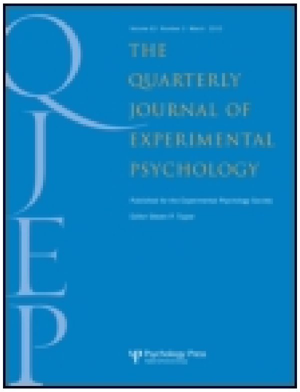

The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Consensus collaboration enhances group and individual recall accuracy

Celia B. Harris [a], Amanda J. Barnier [b] & John Sutton [b]

[a] Center on Autobiographical Memory Research (CON AMORE), Aarhus University, Aarhus, Denmark

[b] Macquarie Centre for Cognitive Science (MACCS), Macquarie University, Sydney, Australia

Available online: 04 Aug 2011

To cite this article: Celia B. Harris, Amanda J. Barnier & John Sutton (2012): Consensus collaboration enhances group and individual recall accuracy, The Quarterly Journal of Experimental Psychology, 65:1, 179-194

To link to this article: http://dx.doi.org/10.1080/17470218.2011.608590

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2012, 65 (1), 179–194

Psychology Press
Taylor & Francis Group

Consensus collaboration enhances group and individual recall accuracy

Celia B. Harris[1], Amanda J. Barnier[2], and John Sutton[2]

[1]Center on Autobiographical Memory Research (CON AMORE), Aarhus University, Aarhus, Denmark
[2]Macquarie Centre for Cognitive Science (MACCS), Macquarie University, Sydney, Australia

We often remember in groups, yet research on collaborative recall finds "collaborative inhibition": Recalling with others has costs compared to recalling alone. In related paradigms, remembering with others introduces errors into recall. We compared costs and benefits of two collaboration procedures—turn taking and consensus. First, 135 individuals learned a word list and recalled it alone (Recall 1). Then, 45 participants in three-member groups took turns to recall, 45 participants in three-member groups reached a consensus, and 45 participants recalled alone but were analysed as three-member nominal groups (Recall 2). Finally, all participants recalled alone (Recall 3). Both turn-taking and consensus groups demonstrated the usual pattern of costs during collaboration and benefits after collaboration in terms of recall completeness. However, consensus groups, and not turn-taking groups, demonstrated clear benefits in terms of recall accuracy, both during and after collaboration. Consensus groups engaged in beneficial group source-monitoring processes. Our findings challenge assumptions about the negative consequences of social remembering.

Keywords: Collaborative recall; Collaborative inhibition; Memory accuracy; Social influences on memory; False memory.

In a tradition dating back at least to Bartlett (1932), social factors have been recognized as important determinants of what and how we remember (Alea & Bluck, 2003; Barnier, Sutton, Harris, & Wilson, 2008; Weldon, 2001). Couples, families, and social groups collaborate in remembering (or forgetting) significant shared and unshared experiences (Harris, Barnier, Sutton, & Keil, 2010; Harris, Keil, Sutton, Barnier, & McIlwain, 2011; Harris, Sutton, & Barnier, 2010; Tollefsen, 2006). The fact that we often recall the past in conversation with others

Correspondence should be addressed to Celia B. Harris, Department of Psychology, Aarhus University, Jens Chr. Skous Vej 4, 8000 Aarhus C, Denmark. E-mail: celia@psy.au.dk

The research we report is part of a larger long-term project on individual and group memory, and the research and the preparation of this manuscript were supported by: (a) an Australian Research Council PhD scholarship to Celia Harris; (b) an Australian Research Council Australian Research Fellowship to Amanda Barnier; (c) an Australian Research Council Discovery-Project Grant (DP0770271) to Amanda Barnier and John Sutton, and (d) the Danish National Research Foundation. We are grateful for that support. This research was conducted at Macquarie University as part of the requirements for Celia Harris's PhD, with the generous support of the Macquarie Centre for Cognitive Science. We are also most grateful to Paul Keil for research assistance and to Suparna Rajaram, Dan Wright, Craig Thorley, and an anonymous reviewer for their helpful comments on a previous draft of this manuscript.

http://www.psypress.com/qjep
http://dx.doi.org/10.1080/17470218.2011.608590

may have consequences for memory both during and after such conversations.

The collaborative recall paradigm

The collaborative recall paradigm (Basden, Basden, Bryner, & Thomas, 1997; Blumen & Rajaram, 2008; Finlay, Hitch, & Meudell, 2000; Rajaram & Pereira-Pasarin, 2010; Weldon & Bellinger, 1997) was designed to assess the "costs and benefits" of remembering in a group (Basden, Basden, & Henry, 2000; for review, see Harris, Paterson, & Kemp, 2008). In collaborative recall, the impact of recalling with others is indexed during collaboration by comparing the output of collaborative groups (a group of three people working together) with the output of nominal groups (the pooled, nonredundant output of three people working separately). This nominal group comparison measures the potential output of the group if collaboration is simply aggregation of individual recall. The impact of recalling with others is indexed after collaboration by comparing the output of individuals who have previously collaborated and those who have not. This comparison measures the ongoing effects of collaboration.

Recall completeness

The costs and benefits of collaboration can be measured in terms of recall completeness: the percentage of the study material accurately recalled. Collaborative recall experiments have repeatedly demonstrated that collaborative groups recall less than nominal groups, an effect termed "collaborative inhibition" (Basden et al., 1997; Basden et al., 2000; Weldon & Bellinger, 1997). The best supported explanation for collaborative inhibition is the "retrieval strategy disruption hypothesis": Exposure to the responses of others interferes with the idiosyncratic organization that each person adopts for their own recall, making each person in the group recall less efficiently (Basden et al., 1997; Rajaram & Pereira-Pasarin, 2010). This is similar to the part-set cueing effect (Andersson, Hitch, & Meudell, 2006), where exposure to some items from a list inhibits recall for the remaining items (see Roediger, 1973). Evidence for this account comes from research showing that collaborative inhibition is abolished when each group member is responsible for recalling a different part of a categorized list (Basden et al., 1997), when recall is cued (Finlay et al., 2000), when group members are forced to organize their recall by category (Basden et al., 1997), and when group members are unable to hear or see the items recalled by other group members (Wright & Klumpp, 2004).

Although social loafing does not explain collaborative inhibition (Weldon, Blair, & Huebsch, 2000), other social and motivational factors—such as whether interaction is face to face or electronic and the perceived output level and pressure in the group—do influence the amount remembered (Ekeocha & Brennan, 2008; Reysen, 2003, Thorley & Dewhurst, 2007). Moreover, recent research suggests that collaboration between expert pilots, who are skilled at communicating in order to perform tasks together, results in facilitation rather than inhibition (Meade, Nokes, & Morrow, 2009). This recent finding is the first experimental demonstration that collaborative recall can exceed nominal recall under certain conditions.

Collaborative recall experiments have also demonstrated that individual memory continues to be influenced by collaboration: people who previously collaborated recall more than people who previously remembered alone (Basden et al., 2000, Blumen & Rajaram, 2008; Rajaram & Pereira-Pasarin, 2010). Extra items tend to be those that were introduced during collaboration; participants who previously collaborated are less likely to remember new items than participants who remembered alone ("inhibition of reminiscence", Basden et al., 2000). This suggests that after collaboration, individual recall is more similar.

Recall accuracy

The costs and benefits of collaboration can also be measured in terms of recall accuracy, or items recalled in error. Notably, in related social memory paradigms such as social contagion (Meade & Roediger, 2002; Roediger, Meade, &

Bergman, 2001) and memory conformity (Gabbert, Memon, & Allan, 2003; Gabbert, Memon, & Wright, 2006), research has emphasized the power of remembering with others to distort recall and "implant" false memories that persist in subsequent individual recall. For example, summarizing 20 years of false-memory work, Loftus wrote: "misinformation has the potential for invading our memories when we talk to other people . . ." (Loftus, 1997, p. 51).

A number of collaborative recall experiments have utilized the study lists of the Deese–Roediger–McDermott (DRM) paradigm (Roediger & McDermott, 1995), or similarly constructed categorized lists, to maximize false recall. These lists contain semantically related words, with the most common exemplars from the category removed. These missing exemplars become "critical lures" (Roediger & McDermott, 1995), and false recall of both critical lures and of other items mentioned in error during collaboration (intrusions) can be measured, although both measures are not always reported. Using DRM lists, Basden, Basden, Thomas, and Souphasith (1998, Experiment 1) reported that false recall of critical lures was similar in collaborative groups than in nominal groups, while overall intrusion rates were higher for collaborative than for nominal groups. Using similarly constructed categorized lists with high-frequency exemplars removed, Basden et al. (1998, Experiment 2) reported that false recall of critical lures was higher for collaborative groups than for nominal groups. In contrast, using DRM lists, Maki, Weigold, and Arello (2008, Experiment 1) found that collaborative groups falsely recalled fewer critical lures than nominal groups and concluded that "the results from the social contagion paradigm do not generalize to collaborating groups" (p. 600). Using similarly constructed lists, Takahashi (2007) also reported (in two of three experiments) that collaborative groups falsely recalled fewer critical lures than nominal groups.

Using nonspecialized materials (without critical lures), Basden et al. (1997) found that collaborative groups produced more intrusions than nominal groups. In contrast, Weldon and Bellinger (1997) reported no difference between the intrusion rates of collaborative and nominal groups, for both word lists and stories. And in two studies, Ross and colleagues (Ross, Spencer, Blatz, & Restorick, 2008; Ross, Spencer, Linardatos, Lam, & Perunovic, 2004) reported that collaborative groups produced fewer intrusions than nominal groups on recall and recognition tasks. Overall, a review of these studies indicates that, regardless of the stimuli, results regarding the accuracy of collaborative groups have been mixed.

In the collaborative recall literature, most studies have not indexed intrusions in subsequent individual recall, so it remains unclear whether collaboration has ongoing effects on recall accuracy and how the findings from the collaborative recall paradigm fit with findings from other paradigms like social contagion and memory conformity. In one exception, using DRM lists, Thorley and Dewhurst (2007) found that two different types of collaborative recall resulted in a reduction in intrusions from collaborative to subsequent individual recall. For critical lures, the effects depended on collaboration type: turn-taking collaboration resulted in a carry-over of critical lures from collaboration to subsequent individual recall, while free-for-all collaboration did not. However, Thorley and Dewhurst did not compare former members of collaborative and nominal groups. We turn now to further discuss this issue of collaboration type.

Group interaction

In previous collaborative recall experiments, the nature of the collaborative groups' interaction has varied, with some experiments adopting a turn-taking procedure in which group members recall a single item in turn (e.g., Basden et al., 1997, 2000), others a free-for-all procedure, in which groups collaborate with no particular instruction about how to proceed (e.g., Weldon & Bellinger, 1997), and, more rarely, a consensus procedure, in which group members agree on each item recalled (e.g., Meudell, Hitch, & Boyle, 1995; Ross et al., 2004). These procedural variations may influence both recall completeness and recall accuracy. In terms of recall completeness, differences in the nature of the group's interaction may influence the disruption of individual recall strategies. Turn-taking collaboration would be expected to maximize disruption of individual recall strategies compared to other methods of collaboration because it forces individuals to only output one item at a time, meaning that they cannot even partially maintain their own recall order. Also, it might be expected that less constrained collaboration like free-for-all would allow the group to use recall strategies to coordinate, cue, and increase collaborative recall (as in Meade et al., 2009, although there is little evidence that nonspecialized groups can cross-cue each other during collaboration, Meudell et al., 1995). Despite these possible differences, Thorley and Dewhurst (2007) found similar levels of collaborative inhibition for turn-taking and free-for-all collaborative groups, suggesting that group interaction does not influence recall completeness during collaboration. After collaboration, Thorley and Dewhurst argued that turn-taking collaboration increased subsequent individual recall, while free-for-all collaboration did not. However, they did not compare the amount recalled by former members of collaborative groups and nominal groups, so the effect of group interaction on subsequent individual recall remains unclear.

In terms of recall accuracy, differences in the nature of the groups' interaction may influence intrusions in two ways. First, Ross et al. (2008) reported that younger and older spouses engaged in error checking during collaboration, which eliminated errors from the group output. They argued that "partners played an important role in helping rememberers discard false memories" (Ross et al., 2008, p. 85). A more constrained procedure like turn taking (compared to free-for-all) does not allow for such group error checking. Second, Ross et al. (2008) argued that collaboration inhibited the production of errors by influencing individuals' response criteria, meaning that they required more certainty before stating an item during collaboration, which would decrease both correct recall and intrusions (see also Ross et al., 2004). Thorley and Dewhurst (2007) suggested that different types of collaboration might influence individuals' response criteria. They found that turn-taking collaboration produced more errors in recall than did free-for-all collaboration. This finding is consistent with the results reviewed earlier, where experiments that use a turn-taking procedure tend to find increased errors in collaborative groups (e.g., Basden et al., 1998), and experiments that use a free-for-all procedure tend to find no difference (e.g., Weldon & Bellinger, 1997) or fewer errors for collaborative groups (e.g., Ross et al., 2008). Thorley and Dewhurst argued that this difference between turn taking and free-for-all was due to differences in group pressure influencing response criterion (Thorley & Dewhurst, 2007; see also Basden et al., 1998), but this was not directly manipulated, and error checking as an alternative explanation cannot be discounted (since free-for-all groups can discuss and correct errors, while turn-taking groups cannot). Additionally, as mentioned above, the effect of group interaction during collaboration on subsequent individual recall accuracy remains unclear.

The current study

Previous collaborative recall experiments have yielded some robust findings and have elucidated processes that influence the recall of people remembering in groups (for review see Harris et al., 2008; Rajaram & Pereira-Pasarin, 2010). However, it remains unclear what the effects of different group interactions may be on both recall completeness and recall accuracy. In the current experiment, we compared costs and benefits for recall completeness and recall accuracy of turn-taking versus consensus collaboration, both during and after collaboration. We focused on both response criterion and error checking in these groups.

We chose to test turn-taking and consensus instructions rather than free-for-all because a free-for-all instruction offers less control over the collaboration. In pilot work we noticed that some groups given free-for-all instructions spontaneously adopted a turn-taking strategy, others adopted a consensus strategy, and others adopted idiosyncratic strategies, which made valid comparison of different collaboration styles difficult (see also varying results for accuracy in free-for-all groups reported by Takahashi, 2007; Weldon & Bellinger, 1997). We used a consensus instruction to ensure that members of these groups genuinely engaged with each item recalled during group discussion.

In terms of recall completeness, during collaboration we expected that both consensus groups and turn-taking groups would show collaborative inhibition, since previous research has found collaborative inhibition across different kinds of collaborative groups. We were interested in whether these costs might be greater for turn-taking groups where individuals experience more strategy disruption during collaboration. However, the findings of Thorley and Dewhurst (2007) suggested that there may not be a difference, since they found no difference in magnitude of collaborative inhibition between turn-taking and free-for-all groups. After collaboration, we expected that prior members of both turn-taking and consensus collaborative groups would recall more than prior members of nominal groups, as they would have had the opportunity to gain items from each other during collaboration (Blumen & Rajaram, 2008). However, we expected that these benefits might be greater for consensus groups if individuals had experienced less retrieval strategy disruption during collaboration. In terms of recall accuracy, during collaboration we expected that consensus groups would be more accurate—producing fewer intrusions—than turn-taking groups (as Thorley & Dewhurst, 2007, found for free-for-all groups). After collaboration, we expected that consensus groups might continue to show benefits in terms of recall accuracy—producing fewer intrusions than turn-taking groups—since they may continue to benefit from reduction in errors provided by the group during collaboration. We were also interested in whether reported recall strategies and indices of response criterion across the different types of collaborating groups might explain differences in their performance.

Method

Participants

We tested 135 undergraduate psychology students (91 women, 44 men) at Macquarie University, Sydney, Australia in a 3 (condition: turn taking vs. consensus vs. nominal) × 3 (recall occasion: 1 vs. 2 vs. 3) mixed-model design. They participated in return for course credit or payment of AU $15 per hour. Of the 135 participants, we tested 45 as three-member turn-taking groups, 45 as three-member consensus groups, and 45 as individuals (later analysed as three-member nominal groups). We always tested participants in either of the two collaborative recall conditions three at a time. We tested participants in the nominal group condition three, two, or one at a time. For the nominal groups, as participants arrived, we assigned them to the next open nominal group and tested them in the appropriate counterbalancing condition until the group was full. We approximately balanced the gender mix across the three conditions; of the 45 groups, 16 were made up of three women, 2 were made up of three men, and the remainder were mixed-gender groups.

Materials

We presented participants with a list of 45 words sequentially on a computer screen. Words appeared in 26-point bold black Arial font in the centre of a white screen, one at a time in a random order. These words each participant saw were a subset from a list of 225 words sourced from the Affective Norms for English Words list[1] (ANEW; Bradley & Lang, 1999) and the Edinburgh Word Association Thesaurus (Kiss, Armstrong, Milroy, & Piper, 1973). Since some previous collaborative recall experiments have used lists of semantically related words to enhance opportunities for cross-cueing (e.g., Basden et al., 2000; Meudell et al., 1995) and have suggested that using categorized lists increases intrusions

[1] We developed the materials in this way to construct sets of semantically related words that were also normed for emotional valence. Of the 45 words each participant saw, 15 were positive, 15 were neutral, and 15 were negative, based on the ANEW database. However, the effects of valence were not reliable or consistent across tasks, and we do not report them here.

(Basden et al., 1998), we used these two sources to organize words into semantically related sets. We created five different lists, such that each participant saw a subset of 45 words: 5 words from nine sets. We counterbalanced which version of the list was learned by each group, so that each list appeared equally often in each group condition: three groups in each of the turn-taking, consensus, and nominal conditions learned each of the five versions of the list.[2]

Procedure

The experimental session consisted of seven phases conducted by a single experimenter: (a) learning; (b) Recall 1; (c) Distraction 1; (d) Recall 2; (e) Distraction 2; (f) Recall 3; and (g) postexperimental inquiry.

Learning. Participants sat at individual computers as they arrived for the experimental session. The experimenter told participants that the experiment was aimed at examining individual differences in memory and other cognitive abilities. She told participants that they would be presented with words on the computer and that they should try to remember them. She did not tell participants that the words belonged to categories. Each participant viewed the list of 45 words in a unique random order. Each word appeared once on the screen for 5 seconds, with a fixation cross appearing for 1 second between words.

Recall 1. Following Finlay et al. (2000, Experiment 1), immediately after learning participants completed an individual free recall task in which they spent 4 minutes writing down as many words as they could remember (Recall 1). The experimenter told the participants to keep trying to remember as many words as possible until they were instructed to stop, and she collected the recall sheets back when time was up.

Distraction 1. After Recall 1, participants in the two collaborative conditions were invited to move from their individual computers to a central round table (diameter 120 cm) in the same room. They sat together in a group of three around the table, and the experimenter introduced them to each other by name. She then explained that they were to complete a Sudoku number puzzle as a group, explained the rules of Sudoku, and told participants they had 10 minutes to complete the task. She then sat in a chair in the corner of the room while participants worked together to attempt to solve the Sudoku. After 10 minutes, she told them that their time was up and collected the puzzle back. We designed this collaborative distraction for participants, who had not met each other before the experimental session, to become comfortable working together.

Participants in the nominal condition remained at their individual computers. The experimenter explained that they were each to complete a Sudoku number puzzle separately, explained the rules of Sudoku, and told participants they had 10 minutes to complete the task. After 10 minutes, she told them that their time was up and collected the puzzle back from each participant.

Recall 2: Turn-taking condition. Following Distraction 1, the experimenter told participants in the turn-taking condition that they were to work as a group to recall as many words from the list as possible (Recall 2). For these groups, the experimenter sat at the table with participants, asked each person in turn to remember an item, and wrote down the items recalled on a central list placed on the table. The specific instructions were:

> For the next task I want you to think back once again to the words you saw on the computer at the beginning of the experiment. This time, I want you to work together to remember as many of the words from the list as you can. I'm going to go around and get you each to remember a word in turn. On

[2] There were five different word lists, and we counterbalanced the assigning of groups to a word list and collaboration condition such that three groups in each condition saw each list. Before we conducted the analyses we report in the Results section, we conducted a five-level (counterbalancing condition) univariate analysis of variance (ANOVA) on all the nine dependent variables. This analysis indicated no significant effects of counterbalancing condition on any of our measures, all Fs < 0.98, all ps $> .438$. Thus, we considered the five lists equivalent, and we collapsed across counterbalancing conditions in the results reported here.

your turn, you need to recall a word that hasn't already been said. If you say a word that has already been said, I'll tell you to try to think of another one. If you can't think of a word within 10 seconds, I'll move on to the next person, and once you've missed your turn three times, I'll stop asking you.

The experimenter turned on the tape recorder and commenced the recall session by turning to ask the participant on her left for the first item. The list of remembered items (recorded by the experimenter as recall proceeded) was on the table, and participants could look at the list to see which items had already been recalled. Turn-taking groups continued to recall until all 3 participants failed to recall an item on their turn. We designed this turn-taking procedure to replicate closely the procedure used in previous experiments that adopted a turn-taking style of collaboration (e.g., Basden et al., 1997; Basden et al., 2000).

Recall 2: Consensus condition. Following Distraction 1, the experimenter told participants in the consensus condition that they were to work as a group to recall as many words from the list as possible (Recall 2). For these groups, the experimenter told participants to reach a consensus about each item and to only write an item down if they all remembered seeing it. The specific instructions were:

For the next task I want you to think back once again to the words you saw on the computer at the beginning of the experiment. This time, I want you to work together to remember as many of the words from the list as you can. I'm going to give you the piece of paper, and someone needs to write down all the words that you remember. It is important that you all agree on each item that is written down, so if you don't all remember seeing an item, don't include it. Keep trying to remember until you can't remember any more.

The experimenter turned on the tape recorder and sat in a chair in the corner of the room; participants wrote the items down. The list of remembered items (recorded by participants as recall proceeded) was on the table, and participants could look at the list to see which items had already been recalled. Consensus groups continued to recall until they could not remember any more, and the experimenter asked them twice if they were sure they were finished once recall appeared to be blocked. We designed this consensus procedure to replicate closely the procedure used in previous experiments that adopted a more interactive style of collaboration (e.g., Finlay et al., 2000; Meudell et al., 1995; Weldon & Bellinger, 1997).

Recall 2: Nominal condition. Following Distraction 1, participants in the nominal condition completed an individual recall task that was identical to Recall 1. They remained at their individual computers and spent 4 minutes writing down as many of the 45 words from the learning phase as they could remember (Recall 2).

Distraction 2. Following Recall 2, all participants completed a 10-minute individual distraction phase at their individual computers, where they identified whether pairs of number strings presented on the computer were the same or different.

Recall 3. Following Distraction 2, participants in all conditions completed an individual recall task that was identical to Recall 1. They remained at their individual computers and spent 4 minutes writing down as many of the 45 words from the learning phase as they could remember (Recall 3).

Postexperimental inquiry. Following final recall, all participants separately wrote answers to a series of questions about the experiment. Among other questions (results not reported here) these questions included:

1. Did your group adopt any strategies to help you recall?
2. What were you doing while other members of your group were recalling words: (a) trying to use the other person's memory as a cue; (b) thinking about if the other person was correct or not; (c) thinking of what you would say next; (d) something else?

Following this postexperimental inquiry, participants were debriefed, given the opportunity to ask questions, and thanked for their participation.

Results

For Recall 1, which is baseline individual recall, we compared future members of turn-taking, consensus, and nominal groups; the *n* was 45 individuals per cell. For Recall 2, which focuses on costs and benefits during collaboration, we compared turn-taking groups, consensus groups, and nominal groups; the *n* was 15 groups per cell. We calculated nominal group recall by pooling the nonredundant items recalled by the three individuals assigned to each nominal group (as in Weldon & Bellinger, 1997).[3] For Recall 3, which focused on costs and benefits after collaboration, we compared former members of turn-taking, consensus, and nominal groups, but to control for possible interdependence in their responses, we analysed them as groups; the *n* was 15 groups per cell.

For each of the three recall occasions, first, we present "recall completeness", the percentage of the study list recalled: Table 1 presents the mean percentage of the 45-word list accurately recalled on Recalls 1, 2, and 3 in the nominal, turn-taking, and consensus conditions. Second, we present "recall accuracy": Table 2 presents mean percentage of total output that was inaccurate on Recalls 1, 2, and 3 in the nominal, turn-taking, and consensus conditions. Finally, we present qualitative data on strategies used by members of turn-taking and consensus groups.

Recall performance before collaboration (Recall 1)

For Recall 1 completeness, participants on average recalled less than half of the words from the original list ($M = 42.04\%$, $SD = 11.83\%$). We calculated pooled, nonredundant group recall scores for groups in all conditions, which was on average about three quarters of the words from the original list ($M = 77.58\%$, $SD = 8.39\%$). For Recall 1 accuracy, participants on average produced less than one intrusion during Recall 1 ($M = 0.42$, $SD = 0.12$), and their percentage inaccuracy was low (see Table 2). Two 3-level (condition: turn taking vs. consensus vs. nominal) univariate analyses of variance (ANOVAs) of Recall 1 completeness and percentage inaccuracy yielded no significant effects, all $Fs < 1.30$, all $ps > .28$. In summary, across our indices there was no evidence of preexisting differences in the performance of individual participants or groups in the different conditions.

Costs and benefits during collaboration (Recall 2)

For Recall 2 completeness, groups on average recalled approximately two thirds of the words from the original list ($M = 66.57\%$, $SD = 11.47\%$). A 3 (condition) univariate ANOVA of Recall 2 completeness yielded a significant effect of condition, $F(2, 42) = 12.16$, $p < .001$, $\eta_p^2 = .37$. Follow-up tests ($p < .05/3$) indicated that both turn-taking groups and consensus groups recalled less of the list than nominal groups, $t(28) = 2.75$, $p = .010$, and $t(28) = 5.40$, $p < .001$, respectively (see Table 1). That is, both turn-taking and consensus groups showed collaborative inhibition. Numerically, consensus groups recalled the least, although the difference between consensus and turn-taking groups was not significant, $t(28) = 1.96$, $p = .060$ (see Table 1).

For recall accuracy, we calculated the percentage of total output that was inaccurate (i.e., did not appear on the study list). On average, groups produced few intrusions on Recall 2 ($M = 1.42$, $SD = 1.57$). However, a 3 (condition) univariate ANOVA of Recall 2 percentage inaccuracy yielded a significant effect of condition, $F(2, 42) = 4.68$, $p = .015$, $\eta_p^2 = .18$. Follow-up tests ($p < .05/3$) indicated that consensus groups were more accurate than both turn-taking and nominal groups, $t(28) = 3.36$, $p = .002$, and $t(28) = 2.53$, $p = .017$, respectively. There was no significant difference between turn-taking and nominal groups, $t(28) = 0.18$, $p = .861$ (see Table 2).

Given that consensus groups produced (numerically) the fewest words overall, as well as the fewest intrusions, we were interested in whether this

[3] Wright (2007) offers an alternative method of forming nominal groups by creating multiple random combinations of individuals. This method increases power but was inappropriate in the current experiment, because participants were grouped based on counterbalancing condition.

Table 1. *Recall completeness*

	Nominal		*Turn taking*		*Consensus*		*Consensus: inclusive*	
	Mean	SD	*Mean*	SD	*Mean*	SD	*Mean*	SD
Recall 1 ($n = 45$ individuals)	40.97	11.63	40.84	11.99	44.31	11.79	—	—
Recall 2 ($n = 15$ groups)	75.41	8.50	65.63	10.85	58.67	8.47	64.74	7.17
Recall 3: summed ($n = 15$ groups) Total possible = 300%	117.33	18.33	136.89	29.50	142.22	25.85	—	—
Recall 3: pooled ($n = 15$ groups)	75.41	7.58	72.44	9.27	71.70	8.94	—	—

Note: Percentage of study list recalled by nominal, turn-taking collaborative, and consensus collaborative groups on Recalls 1, 2, and 3.

Table 2. *Percentage inaccurate recall by nominal, turn-taking collaborative, and consensus collaborative groups on Recalls 1, 2, and 3*

	Nominal		*Turn taking*		*Consensus*		*Consensus: inclusive*	
	Mean	SD	*Mean*	SD	*Mean*	SD	*Mean*	SD
Recall 1 ($n = 45$ individuals)	3.91	5.40	4.74	7.34	3.71	4.33	—	—
Recall 2 ($n = 15$ groups)	5.52	5.56	5.84	4.23	1.52	2.57	4.18	4.13
Recall 3: summed ($n = 15$ groups) Total possible = 300%	21.19	18.47	20.67	10.44	11.10	8.44	—	—

reflected a response bias in consensus groups (see Ross et al., 2004; Thorley & Dewhurst, 2007). By this interpretation, members of consensus groups might have been less likely to mention both correct and incorrect items, as their response criterion may have been different due to different social pressure. To test this, we calculated "inclusive scores" for consensus groups. This involved listening to the audio recordings of the recall sessions and scoring any items that were mentioned during the discussion but subsequently discounted by the group. We then recalculated completeness and accuracy for the consensus groups, based on these inclusive scores (see Tables 1 and 2). A 3 (condition) univariate ANOVA of Recall 2 completeness indicated that the groups still differed in recall completeness when scored inclusively, $F(2, 42) = 7.35$, $p = .002$, $\eta_p^2 = .26$. Follow-up tests ($p < .05/3$) indicated that, exactly as in the analysis above, both turn-taking groups and consensus groups recalled less of the list than nominal groups, $t(28) = 2.92$, $p = .007$, and $t(28) = 3.90$, $p = .001$, respectively, and there were no significant differences between turn-taking and consensus groups, $t(28) = 0.27$, $p = .793$ (see Table 1).

However, when consensus groups were scored inclusively, there was no difference in accuracy across conditions, $F(2, 42) = 0.52$, $p = .600$, $\eta_p^2 = .02$. Taken together, these analyses suggest that members of consensus groups mentioned a similar number of incorrect items as did nominal and turn-taking groups, but that these items were rejected by the group and not included in written group output. Thus individual response bias is not a complete explanation for the different performance of consensus groups.

Costs and benefits after collaboration (Recall 3)

Across our indices of memory performance (recall completeness and recall accuracy), we compared Recall 3 scores using Recall 1 scores as a covariate in both the overall tests and the follow-up pairwise comparisons,[4] to determine whether collaboration (on Recall 2) influenced final memory performance. To take into account the possible interdependence in recall scores of individuals who had previously been members of the same group, we conducted the analyses with groups (rather than individuals) as the unit of analysis and analysed both the summed recall across the three individuals and the pooled recall across the three individuals.

For Recall 3 completeness, participants on average recalled less than half the words from the original list ($M = 44.06\%$, $SD = 12.46\%$), while groups' summed recall averaged 132.15% ($SD = 26.71\%$, out of a possible 300%, given that these were the result of three summed percentages). We also calculated pooled, nonredundant group recall for all groups, which was on average about three quarters of the words from the original list ($M = 73.19\%$, $SD = 8.59\%$). A 3 (condition) univariate analysis of covariance (ANCOVA) of summed Recall 3 output (with summed Recall 1 output for the group as the covariate) yielded a main effect of condition, $F(2, 41) = 12.18$, $p < .001$, $\eta_p^2 = .37$. Follow-up tests for condition ($p < .05/3$) indicated that former members of both turn-taking and consensus groups recalled more than did former members of nominal groups, $F(1, 27) = 28.71$, $p < .001$, and $F(1, 27) = 15.01$, $p = .001$ respectively, and there was no significant difference between turn-taking and consensus groups, $F(1, 27) = 0.92$, $p = .347$ (see Table 1). The effect of the covariate was strongly significant in all comparisons, all Fs > 90.69, all ps < .001, which makes sense, as Recall 1 completeness was likely to be strongly associated with Recall 3 completeness. In terms of recall completeness, taking into account baseline recall, both turn-taking and consensus groups showed postcollaboration benefits (see Table 1).

Interestingly, while total summed recall benefited from collaboration, pooled, nonredundant group output on Recall 3 was lower for turn-taking ($M = 72.44\%$, $SD = 9.27\%$) and consensus ($M = 71.70\%$, $SD = 8.94\%$) groups than for nominal groups ($M = 75.41\%$, $SD = 7.58\%$): A 3 (condition) univariate ANCOVA of pooled Recall 3 output (with pooled Recall 1 output as a covariate) yielded a main effect of condition, $F(2, 41) = 4.62$, $p = .016$, $\eta_p^2 = .18$, as well as a main effect of the covariate, $F(1, 41) = 117.12$, $p < .001$, $\eta_p^2 = .741$. Follow-up tests ($p < .05/3$) indicated that pooled recall of former members of both turn-taking groups (marginally) and consensus groups was lower than pooled recall of nominal groups, $F(1, 27) = 6.23$, $p = .019$, and $F(1, 27) = 11.82$, $p < .002$, respectively, and there was no significant difference between turn-taking and consensus groups, $F(1, 27) = 0.15$, $p = .703$; the covariate was significant in all comparisons, all Fs > 56.69, all ps < .001. This analysis suggests that collaboration produced more overlap in individual recall, such that collaboration increased total recall, but decreased nonredundant recall.

We checked this interpretation by scoring, for each individual, the proportion of items recalled on Recall 3 that they had not recalled on Recall 1, but had been mentioned by any group member on Recall 2—that is, the proportion of items "picked up" during collaboration (with the nominal group members as a baseline control). We then summed these "picked up" items to create a score of total picked items on Recall 3 for each group. A 3 (condition) univariate ANOVA of proportion of total "picked up" items yielded a main effect of condition, $F(6, 42) = 43.84$, $p < .001$, $\eta_p^2 = .68$. Follow-up tests ($p < .05/3$) indicated that both turn-taking groups ($M = 65.01\%$, $SD = 17.85\%$, out of a possible 300%,

[4] We are grateful to Dan Wright for suggesting this analysis strategy.

since they are sum scores) and consensus groups ($M = 53.38\%$, $SD = 19.60\%$) had more "picked up" items in their individual recall than did nominal groups ($M = 13.85\%$, $SD = 5.95\%$), $F(1, 28) = 11.88$, $p < .001$, and $F(1, 28) = 55.87$, $p < .001$, respectively, and there was no significant difference between turn-taking and consensus groups, $F(1, 28) = 2.89$, $p = .100$. Members of nominal groups gained very few items between Recall 1 and Recall 3.

Finally, for recall accuracy we calculated the percentage of total output that was inaccurate (i.e., did not appear on the study list). On average, individuals produced few intrusions on Recall 3 ($M = 1.16$, $SD = 1.39$), while groups' summed inaccuracy averaged 17.65% ($SD = 13.75\%$, out of a possible 300%, given that these were the results of three summed percentages). A 3 (condition) univariate ANCOVA of groups' total Recall 3 inaccuracy (with total Recall 1 inaccuracy as a covariate) yielded a main effect of condition, $F(2, 41) = 3.56$, $p = .038$, $\eta_p^2 = .15$, as well as a main effect of the covariate (Recall 1 scores), $F(1, 41) = 30.80$, $p < .001$, $\eta_p^2 = .43$. Follow-up tests ($p < .05/3$) indicated that former members of consensus groups were more accurate than turn-taking groups and marginally more accurate than nominal groups, $F(1, 27) = 7.65$, $p = .010$, and $F(1, 27) = 5.91$, $p = .022$, respectively, and there was no significant difference between turn-taking and nominal groups, $F(1, 27) = 0.51$, $p = .482$ (see Table 2); the covariate was significant in all comparisons, all Fs > 18.63, all ps $< .001$. Thus, after collaboration, taking into account baseline recall, former members of consensus groups showed benefits in terms of recall accuracy (see Table 2).

Individual and group strategies

Participants' postexperimental comments suggested that individuals in the turn-taking and consensus conditions behaved differently from each other during collaboration. We asked participants what they were doing while other people in the group recalled items. In the turn-taking groups, 13.33% of participants reported that they were checking the accuracy of others' recall, whereas 77.78% of participants reported that they were thinking of what they were going to say next, and 8.89% of participants reported that they were using the recall of others to cue their own memories. In the consensus groups, 42.22% of participants reported that they were checking the accuracy of others' recall, whereas 51.11% of participants reported that they were thinking of what they were going to say next, and 6.67% of participants reported that they were using the recall of others to cue their own memories. This pattern was significantly different across the two groups, $\chi^2(3, N = 90) = 10.87$, $p = .01$, and suggests that consensus groups engaged in more accuracy checking than turn-taking groups, but that they did not adopt a more interactive cross-cueing strategy to increase recall.

Consistent with the individualistic nature of the reported individual strategies, participants' postexperimental comments indicated that in general, groups did not adopt explicit group-level strategies to coordinate and maximize their collaborative recall. Members of turn-taking groups always reported that their group had not adopted any strategies, which is to be expected since they had not been given the opportunity to do so. However, members of only 2 of the 15 consensus groups reported that their group had adopted group-level strategies. In these 2 groups, the participants stated that they had attempted to cue each others' recall using the categories. However, the performance of these 2 groups was no different to the performance of the other consensus groups, so there was no evidence that these reported group-level strategies were effective.

Discussion

This experiment compared the costs and benefits of turn-taking versus consensus collaboration for recall completeness and recall accuracy, both during and after collaboration. We also explored the influence of response criterion and examined reported group and individual strategies during collaboration by conducting a postexperimental inquiry.

Recall completeness

Consistent with previous research, during collaboration both turn-taking and consensus groups showed costs in terms of amount recalled; on Recall 2 these groups recalled fewer words from the list than did nominal groups. That is, both types of collaboration resulted in collaborative inhibition. In general, our findings reinforce the robustness of the collaborative inhibition effect, despite differences in the nature of the instructions given to the different collaborative groups (see Harris et al., 2008, for review). These results suggest that individuals in both kinds of collaborating groups experienced a disruption of individual retrieval strategies and did not cross-cue each others' recall (see Meudell et al., 1995). This interpretation is also supported by participants' postexperimental comments, where consensus groups reported that their group had not used any strategies, and individuals did not report using the recall of others to cue their own recall.

Consistent with some previous research (see Rajaram & Pereira-Pasarin, 2010), we also found that prior collaboration increased subsequent individual recall, regardless of collaboration type. This postcollaboration benefit was likely due to reexposure to list items during collaboration (see Blumen & Rajaram, 2008). In support of this interpretation, we found that for former members of collaborative groups, pooled group recall was actually lower than that for former members of nominal groups, despite their total recall being higher. This suggests that after collaboration, individual recall was more similar and more overlapping. In general, both types of collaborative groups—consensus and turn taking—showed the usual pattern of costs during collaboration and benefits after collaboration (Basden et al., 2000). At least in terms of recall completeness, we found no evidence for advantages or disadvantages of consensus collaboration compared to turn-taking collaboration.

Recall accuracy

Consensus collaboration did, however, provide advantages in terms of increasing recall accuracy. Previous researchers have reported varying results for the accuracy of collaborative group recall relative to nominal group recall, but they have used different instructions for collaboration. We predicted that during collaboration, consensus groups might be more accurate than turn-taking groups because they would benefit from the error-checking opportunities provided by the interactive group. We also predicted that members of consensus groups might continue to benefit from the groups' influence on subsequent individual recall. Consistent with these predictions, we found that consensus groups produced approximately one quarter as many intrusions as turn-taking groups and nominal groups. Although the overall intrusion rate was not high, the difference across types of collaboration was clear and continued after collaboration. On Recall 3, members of consensus groups produced half as many intrusions as did members of turn-taking groups and nominal groups. Thus, consensus collaboration (and not turn-taking collaboration) had benefits for recall accuracy, both during and after collaboration.

This finding is consistent with comments from the postexperimental inquiry. In the consensus groups, 42% of people reported checking the accuracy of others' recall, whereas only 13% in the turn-taking groups reported doing so. Even if only one member of a consensus group was checking accuracy (while the other two thought of what they would recall next) as was the case in the majority of the consensus groups, this was enough to improve the accuracy of the whole group both during the collaboration (when the accuracy checker had immediate influence by providing feedback on errors) and after the collaboration.

To study whether this increased accuracy in consensus groups was due to a more stringent response criterion, we constructed inclusive scores for consensus group recall. This analysis tested whether group members were simply less likely to mention both accurate and inaccurate items. This analysis indicated that, scored inclusively, the accuracy of consensus groups was similar to that of the other groups. That is, in consensus groups, individuals mentioned just as many incorrect items, but these items were removed from the written group output by the error-checking process in the group (similar to Ross et al., 2008). Thus, our findings are more consistent with a group error-checking process than a response bias. This group error checking is especially important because, as well as eliminating errors from group recall output, it improved accuracy on subsequent individual recall.

Interestingly, this "group error checking" is conceptually similar to "source monitoring" in individual memory. Source monitoring is defined as the process by which "activated memory records are evaluated and attributed to particular sources through decision processes performed during remembering" (Johnson, Hashtroudi, & Lindsay, 1993, p. 3). That is, people use the qualities of remembered information to judge the likely source of this information (e.g., as a genuinely experienced event vs. a dream). The emphasis on source monitoring as a dynamic process that occurs during the act of remembering means that the concept has clear parallels for collaborative remembering. In the collaborative context then, the evaluation and attribution of activated memory records is done within or by the group, who can override an individual judgement when it is in the minority or when it is deemed to be incorrect. Our consensus procedure encouraged this process by emphasizing agreement.

Future research could focus on whether groups engage in source monitoring spontaneously in more free-for-all collaborative procedures (without a consensus requirement). However, existing data showing that free-for-all collaboration can reduce errors (at least during the collaboration itself; e.g., Ross et al., 2008) suggest that this is possible. Future research could also focus on the different criteria that groups use for making source attributions, such as the effect of an individual's confidence on the group's decision, and how these relate to the criteria used by individuals. It is also possible that the emphasis on accuracy created a "group-level" response bias in our consensus groups, such that the group adopted more stringent criteria for including items in recall. However, our analysis of items that were rejected by the group indicates that groups were very good at identifying errors and were not simply discarding both correct and incorrect items. Future research could follow up this alternative possibility and, more generally, test whether similar processes operate in individual and group recall.

Our finding that consensus collaboration has benefits for recall accuracy both during and after collaboration has practical implications and challenges some of the assumptions in related literatures on group remembering. In false-memory research, for example, social influence on memory tends to be equated with distortion (Barnier et al., 2008). Studies of cowitness discussion (e.g., social contagion and memory conformity) have emphasized the potential negative effects of witnesses remembering an event together (e.g., Gabbert et al., 2003; Gabbert et al., 2006; Meade & Roediger, 2002; Roediger et al., 2001). Perhaps one recommendation would be to instruct these groups to check accuracy and to reach a consensus, as in our experiment.

However, memory conformity studies involve two or more participants unwittingly witnessing slightly different versions of the same event (e.g., French, Garry, & Mori, 2008; Gabbert et al., 2003). Because participants view different versions of the stimuli in these paradigms, reaching a consensus in collaboration is unlikely to reduce errors since the experimental situation "conspires" against normal error-checking processes. Because there is no single stimulus to be accurately recalled, when coming to a consensus requires one participant to yield to the other—one participant to agree to an inaccurate version of events—nothing can be resolved successfully through discussion (but see Skagerberg & Wright, 2008). Similarly, in social contagion studies (e.g., Roediger et al., 2001) a confederate deliberately introduces errors into collaborative recall. Again, because the confederate is not a genuine participant seeking to maximize collaborative recall, accuracy checking and consensus collaboration are not likely to be effective.

The procedures for memory conformity and social contagion research have been developed to ensure that sufficient memory errors are produced for analysis, and error rates in these paradigms are generally higher than those in collaborative recall experiments. However, while in the memory conformity and social contagion paradigms yielding on the part of one person results in an error (e.g., French et al., 2008), in real-world cases (where there is a single stimulus, and all group members strive to maximize recall), yielding on the part of one person might be quite beneficial to accuracy both during and after collaboration, as in our consensus groups. Indeed, previous research in the memory conformity paradigm (where participants view slightly different stimuli from each other, prior to group recall) suggests that, for stimuli that are common to both participants, discussion can enhance accuracy (Wright, Self, & Justice, 2000) and that when accuracy is a matter of degree, collaboration can result in a blended response where both participants are partially accurate (Skagerberg & Wright, 2008). Although the collaborative recall paradigm results in relatively low error rates, our finding that consensus collaboration improves recall accuracy both during and after collaboration is relevant to real-world concerns about memory accuracy.

Limitations and future directions

While the results of this study are clear-cut, particularly regarding accuracy, the conclusions must be tempered by our sample size. We tested 135 participants, but collaborative recall experiments require the comparison of groups rather than individuals, which drastically reduces the *n* and hence the power to detect differences between groups. Future research could follow up potential differences in the magnitude of collaborative inhibition, since our results suggested that consensus groups may have recalled the least overall. Although the difference was not significant in our study, we may not have had sufficient power to detect it. Secondly, future research could more closely match the collaboration instructions. While we developed the instructions for groups based on instructions used in previous research, our groups varied on a number of dimensions. The turn-taking and consensus groups differed in both the procedure of collaboration and the emphasis on productivity versus accuracy. To some extent these differences are inherent in the comparisons that we are making, but future research could focus on teasing apart the various influences in different kinds of collaborative groups.

Conclusion

Our comparison of turn-taking and consensus collaborative recall yielded important differences. Whereas both resulted in collaborative inhibition and increased individual recall after collaboration, only consensus collaboration resulted in increased recall accuracy both during and after collaboration. That is, consensus collaboration had particular benefits for subsequent individual recall, increasing both recall completeness and recall accuracy. Our analysis suggested that these accuracy benefits were the result of group source monitoring rather than differences in response criterion. There were differences in the strategies used by individuals in the two kinds of collaborating groups, but very few participants reported using the recall of others to cue their own memories. These findings reinforce the robustness of collaborative inhibition, but demonstrate that social influences on memory can enhance both group and individual recall accuracy. Features of the group and the nature of their interaction during collaboration are important determinants of how remembering with others can increase or decrease memory errors.

Original manuscript received 6 November 2010
Accepted revision received 7 July 2011
First published online 22 September 2011

REFERENCES

Alea, N., & Bluck, S. (2003). Why are you telling me that? A conceptual model of the social function of autobiographical memory. *Memory*, *11,* 165–178.

Andersson, J., Hitch, G., & Meudell, P. (2006). Effects of the timing and identity of retrieval cues in individual recall: An attempt to mimic cross-cueing in collaborative recall. *Memory*, *14,* 94–103.

Barnier, A. J., Sutton, J., Harris, C. B., & Wilson, R. A. (2008). A conceptual and empirical framework for the social distribution of cognition: The case of memory. *Cognitive Systems Research*, *9,* 33–51.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. London, UK: Cambridge University Press.

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 & 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.

Basden, B. H., Basden, D. R., Thomas, R. L., III, & Souphasith, S. (1998). Memory distortion in group recall. *Current Psychology*, *16,* 225–246.

Blumen, H. M., & Rajaram, S. (2008). Re-exposure and retrieval disruption during group collaboration as well as repeated retrieval influence later individual recall. *Memory*, *16,* 231–244.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings*. Technical report C-1, Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

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

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

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,* 262–273.

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

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

Harris, C. B., Barnier, A. J., Sutton, J., & Keil, P. (2010). How did you feel when "The Crocodile Hunter" died? Voicing and silencing in conversation influences memory for an autobiographical event. *Memory*, *18,* 185–197.

Harris, C. B., Keil, P. G., Sutton, J., Barnier, A. J., & McIlwain, D. (2011). We remember, we forget: Collaborative remembering in older couples. *Discourse Processes*, *48,* 267–303.

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

Harris, C. B., Sutton, J., & Barnier, A. J. (2010). Autobiographical forgetting, social forgetting, situated forgetting: Forgetting in context. In S. Della Sala (Ed.), *Forgetting* (pp. 253–284). New York, NY: Psychology Press.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source-monitoring. *Psychological Bulletin*, *114,* 3–28.

Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. In A. J., Aitken, R. W. Bailey, N. Hamilton-Smith (Eds.), *The Computer and Literary Studies*. Edinburgh: University Press. Retrieved May 5, 2006, from http://www.eat.rl.ac.uk/.

Loftus, E. F. (1997). Creating false memories. *Scientific American*, *277,* 50–55.

Maki, R. H., Weigold, A., & Arellano, A. (2008). False memory for associated word lists in individuals and collaborating groups. *Memory & Cognition*, *36,* 598–603.

Meade, M. L., Nokes, T. J., & Morrow, D. G. (2009). Expertise promotes facilitation on a collaborative memory task. *Memory*, *17,* 39–48.

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

Meudell, P. R., Hitch, G. J., & Boyle, M. M. (1995). Collaboration in recall: Do pairs of people cross-cue each other to produce new memories? *The Quarterly Journal of Experimental Psychology*, *48A,* 141–152.

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

Reysen, M. B. (2003). The effects of social pressure on group recall. *Memory & Cognition*, *31,* 1163–1168.

Roediger, H. L. (1973). Inhibition in recall from cueing with recall targets. *Journal of Verbal Learning & Verbal Behavior*, *12,* 261–269.

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

Roediger, H. L., Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review*, *8,* 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,* 85–92.

Ross, M., Spencer, S. J., Linardatos, L., Lam, K.C.? H., & Perunovic, M. (2004). Going shopping and identifying landmarks: Does collaboration improve older people's memory? *Applied Cognitive Psychology*, *18,* 683–696.

Skagerberg, E. M., & Wright, D. B. (2008). The co-witness misinformation effect: Memory blends or memory compliance? *Memory*, *16,* 436–442.

Takahashi, M. (2007). Does collaborative remembering produce false memories? *British Journal of Psychology*, *98,* 1–13.

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.

Tollefsen, D. P. (2006). From extended mind to collective mind. *Cognitive Systems Research*, *7,* 140–150.

Weldon, M. S. (2001). Remembering as a social process. *Psychology of Learning and Motivation*, *40,* 67–120.

Weldon, M. S., & Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. *Journal of Experimental Psychology: Learning, Memory & 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 & Cognition*, *26,* 1568–1577.

Wright, D. B. (2007). Calculating nominal group statistics in collaboration studies. *Behavior Research Methods*, *39,* 460–470.

Wright, D. B., & Klumpp, A. (2004). Collaborative inhibition is due to the product, not the process, of recalling in groups. *Psychonomic Bulletin & Review*, *11,* 1080–1083.

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.