participants have traditionally shown that there are far fewer errors in collaborative groups when compared to nominal groups (Ross, Spencer, Blatz, & Restorick, 2008). However, larger groups might provide an advantage in recreating more accurate memories, as more points of view can be assimilated into the group memory. It is entirely possible that one participant catches a detail that several others do not, which would be missed in a group of only two people. Practical problems arise when studying group memory involving multiple individuals, especially when individuals can communicate freely among each other about their retrieved memory. This is because the large number of social and cognitive factors influencing group memory can make it difficult to isolate the factors of interest.

We propose a simple group-memory procedure in which multiple individuals communicate serially in a chain. This allows us to examine the dynamics of memory errors over multiple individuals and situations where errors occur naturally, without being planted by confederates. Our approach is inspired by the seminal research by Bartlett (1932) on the serial reproduction procedure. His study was structured much the way that the game telephone is played: one person sees the target stimulus (the “target message”) and reconstructs it from memory, and then the next person receives this reconstruction as the target stimulus, studies it in order to reconstruct it again, and so on. He found that the final reconstruction could be very different from the target message presented at the beginning of the chain, because each person introduced minor errors in each reconstruction, which then accumulated over each participant in the chain.

The purpose of Bartlett's study was only to examine how memories change over time, as opposed to how participants correct these errors. Ross *et al.* (2008) examined error correction in spouses, and found that they do engage in this sort of correction, but our procedure seeks to investigate these correction mechanisms in indirect collaboration and with larger group sizes. Our proposed method to study group memory does this by adopting some of the essential features of Bartlett's serial reproduction procedure, but with one key difference. In the original procedure, each individual only receives information from the previous individual in the chain (or the set of target stimuli in case of the first individual). In our variant of this procedure, each individual in the chain receives the target stimuli that need to be remembered but also receives the reconstruction from the previous individual in the chain. Therefore, each individual in the chain (excluding the first individual) has the ability to rely on their own memory of the original information, as well as the retrieved memories from one other individual. Because the "truth" individuals are trying to recall is shown to every individual in the chain, individuals are theoretically capable of remembering the entire truth, and correct any mistakes that arise. This is very unlikely to occur in Bartlett's serial reproduction procedure because only the first individual in the chain has access to the true answer. From here on, we will refer to our procedure as the *serial combination procedure*, to distinguish it from Bartlett's procedure.

In a way, the social contagion studies also examine editing memory in a chain, but they employ "single shot" chains of two in which there is only one chance to accept or reject an answer. In these studies, one person deliberately feeds another person incorrect material, a memory test follows, and the "chain" is terminated. However, our approach will allow us to examine the editing process over time with multiple people, as errors introduced by one person in a chain will be either copied or corrected by subsequent individuals in the chain. Our design also has the additional benefit in that the errors will arise naturally – there is no confederate slipping false answers to see how people will react. Finally, our procedure has the advantage that it can be implemented in non-social situations. Each individual can observe the reconstruction from another individual without being in the presence of that person by utilizing an online transmission of information. This allows us to study the dynamics of the editing process in absence of social factors.

Admittedly, our procedure stretches the definition of "collaboration," as the first individual does not have access to anyone else's answers, and there is no direct communication among the collaborators. Therefore, we are examining a more indirect type of collaboration as opposed to real-time face-to-face or online procedures (e.g. Ekeocha, & Brennan, 2008). Overall, the editing process in our serial combination procedure is related to the collaborative editing process in Wikipedia: any individual has the ability to delete and insert information in articles, and if something is blatantly wrong, someone else will delete the incorrect piece of information (e.g. Viégas, Wattenberg, & Dave, 2004). From time to time, people will add new bits of information, and if they are correct, they will remain in the article, while the additional incorrect information continues to be culled.

In this paper, we first describe an experiment where subjects study lists fashioned after the Deese-Roediger-McDermott paradigm and retrieve memories in the serial combination procedure. We then examine the types of editing processes that people utilize when correcting others' responses, and measure overall performance with regards to correct recall and precision of words recalled as a function of the position in the serial chain

Methods

Participants

A total of 25 participants were recruited for this study. All were UC Irvine undergraduate students recruited through the UCI Experimetrix subject pool. They were from all class ranks, with their ages ranging from 18-21. Females outnumber males 2:1 in the psychology department at UCI, and this disproportionality was reflected in our subject population. All participants were awarded 1.5 credits in their psychology courses for their participation (four half-hour online sessions over the course of two weeks). One participant was omitted for blatant cheating on the recall task. This participant's answers were almost an exact duplicate of the study list (impossible to achieve if the task were being attempted correctly).

Stimuli

The stimuli consisted of four separate study lists that were designed in the fashion of the Deese-Roediger-McDermott paradigm in order to produce false recall of a critical unpresented word, called the "critical lure" (Roediger and McDermott, 1995). For each study list, fifteen target words were selected with strong associative connections from the critical lure that itself was not part of the study list. For example, for one of the lists, the critical lure was “cow,” and words such as “steer,” “pasture,” “graze,” and “farm” were selected for the study list. In addition, each study list also contained 30 additional filler words unrelated to the lure. Overall, the study lists were composed of 30 unrelated target words and 15 related target words randomized together.

Procedure

Participants were presented with one study list at a time and were asked to watch a sequence of study words presented for 0.5 seconds each in the privacy of their own home. They then followed a link to an online form where they completed a free recall test that they self-terminated. The first participant was instructed to recall as many words as possible in no particular order. Every subsequent participant in the chain (the remaining seven), could see what the participant immediately preceding him or her had entered, and used this information to aid their own recall by adding or deleting words from the previous participant’s answers. For example, suppose all participants study the list containing words A, B, C, and D. The first participant correctly might recall A, B, C, but also word X, which was not on the original study list. The second participant would then see these answers, and perhaps agree with them, but add another correct word, D, to the answers. The third participant would then see A, B, C, D, and X, and would realize that X was not on the list, and should be deleted. He would then submit his answers as A, B, C, and D, and the group would have correctly reproduced the study list through the collaborative editing process. Participants were instructed that they should recall as many words from the study list as possible, and were free to accept the answers as given by the previous participant, though they were explicitly warned that some answers might be incorrect. Recall was concurrent with exposure to the previous participant's answers, and at no time were participants allowed to create their own list independent from the one provided by the previous participant.

The 24 participants (minus one cheater) were separated into three chains that were eight participants long, and were assigned to particular lists based on a Latin square design to ensure that no participant would repeatedly be exposed to the same previous participant's answers. All participants completed the task for each of the four lists, though they were in a different position of the chain of eight for each list.

Performance on the recall task was measured by counting the number of words produced that were presented in the study list. A word was counted as correct if the stem was correct (e.g. "Sailed" was counted as correct even if the word presented in the study list was "Sailing"). However, when presenting the recalled words to the subsequent participant during the course of the experiment, the exact word that the participant recalled was used.

Results and Discussion

All analyses are presented as a function of participants' position in the chain (1-8). Results for each position are averaged across the three chains of individuals that participated in the experiment as well as the four lists. At all steps in the chain except for the first, participants could perform four different actions: correct insertion, incorrect insertion, correct deletion, and deletion of a correct item. A correct insertion was defined as a participant correctly recalling a word that was presented in the study list. An incorrect insertion was defined as the addition of any other word aside from those presented in the study list. A correct deletion occurred when participants took out a word that was incorrectly inserted by a participant before them, and a deletion of a correct item was when participants removed a correctly recalled word added by a participant before them. A participant in the first position could perform only correct or incorrect insertions.

(Figure 1 about here)

Insertions and deletions. There were more correct than incorrect insertions, $F(1,15)=23.84$, $MSE = 60.565$, $p < .001$. As can be seen in Figure 1, panel A, correct insertions were very high throughout the chain, with relatively few incorrect insertions. Position 1 has the highest rate of correct insertion due to its unique position in the chain; any words recalled by participants in this position are counted as insertions. Subsequent levels of correct insertion are much lower, though the rates are still higher than incorrect insertions. This suggests that if any words were added by participants (relative to the previous participant), they tended to be correct responses. However, the pattern for deletions (panel B) shows the opposite result. While deletions were much less frequent than insertions overall, the number of deletion of incorrect items was higher than the number of correct deletions, $F(1,15)=16.28$, $MSE = 2.68$, $p < .01$. In other words, participants more often removed words corresponding to the correct responses (e.g. words on the study list) from the previous participant than incorrect responses (e.g. the critical lure and other intrusions). As a result, participants tend not to remove the critical lure once it is introduced by previous participants in the chain and the probability of falsely recalling the critical lure increases as the chain progresses (panel C), leading to greater false recall of the critical lure in the second half of the sequence compared to the first half, $F(1,6)=9.4$, $MSE=.0319$, $p<.02$. Therefore, participants generally accept the lures as correct recalls on the part of the previous participant, and mostly refrain from deleting them.

These findings are consistent with previous research by Vollrath, Sheppard, Hinsz, and Davis (1989), which found that in collaborative groups there is both extensive error-checking and exaggeration of commonly recalled items. While editing a set of answers, participants can fall prey to conformity when they accept the “common” lure words (common because it is related to so many other words, and incorrectly recalled by so many). We find that error-checking is a hindrance when it comes to editing correct recall because words that should be left in the list are deleted by participants more often than words that should be deleted. A possible explanation for this is that participants are over-checking each other with a propensity to delete. However, this checking mechanism falls short when it comes to identifying the critical lures, possibly due to the fact that the critical lures are highly related to the study list; as a result, they pass the checking mechanism undetected.

Overall performance. We analyzed the overall performance on the basis of the total number of correct and incorrect words recalled. Figure 1, panel D, shows the total number of words recalled as a function of the participants’ position in the chain. The average number of words recalled increased over chain position and reached an average of 24.8 words (out of the total 45 words that could be recalled). This indicates that participants recalled a little over half of the list (55%) by the time the chain was terminated, though there was no indication of the trend leveling off. The total number of incorrect words recalled remained low throughout the chaining procedure: there were no more than 2 words incorrectly recalled at any position in the chain. Therefore, even though the trend in panel C shows that the critical lure is more likely to be recalled as a result of chaining, this particular item contributes to at most one error per list. The results in panel D show that the total number of errors (including the critical lure as well as other intrusions) remains low overall, indicating very high precision rates for each position of the chain.

This pattern of results is related to the findings by Basden, Basden, Bryner, and Thomas (1997), which revealed that when a participant is presented with part of a study list during recall, fewer total words are recalled than when the participant is not presented with any subset of the list (the part-set cueing impairment). Correspondingly, when participants 2-8 in this study were presented with a previous set of recalled words, this may have inhibited recall of new words by focusing participants on altering the presented words. However, the data show that participants continued to add new words to the list, even at the later stages of the chain, so it is possible that if the chains were longer, more words would continue to be added.

The high rates of precision that the subjects produced are interesting to consider in this paradigm. The error-checking mechanism (Vollrath *et. al.*, 1989) discussed earlier could be playing a role here. Participants might use this mechanism to make sure that they only add words they are confident were on the study list, and delete any words that could potentially be wrong (even if they were correctly recalled by the previous participant). This leads to a bias in strategy in which fewer words are being added at every step, but there is a high likelihood that they are correct. This ultimately results in higher precision rates, which reveals that participant errors seem to come from over-deletion, and not from incorrect insertion.

General Discussion

Overall, participants took into account the responses from other participants and produced responses based not only on what they themselves remembered, but also based on answers provided by the previous participant. Participants were very good at both recalling words that were presented in the study list and in refraining from adding intrusions, except for the single critical lure item, which is likely to be recalled. Even so, overall correct recall only reached approximately half of the study list by the end of the experiment. This is possibly due to the nature of the task, as free recall studies rely heavily on subjective encoding and retrieval methods that are subject to disruption when presented with cues that do not correlate with the particular retrieval strategy being utilized (Rajaram & Pereira-Pasarin, 2010).

With these results in mind, it is of interest to consider the study by Gabbert *et. al.* (2006), which found that participants who add to a group memory after the initial participant has recalled information are less confident in their answers when they conflict with what that initial participant produced. It is possible that participants in our study were more likely to question their own judgments if they did not match what the previous participant had submitted. Knowing that they would be combining someone else's memories with their own recall might make them less likely to add new words to or change existing words in the group list. Similarly, if participants were aware of the fact that their information was going to be passed on to another participant, social variables might be introduced into the task. Our procedure never explicitly stated that this would occur, but if participants inferred that others would be seeing their answers, social confounds might have been introduced into the study (e.g. participants might try harder than they would normally to “impress” the future participant even though direct interaction would never occur).

In order to address this, it might be interesting to conduct a study in which participants are specifically told that their answers will not be passed on to others when they actually are. Similarly, it would be interesting to make participants aware of their position in the chain, to see if this induces any differences in performance. Individual differences in editing processes at each particular step in the chain could then be analyzed. Additionally, because the first participant never actually collaborates with anyone, it would be useful to conduct a study in which this person is reinserted into the chain, to see if his/her memory is subsequently affected by previous participants’ answers.

Another variation on the current procedure would be to use a recognition memory task, as opposed to a recall task, to see if there are any differences in performance. Including a confidence rating for each word as the participant enters it (or selects it, as in the case of a recognition task) would help to identify how exactly people are using the information presented to them. In the current task, a word can only be added or subtracted – the participant cannot say how confident he or she was in that decision to add or subtract a word. In conclusion, this study demonstrates the idea that group memory benefits from even indirect communication among participants, but also reinforces findings by McDermott (1996) on the extreme robustness of false memory. We have outlined an alternative technique for examining recall memory that allows us to investigate the dynamic group editing process one step at a time, and to analyze errors as they occur during indirect collaboration.. This procedure could potentially have applicability for memory studies involving numerous types of stimuli, and could set the stage for additional techniques involving reinsertion of the target message (the “truth”) for all participants. However, it remains to be seen if improvements can be made in the reduction of false memory in these serial combination procedures.

References

Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology*. Cambridge University Press.

Basden, B.H., Basden, D.R., Bryner, S., & Thomas III, R.L. (1997). A comparison of group and individual remembering: Does collaboration disrupt retrieval strategies? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1176-1189.

Edwards, D., & Middleton, D. (1986). Joint remembering: Constructing an account of shared experience through conversational discourse. *Discourse Processes*, *9*(4), 423-459.

Ekeocha, J. O., & Brennan, S. E. (2008). Collaborative recall in face-to-face and electronic groups. *Memory*, *16*(3), 245-261.

Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps toward influence during a discussion. *Psychonomic Bulletin & Review*, *13*(3), 480-485.

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.

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37, 822–832.

Meade, M. L., & Roediger, H. L., III. (2002). Explorations in the social contagion of memory. *Memory & Cognition*, *30*(7), 995-1009.

McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory and Language, 35*, 212-230.

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

Roediger, H.L., III, & McDermott, K.B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

Roediger, H. L., III, & McDermott, K. B. (2011). Remember when? *Science, 333*, 47-48.

Roediger, H. L., III., Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review, 8*(2), 365-371.

Ross, M., Spencer, S. J., Blatz, C. W., & Restorick, E. (2008). Collaboration reduces the frequency of false memories in older and younger adults. *Psychology and Aging, 23*(1), 85-92.

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, 867–881.

Viégas, F.B., Wattenberg, M., Dave, K. (2004). Studying Cooperation and Conflict between Authors with history flow Visualizations. *Proceedings of the SIGCHI conference on Human factors in computing systems*, 575–582.

Vollrath, D. A., Sheppard, B. H., Hinsz, V. B., & Davis, J. H. (1989). Memoryperformance by decision-making groups and individuals. *Organizational Behavior and Human Decision Processes, 43*, 289-300.

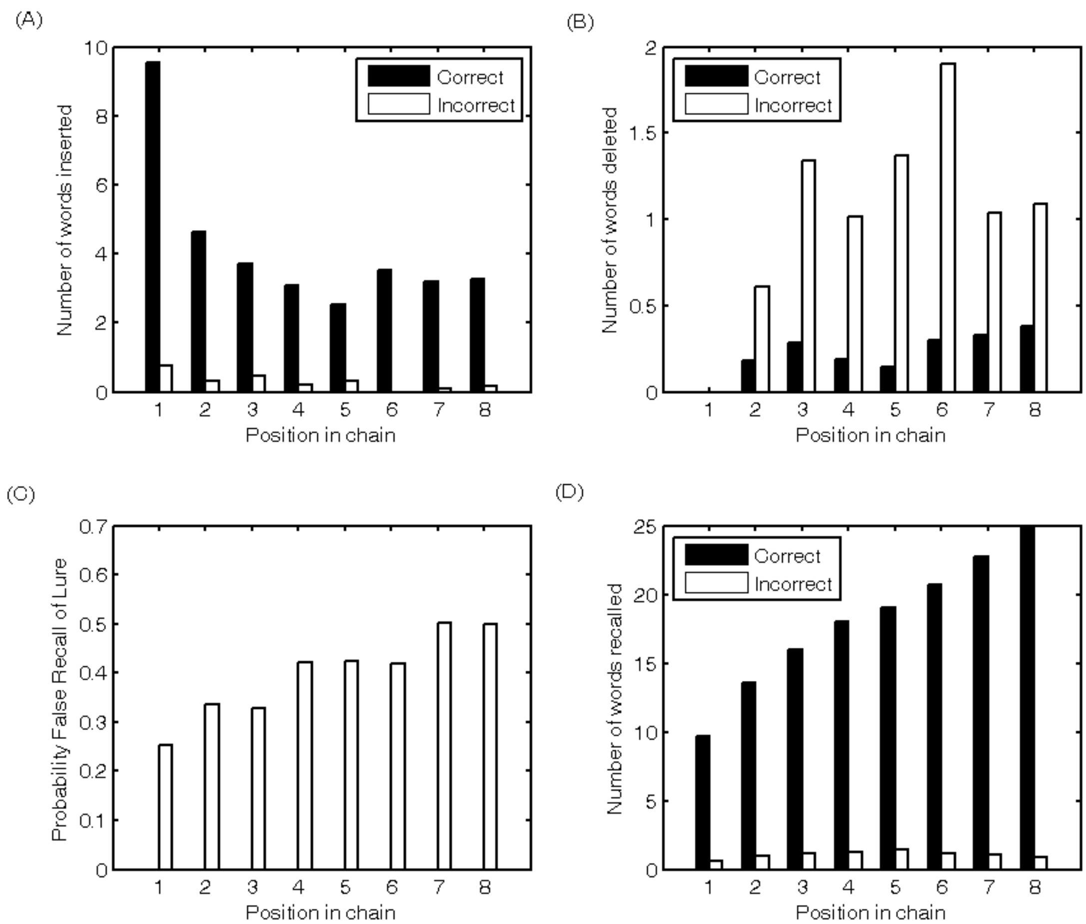
Weldon, M.S., & Bellinger, K.D. (1997). Collective memory: Collaborative and individual processes in remembering. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, 1160–1175.

Weldon, M. S., Blair, C., & Huebsch, P. D. (2000). Group remembering: Does social loafing underlie collaborative inhibition? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1568-1577.

Captions

Figure 1. Results of the experiment as a function of the participants' positions in the chain. (A) The number of correct and incorrect words inserted. (B) The number of correct deletions and deletions of incorrect items. (C) The probability of incorrectly recalling the critical lure. (D) The total number of correct and incorrect words recalled.

[figure 1]


(A)
Number of words inserted
Correct
Incorrect
Position in chain
(B)
Number of words deleted
Correct
Incorrect
Position in chain
(C)
Probability False Recall of Lure
Position in chain
(D)
Number of words recalled
Correct
Incorrect
Position in chain