Psychological Bulletin

A Meta-Analytic Review of Collaborative Inhibition and Postcollaborative Memory: Testing the Predictions of the Retrieval Strategy Disruption Hypothesis

Stéphanie B. Marion and Craig Thorley

Online First Publication, September 12, 2016. http://dx.doi.org/10.1037/bul0000071

Psychological Bulletin

0033-2909/16/$12.00 http://dx.doi.org/10.1037/bul0000071

A Meta-Analytic Review of Collaborative Inhibition and Postcollaborative Memory: Testing the Predictions of the Retrieval Strategy Disruption Hypothesis

Stéphanie B. Marion
Ryerson University

Craig Thorley
University of Liverpool

The *retrieval strategy disruption hypothesis* (Basden, Basden, Bryner, & Thomas, 1997) is the most widely cited theoretical explanation for why the memory performance of collaborative groups is inferior to the pooled performance of individual group members remembering alone (i.e., collaborative inhibition). This theory also predicts that several variables will moderate collaborative inhibition. This meta-analysis tests the veracity of the theory by systematically examining whether or not these variables do moderate the presence and strength of collaborative inhibition. A total of 75 effect sizes from 64 studies were included in the analysis. Collaborative inhibition was found to be a robust effect. Moreover, it was enhanced when remembering took place in larger groups, when uncategorized content items were retrieved, when group members followed free-flowing and free-order procedures, and when group members did not know one another. These findings support the retrieval strategy disruption hypothesis as a general theoretical explanation for the collaborative inhibition effect. Several additional analyses were also conducted to elucidate the potential contributions of other cognitive mechanisms to collaborative inhibition. Some results suggest that a contribution of retrieval inhibition is possible, but we failed to find any evidence to suggest retrieval blocking and encoding specificity impact upon collaborative inhibition effects. In a separate analysis (27 effect sizes), moderating factors of postcollaborative memory performance were examined. Generally, collaborative remembering tends to benefit later individual retrieval. Moderator analyses suggest that reexposure to study material may be partly responsible for this postcollaborative memory enhancement. Some applied implications of the meta-analyses are discussed.

Keywords: collaborative remembering, collaborative inhibition, retrieval strategy disruption, memory

Within cognitive psychology, memory researchers have traditionally focused on how individuals encode, maintain, and retrieve past experiences. Memory, however, is often social in nature and collaborative remembering can occur in a number of different settings, including the workplace (e.g., interview panels jointly recalling a candidate's performance), the courtroom (e.g., jurors jointly recollecting trial evidence during deliberation), and the classroom (e.g., students revising course content during a group study session). It is only in the last two decades that a critical mass of memory research has been conducted examining the impact of collaboration on memory (see Rajaram, 2011).

Intuitively, groups who are remembering a shared experience should be able to yield more information than individuals, because groups working together should have more resources and thus provide greater output than one individual working alone. As it turns out, this is the case: Groups tend to outperform individuals on memory tasks (see Clark & Stephenson, 1989, for a review). However, the comparison between group and individual memory performance does not inform us as to whether or not lone individuals and individuals within the group perform similarly. In other words, do individual group members perform to their full potential when they collaborate? A more valid test of a group's performance requires a comparison between a group's performance and that group's *potential* performance based on the pooled individual performance of each member. Such a group is termed a *nominal group*: In experimental designs, a nominal group is made up of a number of participants remembering in isolation that is equivalent in size to the number of participants in a collaborative group. In a typical collaborative remembering experiment, participants study materials (e.g., items A, B, C, D, E, and F) individually and are then asked to recall the material either individually or collectively as part of a group. The items remembered by those recalling the material individually (the nominal group members) are then combined, but the same items remembered by more than one individual are counted only once. For example if individual 1 remembers items A, B, and C, individual 2 remembers items B, C, and F, and individual 3 remembers A, C, E, and F, then their combined score would be five items (A, B, C, E, and F). This nominal group score is then compared to the number of items remembered by the collaborative group. Several experiments using this paradigm have demonstrated, perhaps counterintuitively, that the memory perfor-

Stephanie B. Marion, Department of Psychology, Ryerson University; Craig Thorley, Department of Psychological Sciences, University of Liverpool.

Correspondence concerning this article should be addressed to Stephanie B. Marion, who is now at the Faculty of Social Science and Humanities, University of Ontario Institute of Technology, Oshawa, Ontario, Canada, L1H 7K4. E-mail: stephanie.marion@uoit.ca

mance of noninteracting nominal groups (as measured by the amount of correctly recalled items) is reliably better than the memory performance of equal-sized collaborative groups (e.g., Basden, Basden, & Henry, 2000; Pereira-Pasarin & Rajaram, 2011; Weldon & Bellinger, 1997). In other words, the individual participants' memory performance is lower if they were part of a collaborative group than if they worked in isolation during a memory task. This phenomenon has been termed *collaborative inhibition* (CI; Weldon & Bellinger, 1997).

Proposed Explanatory Mechanisms Behind the Collaborative Inhibition Effect

Since CI was first reported in the literature, several mechanisms have been proposed to explain it. An overview of the social, motivational, and cognitive factors that have been considered is provided next.

Social and Motivational Mechanisms

As collaborative remembering occurs within social contexts, it is only fitting that social and motivational factors be considered as contributors to CI. Weldon, Blair, and Huebsch (2000) examined whether or not CI is caused by social loafing, whereby group members reduce their individual effort due to a diffusion of responsibility (Karau & Williams, 1993; Latane, Williams, & Harkins, 1979), or evaluation apprehension, whereby group members refrain from contributing to a task in fear of being negatively evaluated by other group members who may perceive their contributions as inadequate or erroneous (Collaros & Anderson, 1969; Mullen, 1983; Mullen & Baumeister, 1987). Both can reduce individual group members' productivity during physical tasks like rope pulling, clapping, and shouting (e.g., Harkins, Latane, & Williams, 1980; Ingham, Levinger, Graves, & Peckham, 1974; Latane et al., 1979; Williams, Harkins, & Latane, 1981) and cognitive tasks like brainstorming (Taylor, Berry, & Block, 1958), perceptual vigilance (e.g., Harkins & Petty, 1982), and maze-learning (e.g., Griffith, Fichman, & Moreland, 1989). Across five experiments, Weldon and colleagues increased group members' motivation to recall items (e.g., by offering them monetary incentives for better performance), decreased evaluation apprehension (e.g., by encouraging them to suggest items they were uncertain of), and reduced the likelihood of diffused responsibility (e.g., by increasing group cohesiveness and individual accountability for items recalled), but found that none of these manipulations reduced CI. Since Weldon et al. (2000) studies, the impact of social and motivational factors on CI have been largely ignored and researchers have instead focused upon cognitive factors.

Retrieval Strategy Disruption

The most widely accepted cognitive mechanism underlying CI, and the one cited in virtually all published research reporting a CI effect, is Basden, Basden, Bryner, and Thomas (1997) *retrieval strategy disruption hypothesis* (RSDH). This explanation derives from Basden and colleagues' earlier research into part-list cuing inhibition (e.g., Basden & Basden, 1995; Basden, Basden, & Galloway, 1977). In part-list cuing inhibition, individual memory performance is decreased when several items of the to-be-remembered material are presented as cues for the rest of the items, in comparison with when no cues are presented (see Nickerson, 1984, for a review). Basden and colleagues explain this effect by suggesting that individuals subjectively organize newly learned information in a way that depends upon their own knowledge, schemas, experiences, and expectations of the retrieval context. For example, when learning a word list of city names, a well-travelled participant may organize this information based upon the order in which he or she visited these cities. Subsequent recall of this information will be greatest when individuals use their subjective organizational structure to guide their retrieval strategy (e.g., if the well-travelled participant can recall the cities in the order they were visited). Part-list cues will often be presented in an order that is inconsistent with an individual's retrieval strategy, forcing this person to abandon his or her own optimal retrieval strategy and switch to a less effective one, thus reducing the volume of information recalled.

Similar to the disruption that presented items cause in a part-list cuing paradigm, Basden et al. (1997) suggested that exposure to other group members' responses during collaborative retrieval may be responsible for CI. They argued that collaborative group members each develop their own subjective organization of studied materials during encoding and consequently their optimal retrieval strategies are likely to differ. During collaborative remembering, the items generated by other participants are equivalent to part-list cues. When participants hear other group members recall information in an order that is inconsistent with their own retrieval strategies (e.g., our well-travelled participant hears another recalling the cities in alphabetical order) they have to change to different, less effective, retrieval strategies, resulting in CI. As nominal group members always work alone they are not exposed to any disruption and are free to rely on their own optimal retrieval strategies, resulting in greater recall for them than collaborative group members.

Basden et al. (1997) provided empirical support for their suggestion that retrieval strategy disruption causes CI. They gave participants categorized word lists to remember and found that collaborative groups recalled fewer words than nominal groups; a clear demonstration of CI. Importantly, collaborative groups showed less clustering (grouping of semantically related words) than individuals, meaning that the collaborative group members switched categories more frequently than individuals in nominal groups. This is consistent with the idea that individual group members attempted to follow their own retrieval strategies but were disrupted by each other's contributions. Basden et al. (1997) also demonstrated that pre- and postcollaborative manipulations designed to align collaborative group members' retrieval strategies reduced CI. For example they manipulated participants' subjective organization of the categorized word lists at encoding by having them Study 15 words from each of six categories or six words from each of 15 categories. Basden and Draper (1973) had previously demonstrated that participants are more likely to have subjectively organized studied materials differently when encoding content from larger categories than smaller categories, making larger categories more susceptible to retrieval strategy disruption from part-list cuing. In line with this, CI disappeared when collaborative groups recalled items from small categories, presumably as individual members' retrieval strategies were better aligned. Basden et al. (1997) were also able to reduce CI by forcing all collaborative

group members to recall words from one category at a time until output for that category was exhausted. This latter manipulation presumably also better aligned their retrieval strategies.

Since Basden et al. (1997) first proposed the RSDH, others have examined whether or not additional pre- and postcollaborative factors designed to strengthen and align retrieval strategies can reduce CI but the results have been mixed (see Rajaram, 2011, and Rajaram & Pereira-Pasarin, 2010, for reviews). For example, some studies have supported the RSDH's prediction that CI should be reduced when participants have similar encoding strategies (and thus similar retrieval strategies, e.g., Barber, Rajaram, & Fox, 2012; Finlay, Hitch, & Meudell, 2000; Garcia-Marques, Garrido, Hamilton, & Ferreira, 2012; Harris, Barnier, & Sutton, 2013) but others have not (Barber & Rajaram, 2011b; Dahlström, Danielsson, Emilsson, & Andersson, 2011). In addition, according to the RSDH, CI should be eliminated when the test format does not allow the test taker to rely on his or her own organizational strategy, such as a cued recall or recognition tests. This is because these test formats are equally disruptive to the organizational strategies of both collaborative and nominal group members. Again, some research supports this prediction (e.g., Barber, Rajaram, & Aron, 2010; Finlay et al., 2000; Thorley & Dewhurst, 2009) but others have reported CI when assessing memory via cued-recall tests (e.g., Kelley, Reysen, Ahlstrand, & Pentz, 2012; Meade & Roediger, 2009) or recognition tests (e.g., Andersson, 2001; Andersson & Rönnberg, 1996).

Alternative Cognitive Mechanisms for CI

As research findings are sometimes inconsistent with the RSDH explanation of CI, several investigators have questioned whether or not there are multiple factors responsible for CI. Although the focus of this meta-analysis is the RSDH, we provide an overview of these alternative explanations for CI as several will be considered in the current meta-analyses.

Barber, Harris, and Rajaram (2015) point out that multiple mechanisms are thought to underlie part-list cuing inhibition (see Nickerson, 1984) and these might also play a role in CI. In addition to *retrieval strategy disruption*, others have suggested that part-list cuing can be caused by *retrieval inhibition,* which occurs when cue words suppress the memory representations of the noncued words in memory, making them permanently inaccessible (e.g., Anderson, Bjork, & Bjork, 1994; Bäuml & Aslan, 2004). It has also been argued that part-list cuing results from *retrieval blocking* whereby cue words are continually brought to mind during retrieval, blocking access to the noncued words (e.g., Rundus, 1973). The impact of these three mechanisms during part-list cuing tasks can be differentiated by examining memory for noncued items on subsequent free recall and recognition tests in which no part-list cues are present. According to the *retrieval strategy disruption account,* there should be no decrease in performance on any subsequent tests as the cues that originally disrupted retrieval are no longer present and participants can return to their own optimal retrieval strategies (e.g., Basden & Basden, 1995). If *retrieval inhibition* is at play, however, there should be a *lasting* decrease in performance on any subsequent test as the part-list cues from the initial test suppressed the memory representations of the noncued words (e.g., Aslan, Bäuml, & Grundgeiger, 2007; Bäuml & Aslan, 2006). Finally, the *retrieval blocking account* suggests that that there should be decrements on free recall tests as the noncued words have been blocked but that this drop in performance will be absent on recognition tests because noncued words are encountered without the test taker needing to access them. To assess the impact of each mechanism on CI, Barber et al. had participants collaborate to recall studied word lists and then work alone to complete a free recall or recognition test. During collaboration, CI was observed. Barber et al. (2015) also found evidence of *retrieval inhibition* on the subsequent memory test, with former collaborative group members recalling and recognizing fewer studied items than participants who had worked alone twice. Importantly, although a decrease in memory performance persisted on a subsequent individual free-recall test, it was attenuated, suggesting that *retrieval strategy disruption* may also have contributed to the initial CI (see also Garcia-Marques et al., 2012, for evidence that retrieval blocking has no impact upon CI).

An additional factor that has been suggested as a contributor to CI is *production blocking* (Diehl & Stroebe, 1987; Nijstad, Stroebe, & Lodewijkx, 2003). The production blocking explanation stems from brainstorming research where it has been observed that collaborative groups generate fewer ideas than equivalent sized nominal groups (e.g., Diehl & Stroebe, 1987; Taylor et al., 1958). Because brainstorming and collaborative remembering both rely on memory—brainstorming requires searching semantic memory for new ideas (Brown & Paulus, 2002; Brown, Tumeo, Larey, & Paulus, 1998; Nijstad, & Stroebe, 2006) and collaborative remembering typically requires searching episodic memory for studied information—it is possible that similar factors could contribute to a decreased performance in both. Production blocking theorists suggest it is the process of turn taking that causes decreased output within brainstorming groups (Diehl & Stroebe, 1987; Nijstad et al., 2003). Due to the implicit rule within collaborative groups that only one person should speak at a time, group members cannot always express their ideas the moment they come to mind. During the delay between generation and expression, group members must hold a suggestion in working memory, monitor the contributions of others, and find an opportunity to present their ideas. Nijstad, Stroebe, and Lodewijkx (2003) suggest that these cognitive processes can limit group members' ability to search for additional ideas. Production blocking therefore differs from the RSDH as the latter suggests it is the content generated by the groups that causes CI, as opposed to the delay between responses.

Several researchers have investigated whether or not production blocking contributes to CI, again with mixed results. Wright and Klumpp (2004) had collaborative pairs take turns to recall studied words, but only half of the pairs could see their collaborative partner's contributions. Contrary to production blocking theory, simply waiting for a partner to respond was not sufficient to produce CI. In line with the RSDH, CI only occurred when partners could see the content of each other's responses. More recently, Hyman, Cardwell, and Roy (2013) investigated whether or not collaborative remembering can result in a restricted exploration of memory, as predicted by production blocking theory. To assess this, they presented dyads with words that were blocked into categorized lists. CI was observed and the collaborative dyads recalled items from fewer categories overall than the nominal dyads, meaning they explored memory less effectively during collaboration and this could have caused their CI. Production blocking, then, may also contribute to CI (see also Andersson,

Hitch, & Meudell, 2006, for evidence suggesting both production blocking and retrieval strategy disruption contribute to CI).

Another potential alternative explanation for CI is that it is an artifact of the traditional paradigm used in collaborative remembering studies, whereby nominal group members benefit from *encoding specificity* (Tulving & Thomson, 1973). In the traditional collaborative remembering paradigm, all participants initially encode material individually, half of whom later remember this material again individually (i.e., nominal group members) and half of whom later remember collaboratively (i.e., collaborative group members). Consequently, there is a mismatch in the encoding and retrieval contexts of collaborative group members but a match in the encoding and retrieval contexts of nominal group members. This discrepancy could provide nominal group members an advantage because of context-dependent learning benefits (i.e., similarity in the physical surroundings of participants at encoding and retrieval) as well as transfer-appropriate processing, or TAP (i.e., the same cognitive processes utilized at encoding are utilized at retrieval). Collaborative group members, on the other hand, experience analogous disadvantages as the physical contexts change between individual study and group remembering tasks, and the cognitive processes involved change between encoding material on their own and remembering while interacting with others.

A small number of researchers have investigated the impact of matched encoding and retrieval contexts on CI. For example, Andersson and Rönnberg (1997) had participants engage in a cued recall task in which they first generated cues for the to-be-remembered words, either individually or collaboratively with their partner. Collaborative retrieval was inferior to nominal group retrieval, but this anticipated CI effect was only found among those who generated cues individually, prompting the authors to suggest that compatibility between the encoding and retrieval contexts can reduce CI. A similar pattern of results was found in other studies in which participants collaborated at encoding (Barber et al., 2012; Finlay et al., 2000) but others failed to show an effect of encoding specificity (Barber et al., 2010; Ross, Spencer, Linardatos, Lam, & Perunovic, 2004). In another study, Harris, Barnier, and Sutton (2013) manipulated interaction at encoding (individual vs. collaborative encoding) and had all participants retrieve the material once on their own on a first recall task before manipulating group collaboration on a second recall task (i.e., nominal vs. collaborative recall). They found that CI was eliminated on the second recall for those who collaboratively encoded the material, but that interaction at encoding did not impact participants' performance on the first individual recall even though cognitive and physical contexts would have been dissimilar on this test for those who encoded collaboratively, casting doubt on the importance of encoding specificity on this type of episodic recall task.

Aims of the Meta-Analysis: Collaborative Remembering

CI seems to be a generally reliable effect. The dominant theoretical explanation for the effect is the RSDH, although the above review demonstrates that other contributing factors may also exist. The RSDH makes clear predictions regarding which participant- and study-level factors should moderate CI. In brief, factors which reduce the extent to which collaborative group members disrupt each other's retrieval strategies should also reduce CI. As discussed, research that either directly or indirectly manipulates some of these variables has produced contradictory results. These contradictory results are perhaps not surprising given that there may be multiple causes of CI. Because the RSDH has been widely cited as an explanatory mechanism of CI, the first aim of this meta-analysis is to test whether the variables predicted by the RSDH to moderate CI actually do so. A systematic review of RSDH's role in CI has never been conducted. Where possible, based on the availability of different study variables, we also test other mechanisms that have been proposed as causal contributors to CI (encoding specificity, retrieval inhibition, and retrieval blocking). A second aim of this report is to determine, with a second meta-analysis, under which circumstances collaboration hurts or helps later individual retrieval (described in the next section on Postcollaborative Memory). The objective and transparent process of a meta-analysis is a much more efficient and valid method of identifying factors that are associated with the magnitude of an effect (Borenstein, Hedges, Higgins, & Rothstein, 2009; Chan & Arvey, 2012). It will allow each moderator to be examined with a greater degree of power than is often observed in individual studies (Cohn & Becker, 2003). These meta-analyses will also help provide future researchers with the effect sizes needed to perform a priori power calculations so they can determine the sample sizes needed to detect an effect. Researchers currently need to scour the, sometimes contradictory, literature to determine these figures.

Hypothesis 1: Collaborative inhibition will increase as group size increases.

The RSDH predicts that the disruptive effects of CI will increase as a function of collaborative group size (Basden et al., 2000). It is suggested that as the number of individualized retrieval strategies being used during collaboration increases, each group member's own preferred retrieval strategy becomes increasingly disrupted and the productivity of the group as a whole decreases (Basden et al., 2000). To date, there have only been two attempts to test this prediction. In one study, Basden et al. (2000) examined the free recall of collaborative and nominal dyads and tetrads. In the other, Thorley and Dewhurst (2007) examined the free recall of collaborative and nominal dyads, triads, and tetrads. In line with the RSDH, both studies found that CI was more pronounced in larger groups. Examining the effect of group size on CI in this analysis provides a useful additional test of this prediction.

Although the limited evidence available suggests CI increases with group size, the smallest group size required to reliably show this effect is not clear. CI has been reliability found in triads (Basden et al., 1997; Blumen & Rajaram, 2008; Congleton & Rajaram, 2011; Pereira-Pasarin & Rajaram, 2011; Weldon & Bellinger, 1997), but it has not been reliability found in dyads, with some studies finding evidence of CI (Andersson & Rönnberg, 1996; Finlay et al., 2000; Thorley & Dewhurst, 2007) and others not (Andersson & Rönnberg, 1995; Basden et al., 2000; Meudell, Hitch, & Boyle, 1995; Meudell, Hitch, & Kirby, 1992). This meta-analysis can help clarify this uncertainty in the literature by establishing whether or not dyads are vulnerable to CI. If they are not, then this demonstrates a possible boundary for the effect. From a pragmatic point of view, it is useful to know if an effect can be reliably obtained

in pairs as it would be more economical for future researchers investigating CI to conduct their studies with dyads than with larger groups.

Hypothesis 2: Study material consisting of story-like information and categorized items will decrease collaborative inhibition compared to uncategorized items.

CI has been reported in studies using several different kinds of study materials (photos, e.g., Ross, Spencer, Blatz, & Restorick, 2008; word lists, e.g., Takahashi, 2007; short stories, e.g., Weldon & Bellinger, 1997, Exp. 2; etc.); however, it is expected that CI would be found to a greater extent in studies using materials that can be organized in many different ways. Lists of words and series of pictures are common study material in collaborative remembering studies. These words or pictures sometimes form one or more categories (e.g., food items, bathroom items, activities, etc.) or sometimes all of the items to be remembered are unrelated to one another. Categorized items already have an underlying organizational structure, and this is true to an even greater extent for short stories and story-like scenarios: Thus, there is a limited way in which they can be organized and encoded by each collaborative group member. In comparison, items from uncategorized lists could theoretically be organized in many more individualized ways. According to the RSDH, it is more likely that each group members' retrieval strategy will be more severely disrupted by others when these strategies are more variable. It was therefore predicted that collaboration would have a stronger inhibitory effect on memory when the material to be remembered consisted of uncategorized items as opposed to either story-like information or categorized items. In line with this, Andersson and Rönnberg (1995) reported stronger inhibition effects when collaborative and nominal pairs recalled unrelated words than when they recalled a short story. These findings, however, are confounded as memory of the unrelated words was assessed via free recall and memory of the story via cued recall. Some researchers have found free recall tests produce CI but cued-recall tests do not (e.g., Finlay et al., 2000, Exp. 2). Consequently, it may be the case that the differences observed resulted from the different memory tests used for each and not the study materials. This is the only study we are aware of that directly compared these two different types of stimuli so including this variable in the meta-analysis provides a useful test of a prediction made by the RSDH.

Hypothesis 3: Larger category sizes will produce greater collaborative inhibition.

When individuals study and rehearse material, they tend to naturally organize the information into categorical clusters (Rundus, 1971). Along the same theoretical lines as Hypothesis 2, the more items there are in a category of to-be-remembered study material, the more ways in which group members can organize the material. For example, in a small category of five items, the items can be associated to one another and organized in fewer ways than the items from a large category of 15 items. The RSDH therefore predicts that group members' strategies are more likely to be aligned with one another, and less likely to disrupt one other, when the items to be remembered form small categories than when they form large categories. Support for this hypothesis has been demonstrated by Basden et al. (1997), who directly manipulated the category size of to-be-remembered word lists in a collaborative remembering experiment. As only one study has directly examined this issue to date, and the evidence from this study is widely used to support the RSDH, including this moderator in the meta-analysis will provide a much-needed additional test of this prediction.

Hypothesis 4: A turn-taking procedure during the retrieval phase will lead to greater collaborative inhibition than a free-flowing procedure.

Two main procedures are used during the collaborative test phase: In the turn-taking procedure, each member of the group retrieves one item at a time and then waits until all other group members have had a chance to respond before contributing another item (e.g., Basden et al., 1997; Wright & Klumpp, 2004). In the free-flowing procedure, all members of the group, in no particular order, may retrieve as many items as they can (e.g., Barber & Rajaram, 2011a, 2011b; Reysen, Talbert, Dominko, Jones, & Kelley, 2011). The RSDH would predict that group members' individual retrieval strategies are more disrupted in the turn-taking procedure, because everyone's strategy will be disrupted after every single item contributed by the group. In the free-flowing procedure, on the other hand, it is possible for one member to retrieve many items sequentially, thus making more use of their own retrieval strategies. To date, only two studies have directly examined this issue and the results are inconsistent with the predictions of the RSDH. Thorley and Dewhurst (2007) and Harris, Barnier, and Sutton (2012, Exp. 2) found that both turn-taking and free-flowing groups suffer from CI and there was little difference between the two groups' performance. Including this moderator in the meta-analysis will therefore provide a more robust test of this prediction.

Hypothesis 5: Free-order memory tasks will produce greater collaborative inhibition than forced-order memory tasks.

It is possible that individual retrieval strategies are most likely to be relied upon during memory tasks in which individuals are free to recall items in any order as opposed to during tasks in which the order of the items to be recalled is in some way imposed upon participants. In cued recall and recognition tasks, the order of the cues or of the items to be judged as old or new is decided by the experimenter (or are presented randomly), and this order is likely inconsistent with participants' retrieval strategies (Finlay et al., 2000). Because retrieval strategies are disrupted for both nominal and collaborative group members, collaboration is expected to have less of an impact on memory performance in studies using such memory tasks. In studies using free-recall tasks, however, collaborative group members would experience retrieval strategy disruption from exposure to others' recall whereas nominal group members would not experience any retrieval strategy disruption. Some empirical evidence supports this notion (e.g., Barber et al., 2010; Finlay et al., 2000, Exp. 2; Thorley & Dewhurst, 2009), however, other studies have observed CI in cued recall and recognition tests (e.g., Andersson & Rönnberg, 1996; Kelley et al., 2012; Meade & Roediger, 2009). As discussed in the previous section, the presence of the effect in forced-order tests (e.g., cued recall and recognition tests) provides evidence suggesting other mechanisms contribute to the CI effect. Despite this, the RSDH

would nevertheless predict stronger CI effects with free-order tests. It was therefore hypothesized that the CI effect would be weakest in studies using forced-order tasks and strongest in studies using free-order tasks.

Hypothesis 6: A higher number of study and test phases will lead to decreased collaborative inhibition.

Interitem association refers to the strength of the association formed during encoding between the items included within the to-be-remembered material (Basden, Basden, & Stephens, 2002). Several studies have shown that repeated study as well as repeated testing tends to increase interitem association (Blumen & Rajaram, 2009; Congleton & Rajaram, 2011; Pereira-Pasarin & Rajaram, 2011). It is thought that increased interitem association decreases the effect of interference on retrieval (Bäuml & Aslan, 2006; Smith, Adams, & Schorr, 1978). Accordingly, increasing the number of study or test phases should decrease the degree of CI because this should reinforce subjects' individual retrieval strategies and make them less susceptible to disruption from other group members' strategies. The RSDH would therefore predict that CI would be strongest when the retrieval phase followed a single study-test cycle, and would decrease as additional study and/or retrieval phases preceded the final retrieval phase. Pereira-Pasarin and Rajaram (2011, Exp. 1) manipulated how often participants studied a list of words. As expected, CI was present when the items were studied once, but was significantly reduced when the items were presented three times. Similarly, Basden et al. (2000, Exp. 2) found a CI effect following a single study-test trial, but reported an absence of CI when participants completed two or three study-test trials prior to the critical collaborative/nominal test trial. However, Congleton and Rajaram (2014) failed to observe a decreased CI effect when two individual test phases preceded the critical test trial compared to when only one individual test phase preceded it. Thus, some contradictory findings exist in the literature, which the testing of this moderator could help clarify.

Hypothesis 7: Incidental encoding will lead to a decrease in collaborative inhibition in comparison to intentional encoding.

Individualized retrieval strategies are believed to be based on the unique way in which participants organize study material *during encoding* and also, in part, on their expectations of the retrieval context (Basden & Basden, 1995; Rajaram, 2011). If so, then the RSDH should predict that the participants' knowledge of an upcoming memory test will impact the strength of the CI effect. Participants who are unaware of an impending test (and thus cannot have any retrieval context expectations) should be less likely to strategically organize the material to which they are exposed, leading to a decrease in CI. Participants who expect a future memory test, however, should be more likely to engage in organized encoding—and later retrieval—of the test material, thus creating a situation where they are susceptible to CI. To date, this moderator has not been directly tested.

Hypothesis 8: Collaborative inhibition will decrease when collaborative group members know one another well.

Rajaram and Pereira-Pasarin (2010) suggest CI may be reduced in collaborative groups of close acquaintances who have a transactive memory. A transactive memory is a shared memory system for encoding, storing, and retrieving information (Wegner, 1986, 1987; Wegner, Giuliano, & Hertel, 1985). Transactive memories develop over time as acquaintances learn about each other's expertise in different knowledge domains. If each knows each other's areas of expertise, they can explicitly (e.g., through negotiated agreement) or implicitly divide responsibility for remembering nonoverlapping elements of new information related to their area of expertise (Wegner, 1987). In an efficient transactive memory system, acquaintances can later access this nonoverlapping information via cross-cuing so they recall more together than they would alone (Wegner et al., 1985).

Wegner, Erber, and Raymond (1991) first demonstrated that close acquaintances can implicitly divide responsibility for learning new information based on their understanding of each other's expertise. They had dyads, composed of dating couples or strangers, study words from six different knowledge domains. No communication was allowed during encoding to prevent the dyads from explicitly dividing responsibility for remembering different domains. When later questioned, dating couples agreed more about the relative expertise of each partner, showed less overlap in their individual recall, and recalled more words overall than strangers (see also Hollingshead, 1998). Several researchers have since examined whether or not the collaborative advantages observed in couples can eliminate CI. To date, the evidence for this is mixed. Some researchers have found a decrease in CI within friends and spouses compared with strangers (e.g., Andersson, 2001; Andersson & Rönnberg, 1995, 1996; Johansson, Andersson, & Rönnberg, 2000) whereas others have not (e.g., Harris et al., 2013; Peker & Tekcan, 2009).

As stated, transactive memory theorists suggest close acquaintances' retrieval is enhanced as they recall nonoverlapping information. In the collaborative remembering literature, it has also been observed that CI is abolished in stranger dyads that are forced to recall nonoverlapping items from categorized word lists (Basden et al., 1997, Exp. 3). Basden et al. (1997) suggest this occurs as participants' own retrieval strategies are unaffected by hearing others recall different items using different strategies. It is therefore possible that CI was reduced in couples in the aforementioned studies because they recalled nonoverlapping information and did not interfere with each other's retrieval strategies.

Due to the conflicting literature, it remains unclear as to whether or not collaborative acquaintances are generally less susceptible to CI than collaborative strangers. The current meta-analysis will therefore examine this issue. Such an effect would be predicted by the RSDH as a result of acquaintances recalling nonoverlapping information and not disrupting each other's retrieval strategies. At this juncture, however, we acknowledge that other mechanisms may also play a role in reducing CI in close acquaintances. For example, Harris, Keil, Sutton, Barnier, and McIlwain (2011) found that communication styles developed in couples that promote successful cross-cuing, relevant elaborations in response to those cues, and mirroring of speech can produce collaborative facilitation. Due to the lack of research on communication styles in collaborative groups, however, it is not possible to assess their impact on CI here.

Exploratory Moderator: Collaborative Inhibition May be Decreased as the Similarity Between the Encoding and Retrieval Contexts Increases

In the traditional CI paradigm, all participants initially encode material individually before retrieving the material either individually (nominal group members) or in a group (collaborative group members). As has been discussed in the previous section on alternative mechanisms for the CI effect, some have argued that the effect found with this traditional paradigm is due to the match between the encoding and retrieval contexts of nominal groups (i.e., the effect really reflects a nominal advantage), as would predict encoding specificity principles (e.g., Tulving & Thomson, 1973). In particular, the concept of transfer-appropriate processing (TAP, e.g., Morris, Bransford, & Franks, 1977) posits that memory performance will be maximized when the cognitive processes involved at encoding match those involved at retrieval (i.e., when the level of interaction between group members is similar), and context-dependent learning postulates that performance will be maximized when the physical context at encoding matches that of the context at retrieval (i.e., when group composition is similar). A few experiments deviating from the traditional paradigm have been described in which either participant interaction at encoding was manipulated or participants encoded in a group setting (with or without interaction). When all participants initially encode material individually, the predictions of the RSDH are the same as those made by the encoding specificity principle: When individuals encode alone, collaborative retrieval will reduce output compared with nominal retrieval because individual members disrupt each other's retrieval strategies, or because the physical context and the cognitive processes they must use during collaborative retrieval do not match those at encoding. However, when individuals initially encode material collaboratively (resulting in the development of a group encoding strategy), the two mechanisms predict different outcomes: The RSDH would predict that both nominal and collaborative group members could use the group encoding strategy at retrieval, resulting in equal performances for both groups and an elimination of CI. Encoding specificity, on the other hand, would predict that the context and cognitive processes involved for nominal group members differ at encoding and retrieval (as the collaborative element is no longer present at retrieval). In this case, nominal group members would be at a disadvantage, resulting in a collaborative advantage. Although no individual study to date has reported a collaborative retrieval advantage following collaborative encoding, some groups of researchers have observed an elimination of CI following collaborative encoding (Andersson & Rönnberg, 1997; Barber et al., 2012; Finlay et al., 2000). This synthesis, however, may be better powered to demonstrate such an effect. This exploratory analysis will compare the presence and magnitude of CI in studies with varying degrees of consistency in the cognitive processes engaged at encoding and retrieval (individual vs. collaborative interaction), and the physical contexts created at encoding and retrieval (lone vs. group settings).

Postcollaborative Memory

Some researchers have examined not only whether collaboration can affect group remembering, but also whether collaboration can have an impact on later individual memory (e.g., Barber & Rajaram, 2011a, 2011b; Blumen & Stern, 2011; Harris et al., 2012; Harris et al., 2013; Weldon & Bellinger, 1997). In collaborative remembering studies, participants are sometimes asked to retrieve information twice (or more): once in nominal or collaborative groups, and again individually. Thus, members of nominal groups retrieve the study material on their own twice whereas members of collaborative groups retrieve the material first in a group context and again individually. Memory performance on this later individual retrieval can thus be compared between previous collaborative and nominal group members to determine the lasting effects, if any, of collaboration on individual remembering.

Following CI, three possible outcomes can result from postcollaborative retrieval. The first outcome is that the negative impact of collaboration persists, resulting in previous collaborative group members remembering less than previous nominal group members. This *forgetting effect* occurs when individuals fail to retrieve items on a postcollaborative test that were previously retrieved during collaborative remembering (e.g., Blumen & Rajaram, 2008; Congleton & Rajaram, 2011). Another type of forgetting can happen when the items that were not retrieved during collaboration are also not retrieved during later individual recall. This phenomenon has been attributed to *socially shared retrieval-induced forgetting* (Rajaram, 2011), which posits that items retrieved by one person during collaboration are also covertly retrieved by all other group members, leading to the inhibition of other related (vs. unrelated) items that are yet to be retrieved by the group (e.g., Coman, Manier, & Hirst, 2009; Cuc, Koppel, & Hirst, 2007). This outcome is also compatible with retrieval inhibition (discussed in the previous section on alternative cognitive mechanisms of CI), which is assumed to occur when cue words permanently suppress the retrieval of target words, making them unavailable for later retrieval (e.g., Anderson et al., 1994; Bäuml & Aslan, 2004).

The second possible outcome in postcollaborative retrieval is that the average memory performance of former collaborative group members is no better or worse than the average memory performance of former nominal group members. This *rebound effect* occurs when individuals remember items that were not retrieved during collaboration due to a release from the negative impact of collaboration on individual retrieval strategies (Rajaram & Pereira-Pasarin, 2010). In this case, the RSDH posits that although group members initially encoded the items, these were not retrieved during collaboration due to group members' disrupted retrieval strategies. However, during postcollaborative remembering, individuals are released from this disruption, and are able to revert back to their individual retrieval strategy. Consequently, their memory performance on this subsequent test is equal to nominal group members who never engaged in collaborative remembering. A rebound effect is akin to a *release effect* in part-list cuing inhibition that is observed when participants' memory of the target (i.e., noncued) items improves on a subsequent test in which the cue items are no longer present (e.g., Basden & Basden, 1995).

The third possible outcome after collaborative retrieval is that a *postcollaborative advantage effect* is observed. That is, the average memory performance of former collaborative group members is better than the average memory performance of former nominal group members. It is common to observe an increase in memory performance after repeated testing (Payne, 1987), but in this case

the former collaborative group members experience a benefit over and beyond this effect of repeated testing. This superior performance can be attributed to a release from CI (i.e., a rebound effect) in addition to the beneficial effect of reexposure to study items from other group members' retrieval during previous collaboration (e.g., Blumen & Rajaram, 2008), to which of course previous nominal group members were not exposed. In order for a postcollaborative advantage to be observed, the beneficial effect of reexposure must be stronger than any detrimental effect of forgetting.

Significant net effects of persistent collaborative impairment (i.e., forgetting) have seldom been reported (but see Barber et al., 2015); however, greater forgetting among collaborative group members than among nominal group members has sometimes been observed (e.g., Basden et al., 2000; Finlay et al., 2000; Henkel & Rajaram, 2011). Further, whereas several studies have shown equal performances between previous collaborative group members and previous nominal group members (e.g., Finlay et al., 2000, Exp. 1 and 3; Meade & Roediger, 2009; Wright & Klumpp, 2004), several others have shown clear postcollaborative advantage effects (e.g., Barber & Rajaram, 2011a, 2011b; Blumen & Stern, 2011; Choi et al., 2014; Stephenson & Wagner, 1989; Weldon & Bellinger, 1997). Finally, some researchers have reported mixed results, in which postcollaborative advantage effects may depend on certain methodological elements, such as group size (Basden et al., 2000; Thorley, 2007; Thorley & Dewhurst, 2007), encoding procedure (Congleton & Rajaram, 2011; Harris et al., 2013), testing procedure (Finlay et al., 2000, Exp. 2; Thorley & Dewhurst, 2007), and group member relationship[1] (Harris et al., 2013; Peker & Tekcan, 2009).

In summary, individuals who are asked to retrieve study material following collaborative retrieval can experience either continued memory inhibition (i.e., forgetting), a release from CI (i.e., a rebound effect), or improved memory performance (i.e., postcollaborative advantage). The reason behind these contradictory outcomes of postcollaboration on individual memory is still not clear (Rajaram, 2011). In an attempt to address these seemingly contradictory findings, a second goal of this review is to conduct a secondary meta-analysis of the data on postcollaboration memory performance.

Hypothesis 9: Collaboration benefits later individual retrieval.

As stated above, the beneficial effect of collaboration on postcollaborative memory have not always been replicated (e.g., Finlay et al., 2000; Meade & Roediger, 2009; Wright & Klumpp, 2004); however, given that most research has found at least partial evidence for a postcollaborative advantage, it was hypothesized that former collaborative group members would perform better on an individual retrieval task than former nominal group members. This prediction is in line with the RSDH. If retrieval disruption is to blame for CI, then a subsequent individual retrieval task should enable former collaborative group members to be released from this disruption and revert back to their own, more efficient, retrieval strategies. In addition, because groups do outperform individuals in remembering tasks (Clark & Stephenson, 1989), it follows that, at the very least, former collaborative group members would be reexposed to more of the study items during previous retrieval events than would former nominal group members, allowing for greater relearning.

Hypothesis 10: A high level of interitem association among study items will lead to a larger rebound effect in postcollaborative memory.

One possible explanation for some of the contradictory findings found among postcollaborative remembering effects comes from the part-list cuing literature. As previously stated, the RSDH in collaborative recall is functionally equivalent to the retrieval strategy disruption explanation offered to explain part-list cuing inhibition (Basden & Basden, 1995). Just as a rebound effect is often found in collaborative remembering studies, a *release effect* has sometimes been observed in part-list cuing inhibition (e.g., Basden & Basden, 1995). However, Bäuml and Aslan (2006) have examined methodological differences between part-list cuing studies in which a release effect was found and studies in which it was not found. They proposed, and demonstrated through a series of experiments, that a release from strategy disruption was likely when there was a high degree of interitem association among study items, but that memory performance impairments were likely to persist when there was a low degree of interitem association among study items. Given the similarities between the findings and theoretical explanations of part-list cuing and collaborative remembering studies, it is possible that interitem association also has an impact on rebound and forgetting effects observed in postcollaborative memory performance.

Thus, study material that is high in interitem association might make participants immune to the potential long-lasting effects of CI (i.e., forgetting) on postcollaborative memory. Study designs that incorporate procedures that increase the interitem association of the study material may therefore be more likely to lead to a rebound effect, and a potential postcollaborative memory advantage, compared with study designs that do not encourage high interitem association. Items from categorized lists and ideas from story-like material are by nature more strongly associated to one another than are items from uncategorized lists. Also, interitem association is increased when participants study the material multiple times or when they retrieve the material multiple times, as each participant's individual retrieval strategy is reinforced (Basden et al., 2000; Congleton & Rajaram, 2011). Two studies from the CI literature provide indirect evidence of a possible effect. Blumen and Rajaram (2008) and Congleton and Rajaram (2011) found a greater collaborative advantage among group members who had more stable retrieval strategies prior to the critical CI trial (e.g., as measured by a higher degree of clustering in the items retrieved over several trials). They proposed that a more secure organization of the study material enabled collaborative group members to benefit from reexposure to a greater extent than those who had a less secure organization of the material, which in turn resulted in a postcollaborative advantage compared to previous nominal group members not exposed to others' retrieved items. No researcher, however, has ever directly tested the impact of interitem association (as measured by either type of study material or number of study and/or retrieval events) on postcollaborative memory, and thus this analysis will be the first to test this potential moderating effect. It was hypothesized that study designs that

[1] Note that this variable could not be tested as a moderator in our analysis due to a lack of variability among the studies meeting inclusion criteria.

include more procedures which increase interitem association (story-like material, categorized lists, repeated study, and repeated testing) will show greater postcollaborative advantage than studies that use less or none of these procedures.

Hypothesis 11: A free-flowing retrieval procedure will lead to greater rebound effects than a turn-taking retrieval procedure

Because the act of retrieving encoded information can strengthen a retrieval strategy, collaborative as well as nominal group retrieval is likely to increase the interitem association of study material formed by individual group members. However, within collaborative groups, a free-flowing procedure may be more likely to facilitate this reinforcement of interitem associations since it provides group members the opportunity to rehearse several study items at once during each turn. In a turn-taking procedure, on the other hand, group members are limited to retrieve one item at a time, with interfering items interspersed between each of their retrieval efforts. Thus, turn-taking retrieval is less likely to have a strengthening effect on interitem association, which in turn is less likely to promote a benefit of prior collaboration. Only two studies by Thorley (2007) manipulated interaction at retrieval and also measured postcollaborative remembering. The results obtained in these studies were somewhat mixed but generally contradicted expectations: A postcollaborative advantage was found in turn-taking group dyads, triads, and tetrads, but was only consistently observed in free-flowing triads and tetrads (not dyads). This moderator analysis will therefore provide an important further test of the hypothesis that a free-flowing procedure facilitates rebound (and a postcollaborative advantage) compared with a turn-taking procedure.

Exploratory Moderator: Group Size May Moderate the Effect of Collaboration on Later Individual Remembering

Although retrieval itself can strengthen a retrieval strategy (Congleton & Rajaram, 2011), according to the RSDH collaborative retrieval is likely to do so to a lesser extent than individual retrieval, since any one person's retrieval strategy is likely to interfere with another's. Having said that, retrieval in smaller collaborative groups (e.g., dyads) may be more effective at strengthening collaborative group members' individual retrieval strategies (and thereby increasing interitem association) than retrieval in larger collaborative groups (e.g., triads or quartets), because less disruption is likely to occur in the former (see Hypothesis 1). On the other hand, however, larger collaborative groups may reexpose individuals to a larger portion of the original material compared to smaller groups. The effect of group size on postcollaborative memory has been reported in two studies, with mixed results. Thorley (2007) reports a postcollaborative advantage among previous members of collaborative dyads, triads and quartets in most—but not all—of his four experiments, and Basden et al. (2000, Exp. 1) reports greater postcollaborative advantage among quartets than dyads. Thus, it is possible that a greater rebound effect in postcollaborative memory may be seen amongst smaller collaborative groups due to a greater strengthening of individuals' differing retrieval strategies, or amongst larger collaborative groups due to greater reexposure to the material. The size of the nominal and collaborative groups was thus included as an exploratory moderator variable.

Exploratory Moderator: Type of Postcollaborative Memory Test May Moderate the Effect of Collaboration on Later Individual Remembering

As discussed in a previous section, retrieval strategy disruption is only one, albeit the most cited, of several other proposed cognitive mechanisms behind CI. Specifically, Barber et al. (2015) described how *retrieval inhibition* and *retrieval blocking* could also contribute to CI effects. They tested the contribution of each of these three mechanisms using cleverly designed experiments in which they compared the performance of previous nominal and collaborative group members on individual free recall and recognition tests. They theorized that if the detrimental effect of collaboration disappeared on both subsequent individual free recall and recognition tests (i.e., rebound effects), then retrieval disruption likely caused CI. If, however, the detrimental effect of collaboration persisted on both subsequent individual free recall and recognition tests (i.e., forgetting effects), then retrieval inhibition was likely at play. Finally, if the detrimental effect of collaboration persisted on a later free recall test but disappeared on a later recognition test, then retrieval blocking contributed to CI. Thus, in addition to clarifying the contradictory findings of postcollaborative memory found within the CI literature, this secondary meta-analysis also has the potential to elucidate the relative contribution of each of these three mechanisms toward CI by comparing the postcollaborative effects in studies using free recall and recognition tests. The type of memory test used to measure individual postcollaborative memory was included as an exploratory moderator of the effect of collaboration on later individual remembering.

Method

Inclusion and Exclusion Criteria

The first meta-analysis included studies that compared the memory performance of collaborative groups with that of equal-sized nominal groups (CI effect analysis). Nominal group memory performance had to be measured by pooling the items remembered by group members working individually and disregarding redundant answers. This excluded studies that compared only individual and group performances (e.g., Coman et al., 2009; Stephenson, Knivetton, & Wagner, 1991) and studies that provided "nominal scores" consisting simply of the arithmetic means of groups of individual participants' scores (e.g., Stephenson & Wagner, 1989). A separate meta-analysis included studies that compared the memory performance of individuals subsequent to a collaborative remembering task to the memory performance of individuals subsequent to an individual remembering task (postcollaborative memory analysis). For such studies to be included, memory performance had to be measured for each individual in both groups, and the study material had to be the same for both the initial and the subsequent postcollaborative (or postindividual) memory tasks. Studies in which an effect size for CI could not be obtained could still potentially provide an effect size for postcollaborative memory. For both analyses, studies that measured memory of healthy individuals with free recall, cued recall, recognition, and visual and

spatial recognition tasks were included. However, the dependent measures of interest were limited to those that captured "recollection completeness" (Harris et al., 2012). This includes studies that reported the proportion, percentage, or absolute number of study material items that were correctly recalled (in cued recall, free recall, and reconstruction tasks), or the hit rate (in recognition tasks). "Items" to be remembered could include words, pictures, locations, abstract images or patterns, scenarios, and stories. Studies utilizing either intentional learning paradigms (when participants are explicitly told that their memory will be tested following material presentation, e.g., Andersson, 2001; Barber & Rajaram, 2011a) or incidental learning paradigms (when participants are not explicit told that their memory will be tested, e.g., Barber & Rajaram, 2011b; Harris et al., 2013) were included.

As a starting point, relevant studies were located by examining the online index PsycINFO using the following search: 'collaborative inhibition' OR 'collaborative memory' OR 'collaborative remembering' OR 'group memory' OR 'group remembering' OR 'joint memory' OR 'joint remembering.'. The reference sections of the retrieved articles, as well as the references of recent narrative reviews (Rajaram, 2011; Rajaram & Pereira-Pasarin, 2010) were then searched for additional studies matching the inclusion criteria. In addition, a search of the online archives of www.psychfiledrawer.org was conducted in an attempt to find any unpublished replication of CI and postcollaborative memory studies (none were found). Search efforts were terminated in May, 2014.

Within-subjects design studies that confounded group composition with the testing order of the same study material (e.g., all participants recalled List A nominally at Test 1, then recalled List A again collaboratively at Test 2) were excluded from the analysis (e.g., Blumen & Rajaram, 2009; Ekeocha & Brennan, 2008). The rationale for exclusion was that, for this given scenario, a collaborative advantage could be due to repeated testing and a collaborative disadvantage could be due to the increased elapsed time since study. Studies using a within-subjects design that avoided these confounds by using different study material for nominal and collaborative remembering tasks were also excluded if the correlations between participants' nominal and collaborative recollection completeness scores, information necessary to compute an effect size, were not reported and could not be obtained (e.g., Andersson, 2001; Johansson, Andersson, & Rönnberg, 2005).

Other studies that otherwise fit the inclusion and exclusion criteria were excluded due to methodological procedures that made it difficult to meaningfully compare their data with the other effect sizes included in the sample. For example, this included one study by Basden et al. (1997, Exp. 4) in which each group member was asked to recall a different subset of the study material. In this case, it is unlikely that the participants' individualized recall strategies could have been disrupted by others' responding since everyone was recalling a different list of items. Similarly, Wright and Klumpp (2004) included a condition in which collaborative group members could neither see nor hear other group members' responses. Although the results from these studies would have been useful in elucidating social factors that impact recollection completeness (e.g., social facilitation), this was not of interest to this synthesis.

Finally, 17 CI and six postcollaborative memory effect sizes were excluded because the necessary statistics were not provided in the research report, and either they had been published too long ago to reasonably expect authors to still have access to the data, or efforts to obtain the required information from the authors were not successful.

To ensure that all effect sizes in the final sample were independent, care was taken to avoid including more than one study that reported the same data (e.g., Pereira-Pasarin, 2007 Dissertation, Exp. 2, and Pereira-Pasarin & Rajaram, 2011, Exp. 1). The search yielded a final sample of 59 studies, published between 1989 and 2014, and five dissertation studies that met one or both sets of inclusion criteria. This sample consisted of 102 independent effect sizes (75 effect sizes relevant to CI and 27 effect sizes relevant to postcollaborative memory). References for the included studies are marked with an asterisk in the references section.

Calculation of Effect Sizes

Comprehensive Meta-Analysis (CMA) 2.0 software was used to perform all analyses. Standardized mean differences were calculated between (a) the recollection completeness score of collaborative and nominal groups, in the first meta-analysis, and (b) the recollection completeness score of participants who had previously recalled the material collaboratively and participants who had previously recalled the material on their own, in the second meta-analysis. When group means, standard deviations, and group sample sizes were reported in the studies, these statistics were entered into CMA. (If standard errors were provided but standard deviations were not, the latter was calculated according to the formula provided by Lipsey & Wilson (2001, p. 200). When such information was not available, *t*-values, group means and group sample sizes were used, if possible. If neither of these sets of statistics were available but the *F* values from a one-way or two-way ANOVA (all factors between-subjects), group means, and group sample sizes were available, the standardized mean difference was estimated using the formula provided by Lipsey and Wilson (2001, pp. 173–185), and the standardized mean difference standard error was calculated according to the formula provided by Borenstein et al. (2009, p. 27).

For the CI meta-analysis, effect sizes were calculated by subtracting the mean recollection completeness score of the nominal group from the mean recollection completeness score of the collaborative group, such that a negative effect size indicated lower performance by collaborative groups (i.e., CI) and a positive effect size indicated the opposite. In study designs using multiple test phases (i.e., when collaborative and nominal groups remembered the studied material more than once, e.g., Finlay et al., 2000, Exp. 1), the first test phase was chosen as the critical test from which memory scores were used to compute the effect sizes. For the postcollaborative meta-analysis, effect sizes were calculated by subtracting the mean recollection completeness score of participants who had previously remembered collaboratively from the mean recollection completeness score of participants who had previously remembered individually, such that a positive effect size indicated higher performance by participants who previously remembered collaboratively (i.e., postcollaborative advantage). In study designs where participants were individually tested multiple times after the critical collaborative or nominal testing phase (e.g., Blumen & Stern, 2011), data from the first individual test phase was used to compute the effect sizes.

A mixed-effects model was used for the analyses: A random-effects model was used to compute the summary effects as it is assumed that different true effect sizes exist among the included studies, due in part to variations in participants and study design characteristics. A fixed-effect model was used, however, to determine heterogeneity across subgroups (i.e., for moderator variable analyses, Borenstein et al., 2009). To increase the precision of the summary effect size estimate we computed weighted means of the summary effect sizes, where each study effect size was weighted by the inverse of its variance. *Q* statistics are reported as a measure of heterogeneity in true effects, for both summary effects and moderator variable analyses (Borenstein et al., 2009).

Coding of Moderator Variables

Eight potential moderator variables relating to study design characteristics (group size, category size, study material, interaction at retrieval, memory task, number of study/test phases, encoding task, and encoding context) and one potential moderator variable relating to participant characteristics (relationship of group members) were coded from the collaborative remembering studies. Four potential moderator variables relating to study design characteristics (interitem association, interaction at retrieval, group size, and type of memory test) were coded from the postcollaborative memory studies. The first author coded the variables for all studies, and a random subset (44%) was also coded by the second author to assess interrater reliability. Kappa's were computed for nominal variables and intraclass correlation coefficients were computed for ordinal and interval variables. Agreement ranged from moderate (relationship, $\kappa = .46$) to perfect (group size and category size, $ICC = 1.0$), with 10 out of 11 variables yielding substantial agreement (κs and ICCs $> .75$, Landis & Koch, 1977). All disagreements were resolved through discussion.

Group size. The number of individuals making up the collaborative and nominal groups was coded for each effect size. The collaborative and nominal groups were always of equal size for each comparison (i.e., effect size) included in the analyses. All selected effect sizes originated from studies that used groups of either two or three participants except for two (Basden et al., 2000, Exp. 1; Thorley, 2007, Exp. 2), which used groups of four. Group size was therefore dichotomized and coded as either (a) *two members per group* or (b) *three or four members per group*. Group size was coded in the same manner for CI and postcollaborative memory effect sizes.

Study material. The type of material to be remembered by members of collaborative and nominal groups was coded into three categories: (a) story-like information, (b) categorized items, and (c) uncategorized items. These categories were created to reflect the organization potential of the material. Story-like information included short stories and scenarios (e.g., Andersson, 2001; Takahashi & Saito, 2004), categorized items included lists of words and series of pictures which could somehow be related to one another and could belong to one or several categories (e.g., food, household items, behaviors, etc.), and uncategorized items included words, pictures or patterns that were unrelated to one another. One study was not coded for this variable and was excluded from the relevant analysis, as the material used could not clearly fit into any category (unrelated sentences, Kelley et al., 2012).

Category size. If the type of material to be remembered was coded as categorized items (as defined above), the number of target items per category was also coded. If the study material consisted of several categories with an unequal number of items, the average number of items per category was calculated. Forty-five out of the 75 CI effect sizes were coded as having categorized study material, and category size was reported for all but one of these. Category size was coded as a continuous variable and ranged from 4 to 30, with an average of 13.6 ($SD = 7.3$) items per category. Effect sizes for which study material was coded as uncategorized items or story-like information were not coded for category size.

Interaction at retrieval. The interaction of collaborative group members during the memory task was coded as either (a) *free-flowing* or (b) *turn-taking*. Studies were coded as using a free-flowing procedure if any member of the collaborative group could provide any answer at any time without any restrictions as to how many answers one person could provide at one time. If, on the other hand, group members were instructed to take turns either recalling items (in free and cued recall tasks) or making old/new judgments (in recognition tasks), the study was coded as employing a turn-taking procedure. Each group member could only recall one item or make one old/new judgment at a time.

Memory task. During the critical test phase, participants completed (a) a forced-order memory test, or (b) a free-order memory test. In free-order tests participants were simply asked to recall the presentation or location of as many items or ideas (from story-like material) as they could (i.e., free recall). Forced-order tests consisted of either cued recall tasks in which participants were provided with cues to help them recall the material (e.g., category names, Basden et al., 1997, Exp. 3; questions, Andersson, 2001, Exp. 2; cue items, Barber et al., 2010, Exp. 1), recognition tests in which participants were presented with target and nontarget items and asked to make a new/old decision, or spatial reconstructive tasks (e.g., rearranging pictures to match an original configuration, Andersson, 2001). One study was not coded for this variable and excluded from the relevant analysis as the memory task used did not clearly fit into either category (shopping task, Ross et al., 2004).

Number of study/test phases. The number of times the study material was presented to participants (study phases) and the number of times that participants recalled the material or made recognition judgments about the study material (test phases) *before* the critical test phase was coded for each effect size. The number of study phases was added to the number of precritical test phases to create this variable. Only precritical test phases where all participants (members of both nominal and collaborative groups) were tested individually were included in this count. Because this variable was intended to act as a measure of the strength of association made between items of the study material (with more study/test phases resulting in stronger associations), precritical collaborative test phases were not included because it was assumed that collaborative remembering would disrupt individual group members' retrieval strategies (Rajaram & Pereira-Pasarin, 2010) and therefore would *not* strengthen interitem association (Hypothesis 6). The number of test/study phases was coded as a continuous variable and ranged from 1 to 3 with a mode of 1.

Encoding task. The manner in which participants encoded the material to be remembered was coded as either (a) incidental or (b)

intentional. In incidental encoding tasks, participants were not asked the memorize the material and instead were given a deceptive rationale for their exposure to it, such as observing their social interactions after viewing an emotional video (Wessel, Zandstra, Hengeveld, & Moulds, 2015) or their performance on a sentence-formation task (Barber et al., 2012). Participants in studies using intentional encoding tasks, on the other hand, were explicitly informed of the impending memory test before their exposure to the material. Two studies were not coded for this variable and were excluded from the relevant analysis as the methodological detail was unclear (Andersson & Rönnberg, 1996, Exp. 1 and 2).

Relationship of group members. Members of collaborative groups were coded as being either (a) strangers or (b) nonstrangers. Nonstrangers included participants described as spouses, friends, and classmates. For studies in which the relationship of group members were not specified, but in which it was stated that group membership was randomly assigned, it was assumed that group members were strangers. When it was stated that participants forming collaborative groups signed up for their study session together (i.e., formed their own groups), it was assumed that they were friends and the relationship was coded as nonstrangers. Four studies were not coded for this variable (Finlay et al., 2000, Exp. 1; Meade et al., 2009; Takahashi & Saito, 2004, Exp. 1 and 2), as the relationship between group members varied or was ambiguous (some were friends and some were not). The relationship of members of nominal groups was not coded as these individuals did not interact with each other during testing phases nor were they aware that their recollection completeness score would be combined with other participants' scores.

Encoding context. The disparity of the match in cognitive and physical contexts during encoding and retrieval tasks between nominal and collaboration groups was coded by considering two characteristics of the encoding task: group interaction and group composition. Group interaction refers to whether or not participants interacted with their future retrieval group members (whether collaboratively or nominally) during the encoding task. For example, some groups were asked to agree on cues for target words (Andersson & Rönnberg, 1997), collaboratively create sentences containing target words (Barber et al., 2010, 2012), or collaboratively study, with a partner, words printed on a single stack of cue cards (Abbe, 2004). Group composition refers to whether or not participants completed the encoding task in the presence of the other group members, regardless of whether they interacted to complete the task. For example, some group members encoded material alone at their own computer cubicle (e.g., Reysen et al., 2011), whilst others were part of a group who sat together around a single computer monitor (e.g., Weldon et al., 2000). The studies included in the analyses fell into one of five categories: (a) Both nominal and collaborative group members encoded interactively with their group; (b) nominal group members encoded individually and collaborative group members encoded interactively with their group; (c) nominal group members encoded individually and collaborative group members encoded in noninteractive groups; (d) both nominal and collaborative group members encoded in noninteractive groups; and (e) both nominal and collaborative group members encoded individually. Eleven studies were not coded on this variable and excluded from relevant analyses because the encoding methodology used was unclear.

Interitem association. Several study design characteristics were considered to have an impact on the strength of the association between items or ideas of the study material. Story-like information and categorized items were assumed to increase interitem association strength, as was repeated study phases and repeated test phases (e.g., Congleton & Rajaram, 2011). Effect sizes from studies that used uncategorized items as study material and only a single study phase and a single (critical) test phase were coded as having "low" interitem association. Effect sizes from studies in which categorized items or story-like information was used *or* multiple study or precritical test phases were used were coded as having "moderate" interitem association. Effect sizes from studies in which categorized items or story-like information was used *and* multiple study or precritical test phases were used were coded as having "high" interitem association. However, because only two studies included in the postcollaborative remembering analyses were coded as having "low" interitem association, the "low" and "moderate" interitem association categories were combined for analyses.

Type of postcollaboration memory test. Each effect size included in the postcollaborative remembering analyses was coded for the type of memory test given to individual participants *after* the critical collaborative versus nominal memory test. The type of memory test was coded as either (a) free recall or (b) recognition, because only a comparison of these two types of tests was of interest for this moderator analysis. A test was categorized as free recall when participants were asked to recall the material in any order, without any cues (i.e., without word cues, category names, or worded questions), and categorized as recognition when participants were asked to make old/new judgments of presented items. Only two studies (Stephenson & Wagner, 1989, and Wright & Klumpp, 2004) employed a cued-recall test and these two effect sizes were not included in this analysis.

Studies With Multiple Effect Sizes

It was possible that more than one effect size could have been selected from a single study. However, in order to ensure that every effect size in the meta-analyses were independent from each other, only one per sample of participants was included. In studies that compared multiple collaborative groups to a single nominal group (e.g., collaborating friends vs. collaborating strangers vs. nominal group, Andersson & Rönnberg, 1996), one comparison was randomly selected and the respective effect size was included. When studies included subsets of participants (e.g., spouses and strangers, Abbe, 2004, Exp. 3) that each had their own nominal group for comparison, all relevant effect sizes were included because the data from each nominal and collaborative group was used in the calculation of only one effect size.

Where studies compared samples based on a methodological difference that was of interest to this synthesis (e.g., cued recall vs. free recall, Basden et al., 1997, Exp. 2), separate effect sizes were included if the variable of interest was measured or manipulated between-subjects (provided that each group had its own nominal comparison group, and that the appropriate statistics were available), such that each participant contributed to the calculation of only one effect size. If the variable of interest was a within-subjects factor, only one level of the factor was randomly selected and the respective effect size was included (e.g., one vs. three

study phases, Pereira-Pasarin & Rajaram, 2011, Exp. 1). In studies that compared samples based on a methodological difference that was *not* of interest in this study (e.g., immediate recall vs. 2-hr delay, Congleton & Rajaram, 2011), one effect size that compared all collaborative and nominal groups was included if appropriate statistics were provided, but multiple effect sizes were included if statistics for combined samples were not provided or could not be computed.

Results

Collaborative Inhibition Effect

Summary effect size. Forty-one research reports (64 studies and 75 independent effect sizes) were included in the meta-analysis examining the CI effect, with a total of 1,507 nominal groups and 1,575 collaborative groups compared across all studies. Figure 1 lists all samples included. The effect sizes (standardized difference in means) ranged from −2.12 to +0.98. Of these, the majority (66, or 88%) were negative, two (3%) were exactly zero, and only seven (9%) were positive. The overall mean effect size was −0.78, which was significantly different from zero, $z = -11.77$, $p < .001$; $CI_{0.95}$ [−0.91, −0.65]. As research reporting significant findings is more likely to be published than research reporting nonsignificant findings (Dickersin, 2005), the analysis was tested for the effect of publication bias using Duval and Tweedie's *trim and fill* procedure. This analysis estimates, based on the number and size of the effects in the meta-analysis, the number of unpublished, nonsignificant results that are likely to exist (see Duval & Tweedie, 2000, for a thorough description of this method). This procedure results in a more conservative adjusted estimate of the overall mean effect size, $d = -0.56$, $CI_{0.95}$ [−0.69, −0.42], by including 18 studies that would be missing if the asymmetry in the distribution of obtained effect sizes (around a central null effect) was due to a publication bias (see Figure 2 for funnel plot). This medium-size (Cohen, 1988) adjusted summary effect further suggests that collaboration has a negative effect on recollection completeness. There was significant heterogeneity in the sample of effect sizes, $Q = 216.03$, $p < .001$, $I^2 = 65.75$, and thus some of the variance observed in effect sizes could potentially be explained by moderating variables.

Moderator analyses. To determine whether categorical variables pertaining to study design and participant characteristics had moderating effects on CI, an overall standardized difference in means was obtained for each level of the variables and then

Study name	subgroup	Statistics for each study						Sample size		Weight
		Std diff in means	Std error	Lower limit	Upper limit	Z-value	p-value	Nominal groups	Collab. groups	Rel. weight
Abbe,2004_3	Couples	-0.32	0.28	-0.87	0.23	-1.15	0.25	26	26	1.55
Abbe,2004_3	Strangers	-1.36	0.31	-1.97	-0.76	-4.43	0.00	26	26	1.46
Andersson&Ronnberg,1996_1		-0.61	0.32	-1.25	0.02	-1.89	0.06	20	20	1.41
Andersson&Ronnberg,1996_2		-1.34	0.39	-2.11	-0.58	-3.43	0.00	16	16	1.22
Andersson&Ronnberg,1997_1	Group encoding	0.16	0.35	-0.54	0.85	0.44	0.66	16	16	1.33
Andersson&Ronnberg,1997_1	Ind. encoding	-1.43	0.40	-2.20	-0.65	-3.60	0.00	16	16	1.21
Barber&Rajaram,2011A_2		-0.86	0.26	-1.38	-0.35	-3.31	0.00	32	32	1.60
Barber&Rajaram,2011B		-0.60	0.20	-1.00	-0.20	-2.94	0.00	50	50	1.78
Barber et al.,2012	Group encoding	-0.37	0.37	-1.09	0.35	-1.01	0.31	15	15	1.29
Barber et al.,2012	Ind. encoding	-1.87	0.44	-2.72	-1.01	-4.27	0.00	15	15	1.10
Barber et al.2010_1	Group encoding	0.00	0.25	-0.49	0.49	0.00	1.00	24	48	1.64
Barber et al.,2010_1	Ind. encoding	-0.15	0.29	-0.71	0.42	-0.51	0.61	24	24	1.52
Basden et al.,1998_1		-0.80	0.31	-1.41	-0.19	-2.57	0.01	21	24	1.45
Basden et al.,1998_2		-0.18	0.25	-0.66	0.31	-0.71	0.48	32	33	1.64
Basden et al.,2000_1		-0.83	0.33	-1.48	-0.18	-2.52	0.01	20	20	1.40
Basden et al.,1997_1		-0.74	0.27	-1.27	-0.21	-2.75	0.01	29	30	1.58
Basden et al.,1997_2	Cued recall	-1.95	0.43	-2.79	-1.11	-4.54	0.00	15	17	1.12
Basden et al.,1997_2	Free recall	-0.43	0.36	-1.14	0.28	-1.18	0.24	15	16	1.30
Basden et al.,1997_3		-1.16	0.34	-1.82	-0.50	-3.44	0.00	20	21	1.37
Basden et al.,1997_4		0.30	0.43	-0.54	1.14	0.70	0.48	11	11	1.13
Blumen&Rajaram,2008	ICI vs. CCI	-1.38	0.39	-2.15	-0.60	-3.50	0.00	16	16	1.22
Blumen&Rajaram,2008	III vs. CII	-1.41	0.39	-2.18	-0.63	-3.56	0.00	16	16	1.21
Blumen&Stern,2011	Older adults	-0.98	0.47	-1.91	-0.06	-2.08	0.04	10	10	1.02
Blumen&Stern,2011	Young adults	-0.92	0.47	-1.85	0.00	-1.96	0.05	10	10	1.03
Choi et al.,2014		-1.82	0.34	-2.49	-1.16	-5.40	0.00	18	36	1.37
Congleton&Rajaram,2011	Repeated Study	-1.39	0.32	-2.02	-0.76	-4.32	0.00	24	24	1.42
Congleton&Rajaram,2011	Repeated Test	-0.36	0.29	-0.93	0.21	-1.24	0.22	24	24	1.51
Congleton&Rajaram,2014		-0.93	0.43	-1.77	-0.09	-2.17	0.03	12	12	1.12
Finlay et al.,2000_1		-1.32	0.45	-2.20	-0.44	-2.93	0.00	12	12	1.07
Finlay et al.,2000_3		-1.89	0.44	-2.75	-1.03	-4.30	0.00	15	15	1.10
Garcia-Marques et al.,2012_2		-0.84	0.32	-1.47	-0.22	-2.64	0.01	21	22	1.43
Garrido et al.,2012		-1.33	0.45	-2.21	-0.45	-2.95	0.00	12	12	1.07
Harris et al.,2012		-1.00	0.39	-1.76	-0.24	-2.59	0.01	15	15	1.23
Harris et al.,2013_1	Group encoding	-0.28	0.45	-1.16	0.60	-0.63	0.53	10	10	1.08
Harris et al.,2013_1	Ind. encoding	-1.77	0.49	-2.73	-0.80	-3.57	0.00	13	10	0.97
Harris et al.,2013_2	Group encoding	0.00	0.45	-0.88	0.88	0.00	1.00	10	10	1.08
Harris et al.,2013_2	Ind. encoding	-1.32	0.49	-2.29	-0.35	-2.68	0.01	10	10	0.97
Harris et al.,2011		-0.31	0.29	-0.88	0.25	-1.10	0.27			1.53

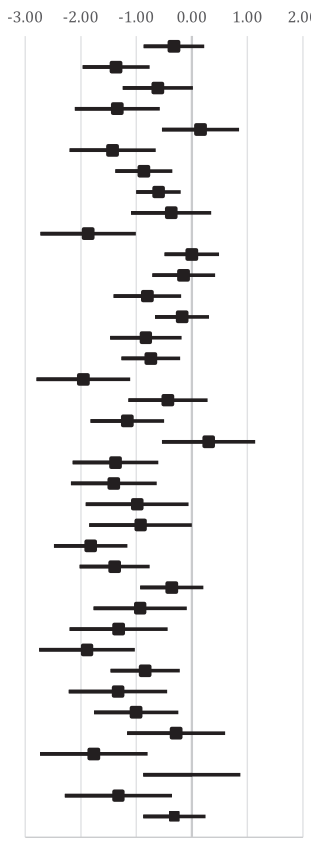

Figure 1. Forrest plot of effect sizes included in collaborative inhibition meta-analysis. The number following the publication year refers to the study number (e.g., "1" refers to Study 1, etc.). The summary effect values appear on the last row (values in bold).

Study name	subgroup	Std diff in means	Std error	Lower limit	Upper limit	Z-value	p-value	Nominal groups	Collab. groups	Rel. weight
Henkel&Rajaram,2011	Older adults	-0.75	0.37	-1.47	-0.03	-2.05	0.04	16	16	1.29
Henkel&Rajaram,2011	Young adults	-0.75	0.37	-1.47	-0.03	-2.05	0.04	16	16	1.29
Hyman et al.,2013_1		-0.72	0.25	-1.22	-0.23	-2.86	0.00	32	35	1.63
Hyman et al.,2013_2		-0.70	0.20	-1.08	-0.31	-3.53	0.00	55	54	1.80
Hyman et al.,2013_3		-0.02	0.21	-0.43	0.40	-0.07	0.94	41	48	1.75
Kelley et al.,2012		-0.57	0.28	-1.11	-0.03	-2.08	0.04	24	32	1.56
Kelley et al.,2014_1		-0.91	0.25	-1.39	-0.42	-3.67	0.00	36	36	1.65
Kelley et al.,2014_2		-1.15	0.27	-1.67	-0.63	-4.31	0.00	30	36	1.59
Kelley et al.,2014_3		-0.59	0.30	-1.18	-0.01	-1.98	0.05	28	20	1.49
Meade&Roediger,2009_2		-0.89	0.34	-1.55	-0.24	-2.66	0.01	20	19	1.38
Meade et al.,2009		0.07	0.20	-0.33	0.47	0.36	0.72	48	48	1.78
Pareira-Pasarin&Rajaram,2011_1		-0.63	0.36	-1.34	0.08	-1.74	0.08	16	16	1.30
Pareira-Pasarin&Rajaram,2011_2		-0.75	0.26	-1.26	-0.24	-2.90	0.00	32	32	1.61
Pareira-Pasarin,2007	Full attention	-1.10	0.38	-1.84	-0.36	-2.90	0.00	16	16	1.25
Peker&Tekcan,2009		-1.24	0.41	-2.03	-0.44	-3.05	0.00	14	15	1.18
Reysen et al.,2011_1		-0.57	0.32	-1.20	0.06	-1.76	0.08	20	20	1.42
Reysen et al.,2011_2		-0.88	0.37	-1.61	-0.15	-2.38	0.02	16	16	1.28
Reysen et al.,2011_3		-0.68	0.33	-1.31	-0.04	-2.08	0.04	20	20	1.41
Ross et al.,2004		-0.82	0.27	-1.35	-0.29	-3.02	0.00	30	29	1.57
Ross et al.,2008		-0.56	0.25	-1.06	-0.06	-2.18	0.03	31	33	1.62
Takahashi&Saito,2004_1		-0.95	0.47	-1.88	-0.03	-2.02	0.04	10	10	1.02
Takahashi&Saito,2004_2		0.98	0.50	0.00	1.96	1.97	0.05	9	9	0.96
Takahashi,2007_1		0.20	0.25	-0.28	0.68	0.81	0.42	33	33	1.65
Takahashi,2007_2		-0.65	0.31	-1.26	-0.04	-2.10	0.04	22	22	1.46
Takahashi,2007_3		-1.40	0.35	-2.10	-0.71	-3.98	0.00	20	20	1.33
Thorley&Dewhurst,2007		-1.60	0.47	-2.52	-0.68	-3.41	0.00	12	12	1.03
Thorley,2007_2		-1.56	0.47	-2.48	-0.65	-3.35	0.00	12	12	1.04
Thorley,2007_3		0.14	0.41	-0.66	0.94	0.35	0.73	12	12	1.18
Thorley,2007_4		0.55	0.42	-0.27	1.36	1.32	0.19	12	12	1.16
Weigold et al.,2014		-1.65	0.38	-2.38	-0.91	-4.39	0.00	19	19	1.27
Weldon et al.,2000_1		-0.26	0.26	-0.77	0.25	-0.99	0.32	29	30	1.60
Weldon et al.,2000_2		-1.14	0.48	-2.08	-0.19	-2.36	0.02	10	10	1.00
Weldon et al.,2000_3		-0.88	0.38	-1.63	-0.13	-2.31	0.02	15	15	1.25
Weldon et al.,2000_4		-0.33	0.32	-0.95	0.30	-1.03	0.30	20	20	1.43
Weldon et al.,2000_5		-0.89	0.42	-1.71	-0.07	-2.12	0.03	12	13	1.15
Wessel et al.,2014		-2.12	0.41	-2.93	-1.31	-5.16	0.00	18	19	1.17
Wright&Klumpp,2004		-1.61	0.51	-2.61	-0.60	-3.12	0.00	10	10	0.93
		-0.78	**0.07**	**-0.91**	**-0.65**	**-11.77**	**0.00**			

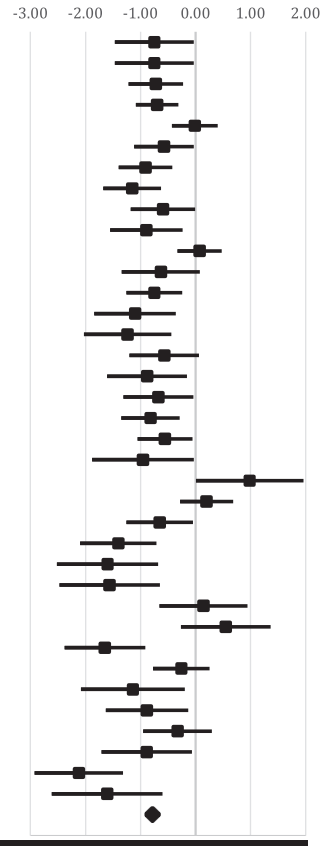

Figure 1. (continued).

compared in an analysis akin to the ANOVA in primary studies. To determine whether continuous potential moderator variables significantly predicted the strength of CI, such variables were individually entered into a metaregression analysis (akin to a regression analysis in primary studies). Table 1 lists all moderator variables with relevant statistics.

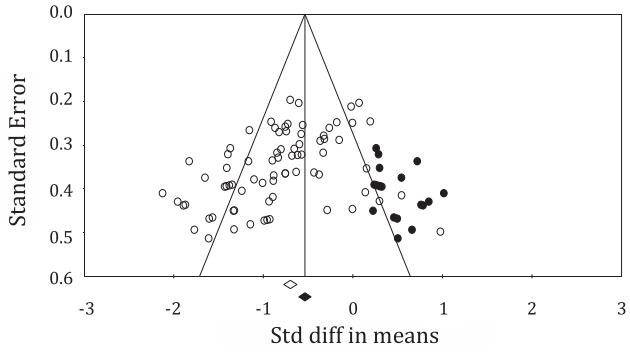

Figure 2. Duval and Tweedie Trim and Fill Imputed Funnel Plot. Clear circles are those studies that were included in the meta-analysis. Filled circles represent imputed studies that are thought to be missing due to publication bias (i.e., due to null or reversed effects). The clear diamond represents the original effect size, while the filled diamond represents the adjusted effect size. The widths of the diamonds represent variance.

Five of the eight variables related to study design had significant or trending moderating effects on CI. Hypothesis 1 was supported: Both groups of two and groups of three or four showed significant CI ($ps < .001$), but the mean CI effect was larger for groups of three or four, $d = -0.86$, than for groups of two, $d = -0.55$, $Q = 17.08$, $p < .001$. Hypothesis 5 was also supported, as the CI effect was stronger in studies that used a free-order memory task (e.g., free recall), $d = -0.76$, than in studies using a memory task that forced an order on the material to be remembered (e.g., recognition test), $d = -0.40$, $Q = 12.77$, $p < .001$ (although once again both types of memory tasks produced significant CI, $ps < .001$). Significant CI was also found in studies using story-like, categorized, and uncategorized study material ($ps < .001$), but the type of material to be remembered moderated the CI effect, $Q = 5.53$, $p = .063$. Although this omnibus analysis did not quite reach the traditional significance level, follow-up analyses were nevertheless computed using a Bonferonni-corrected alpha level for the two predicted two-level comparisons: Studies that used uncategorized material yielded a stronger average CI effect, $d = -0.80$, than studies that used story-like material, $d = -0.47$, $Q = 5.53$, $p = .019$. No difference was observed between uncategorized and categorized material, $d = -0.70$, $p = .26$. Thus, Hypothesis 2 was

Table 1
Analysis of Potential Moderator Variables of Collaborative Inhibition Effect

Variable	Number of effect sizes	Mean effect size (*d*)	Regression slope estimate	95% confidence interval	*Q*-test
Group size					17.08**
2	34	−.55		−.65, −.45	
3/4	41	−.86		−.97, −.76	
Study material					5.53*
Story-like	8	−.47		−.70, −.25	
Categorized items	45	−.70		−.79, −.61	
Uncategorized items	21	−.80		−.94, −.65	
Category size	44		−.004	−.01, .02	.26
Interaction at retrieval					2.77*
Free-flowing	60	−.67		−.75, −.59	
Turn-taking	15	−.84		−1.01, −.66	
Memory task					12.77**
Free-order	61	−.76		−.84, −.68	
Forced-order	13	−.40		−.58, −.22	
Number of study/test phases	74		.04	−.10, .17	.30
Encoding task					2.30
Incidental	22	−.79		−.94, −.64	
Intentional	51	−.66		−.75, −.57	
Encoding context					16.55**
Group 1	6	−.26		−.52, .01	
Group 2	3	−.73		−1.06, −.39	
Group 3	25	−.67		−.80, −.54	
Group 4	9	−.69		−.91, −.46	
Group 5	19	−.88		−1.03, −.73	
Relationship					8.45**
Strangers	63	−.77		−.85, −.69	
Nonstrangers	8	−.44		−.65, −.22	

* $p < .10$. ** $p < .05$.

partially supported. Finally, in regards to Hypothesis 4, a trend emerged wherein a turn-taking interaction between collaborative group members yielded a stronger average CI effect, $d = -0.84$, compared with a free-flowing interaction, $d = -0.67$, $Q = 2.77$, $p = .096$ (again both types of retrieval interaction yielded significant CI, $ps < .001$). Category size ($p = .62$), number of study/test phases ($p = .58$), and encoding task ($p = .13$) did not significantly moderate the CI effect.

The exploratory analyses on the effect of encoding specificity on CI produced some significant results. The encoding context (i.e., whether participants worked in the presence of, or interacted with, their group members at encoding) moderated the CI effect, $Q = 16.55$, $p = .002$. All groups demonstrated significant or near-significant CI, with the strongest effect being found among Group 5 (the traditional design, in which both nominal and collaborative group members encoded individually), $d = -0.88$, $p < .001$, and the weakest effect being found among Group 1 (in which both nominal and collaborative group members encode in interactive groups), $d = -0.26$, $p = .058$. Our exploratory analyses of interest were the differences in effect sizes between Groups 1 and 2, and between Group 5 and all other groups. Group 2 (in which nominal group members encoded individually and collaborative group members encoded in interactive groups) yielded a stronger CI effect, $d = -0.73$, than Group 1, $Q = 4.72$, $p = .030$. Group 5 yielded a significantly stronger effect than Group 1, $Q = 16.15$, $p < .001$, and Group 3 (in which nominal group members encoded individually and collaborative group members encoded in noninteractive groups), $d = -0.67$, $Q = 4.18$, $p = .041$, but yielded a nonsignificantly stronger effect than Group 2, $p = .42$, and Group 4 (in which both nominal and collaborative group members encoded in noninteractive groups), $d = -0.69$, $p = .15$. Overall, these results lend little support to the suggestion that the CI effect found in traditional study paradigms is due to encoding specificity principles. We elaborate on the interpretation of these findings in the Discussion.

The single variable related to participant characteristics, the relationship between collaborative group members (Hypothesis 8), was a significant moderator of CI, $Q = 8.45$, $p = .004$. Both groups of strangers and groups of nonstrangers experienced CI ($ps < .001$), but the mean CI effect was larger when group members were strangers, $d = -0.77$, than when group members were friends or spouses, $d = -0.44$.

Postcollaborative Memory

Summary effect size. Seventeen research reports (22 studies and 27 independent effect sizes) were included in the meta-analysis examining the effect of prior collaboration on subsequent individual memory performance, with a total of 2,446 participants across all studies. Figure 3 lists all samples included. The effect sizes (standardized difference in means) ranged from −0.33 to +2.60. Of these, the majority (23 or 85%) was positive, and four (15%) were negative. The overall mean effect size was $d = 0.59$, which was significantly different from zero, $z = 5.58$, $p < .001$; $CI_{0.95}$ [0.38, 0.79]. Duval and Tweedie's *trim and fill* procedure suggests no asymmetry in the distribution of effect sizes and thus

Study name	subgroup	Statistics for each study						Sample size		Weight
		Std diff in means	Std error	Lower limit	Upper limit	Z-value	*p*-value	Nominal groups	Collab. groups	Rel. weight
Barber&Rajaram,2011A_2	-	0.76	0.15	0.47	1.05	5.08	0.00	96	96	4.25
Barber&Rajaram,2011B	-	1.06	0.12	0.82	1.30	8.59	0.00	150	150	4.37
Blumen&Rajaram,2008	-	0.51	0.21	0.10	0.91	2.44	0.01	48	48	3.94
Blumen&Stern,2011	Older adults	0.66	0.27	0.14	1.18	2.50	0.01	30	30	3.59
Blumen&Stern,2011	Young adults	1.28	0.28	0.72	1.83	4.51	0.00	30	30	3.48
Choi et al.,2014	-	0.77	0.17	0.44	1.11	4.49	0.00	54	108	4.13
Congleton&Rajaram,2011	Repeated Study	0.25	0.17	-0.08	0.58	1.49	0.14	72	72	4.16
Congleton&Rajaram,2011	Repeated Test	0.62	0.17	0.28	0.95	3.62	0.00	72	72	4.14
Congleton&Rajaram,2014	-	0.67	0.24	0.19	1.14	2.75	0.01	36	36	3.73
Finlay et al.,2000_1	-	0.31	0.29	-0.26	0.88	1.07	0.29	24	24	3.43
Finlay et al.,2000_3	-	-0.13	0.26	-0.64	0.38	-0.50	0.62	30	30	3.63
Harris et al.,2013_1	Group encoding	1.12	0.28	0.58	1.67	4.04	0.00	30	30	3.51
Harris et al.,2013_1	Ind. encoding	-0.29	0.24	-0.77	0.19	-1.19	0.23	39	30	3.72
Harris et al.,2013_2	Group encoding	0.81	0.27	0.28	1.34	3.02	0.00	30	30	3.57
Harris et al.,2013_2	Ind. encoding	-0.33	0.26	-0.83	0.18	-1.25	0.21	30	30	3.62
Henkel&Rajaram,2011	Older adults	0.29	0.21	-0.11	0.69	1.40	0.16	48	48	3.95
Henkel&Rajaram,2011	Young adults	0.50	0.21	0.10	0.91	2.43	0.02	48	48	3.94
Meade&Roediger,2009_2	-	0.09	0.23	-0.35	0.54	0.41	0.68	40	38	3.82
Peker&Tekcan,2009	-	0.20	0.22	-0.22	0.63	0.95	0.34	42	45	3.89
Stephenson&Wagner,1989	-	0.90	0.49	-0.07	1.87	1.82	0.07	9	9	2.29
Thorley&Dewhurst,2007	-	0.00	0.26	-0.52	0.52	0.00	1.00	36	24	3.60
Thorley,2007_2	-	2.60	0.28	2.06	3.15	9.39	0.00	48	48	3.51
Thorley,2007_3	-	0.76	0.26	0.25	1.26	2.95	0.00	48	24	3.63
Thorley,2007_4	-	1.11	0.27	0.59	1.63	4.17	0.00	48	24	3.58
Weldon&Bellinger,1997_1	-	1.15	0.22	0.72	1.59	5.23	0.00	48	48	3.86
Weldon&Bellinger,1997_2	-	0.57	0.29	-0.01	1.14	1.92	0.05	24	24	3.41
Wright&Klumpp,2004	-	-0.30	0.32	-0.92	0.33	-0.93	0.35	20	20	3.26
		0.59	**0.11**	**0.38**	**0.79**	**5.58**	**0.00**			

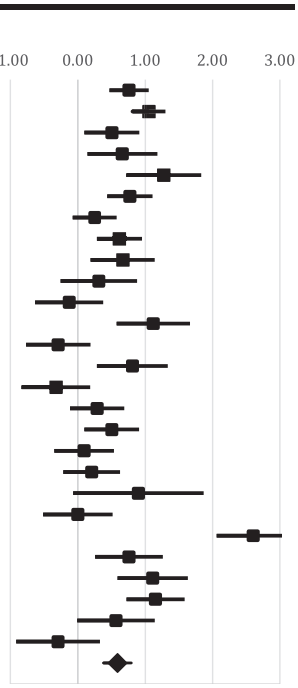

Figure 3. Forrest plot of effect sizes included in postcollaborative memory meta-analysis. Note: The number following the publication year refers to the study number (e.g., "1" refers to Study 1, etc.). The summary effect values appear on the last row (values in bold).

the overall mean effect size does not need to be adjusted (see Figure 4). This medium effect size (Cohen, 1988) suggests that, compared with a prior individual recollection event, prior collaboration on a memory task benefits future individual memory, supporting the postcollaborative advantage hypothesis (Hypothesis 9). There was significant heterogeneity in the sample of effect sizes, $Q = 152.32$, $p < .001$, $I^2 = 82.93$, and thus some of the variance observed in effect sizes could potentially be explained by moderating variables.

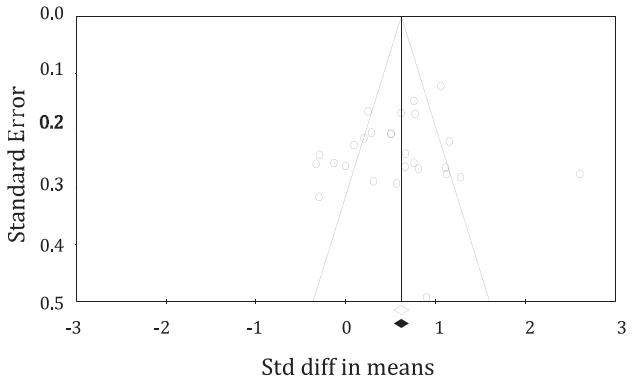

Figure 4. Duval and Tweedie Trim and Fill Imputed Funnel Plot. Clear circles are those studies that were included in the meta-analysis. Filled circles represent imputed studies that are thought to be missing due to publication bias (i.e., unpublished studies due to null or reversed effects). The clear diamond represents the original effect size, while the filled diamond represents the adjusted effect size (which in this case are superimposed). The widths of the diamonds represent variance. (Note that in this analysis, the effect size did not need adjustment and no missing studies were imputed.)

Moderator analyses. To determine whether variables pertaining to study design characteristics had moderating effects on the postcollaborative advantage effect, an overall standardized difference in means was obtained for each level of the variables and then compared in an analysis akin to the ANOVA in primary studies. All potential moderator variables for this meta-analysis were categorical. Table 2 lists all moderator variables along with relevant statistics.

Analyses yielded significant results for two potential moderators, interitem association and group size. However, in regards to Hypothesis 10, an effect opposite to that predicted was observed: Although both levels of interitem association yielded significant postcollaborative advantage effects (ps $< .001$), study designs that fostered a high association between study items resulted in a smaller benefit of prior collaboration, $d = 0.40$, compared with study designs with a low or moderate interitem association potential, $d = 0.75$, $Q = 16.34$, $p < .001$. Regarding our exploratory group size variable, previous members of collaborative dyads ($d = 0.30$) as well as previous members of collaborative triads and quartets ($d = 0.69$) significantly benefitted from collaboration compared to previous members of nominal groups of respective sizes (ps $< .002$). However, we observed that participants who formed collaborative dyads benefitted less from this collaboration, $d = .30$, than participants who formed collaborative triads or quartets, $d = .69$, $Q = 12.49$, $p < .001$. Finally, the interaction of

Table 2
Analysis of Potential Moderator Variables of Postcollaborative Memory

Variable	Number of effect sizes	Mean effect size (*d*)	95% confidence interval	*Q*-test
Interitem association				16.34**
Low/Med	16	.75	.65, .86	
High	11	.40	.27, .53	
Interaction at retrieval				.09
Free-flowing	23	.62	.53, .71	
Turn-taking	4	.58	.34, .83	
Group size				12.49**
2	8	.30	.11, .50	
3	19	.69	.60, .78	
Memory test				.01
Free recall	22	.63	.54, .72	
Recognition	3	.62	.35, .90	

** $p < .05$.

the collaborative group at retrieval was not a significant moderator of the relationship between prior collaborative remembering and subsequent individual memory, $p = .77$, and neither was the type of individual memory test given to participants postcollaboration, $p = .95$, although postcollaborative advantages were found at all levels of both variables (all ps $< .001$).

Discussion

The meta-analyses reported here had two aims. The purpose of the first meta-analysis was to systematically test variables that were predicted by the RSDH to moderate the relationship between collaboration and memory performance. The second meta-analysis focused upon postcollaborative memory and aimed to establish which factors contribute to a rebound from the inhibitory impact of collaboration on retrieval and which contribute to the persistence of CI in later individual recall (i.e., forgetting). Both meta-analyses also assessed the impact of three alternative cognitive mechanisms on CI (encoding specificity, retrieval inhibition, and retrieval blocking). We start with a discussion of our findings related to CI, followed by a discussion of our findings related to postcollaborative memory.

Retrieval Strategy Disruption as a Cause of CI

The first meta-analysis revealed a significant summary effect of CI: Collaborative groups remembered less studied information than equivalent sized nominal groups. The dominant theoretical explanation for this effect is the RSDH. This theory suggests that collaborative group members each attempt to recollect information in an order consistent with their own optimal retrieval strategies. As each member's strategy differs, they disrupt each other and the group's recollection is impaired. As nominal group members work alone, they can use their own optimal retrieval strategies without any disruption and therefore recollect more than collaborative groups. If the RSDH is correct, factors that influence the degree to which individuals can utilize their own retrieval strategies during collaborative remembering tasks should moderate CI. The veracity of the theory was therefore tested by examining the impact of eight such factors on CI. Consistent with the RSDH, five moderated CI and these are considered next.

First, collaborative group size moderated CI, with the effect being most pronounced in larger groups (see also Basden et al., 2000; Thorley & Dewhurst, 2007). This is consistent with a RSDH account of CI as larger groups contain a greater number of competing retrieval strategies, intensifying the disruption each individual experiences when the group collaborates. It is important to note that although the RSDH predicts more pronounced CI in larger groups, no study to date has measured the effect in groups larger than four. Thus, it is unknown whether CI increases linearly as group size increases beyond four or whether the effect plateaus. This warrants future investigation. Importantly, CI was evident in collaborative pairs, a finding reported in some previous studies (e.g., Thorley & Dewhurst, 2007) but not others (e.g., Andersson & Rönnberg, 1995). In fact, some researchers have reported testing triads, as opposed to more economical dyads, because of the published evidence suggesting that a CI effect does not occur amongst dyads. Our findings, however, suggest that dyads can effectively be used for research on CI.

Second, the type of retrieval a collaborative group engaged in moderated CI, with larger effects observed during turn-taking retrieval than free-flowing retrieval. This is also consistent with the RSDH's explanation of CI: Turn-taking retrieval is highly disruptive to collaborative group members' retrieval strategies as they can only contribute one item at a time, making it difficult to recollect information in their own preferred order. Free-flowing retrieval is less disruptive as it permits multiple contributions at once, allowing group members to recollect some information in their own preferred order. Only two studies had previously compared the impact of both retrieval types on collaborative remembering and no differences were observed (Harris et al., 2012, Exp. 2; Thorley & Dewhurst, 2007). Our finding is therefore novel but aligns with the RSDH.

Third, the type of memory test a collaborative group completed also moderated CI, with weaker effects occurring during forced-order tests than free-order tests. Again, this finding is consistent with the RSDH. Forced-order tests impose a retrieval strategy on both collaborative and nominal group members such that neither can use their own retrieval strategies, making the test format detrimental for both. Free-order tests, on the other hand, impose no retrieval strategy and members from both groups can use their own (with more efficient use for nominal group members). CI was, however, also observed during forced-order tests. The RSDH would not predict this as forced-order tests should be equally disruptive for collaborative and nominal groups. Inconsistent findings have been reported in the literature, with some researchers observing CI during forced-order tests (e.g., Meade & Roediger, 2009) but others not (e.g., Thorley & Dewhurst, 2009). This suggests additional mechanisms may be contributing to CI during forced-order tests, such as retrieval inhibition (e.g., Anderson et al., 1994; Bäuml & Aslan, 2004) or retrieval blocking (Rundus, 1973). Future research is needed to clarify the conditions under which forced-order tests induce CI.

Fourth, study materials moderated CI, with smaller effects occurring for story-like materials than uncategorized materials (i.e., unrelated words or items). According to the RSDH, this would be expected as story-like materials, unlike uncategorized materials, have an inherent structure that align group members' retrieval strategies during encoding so that they are less likely to disrupt one another during retrieval. This finding clarifies uncertainty sur-

rounding this issue as only Andersson and Rönnberg (1995) had previously compared recollection of both study materials and, whilst their findings are in line with ours, their study was confounded as memory of a story was assessed via cued recall (which can eliminate CI) and memory of uncategorized words was assessed via free recall (which can facilitate CI). An additional finding of interest here is that CI was equivalent for uncategorized materials and categorized materials (e.g., related words), suggesting categorized materials lack enough structure to sufficiently align group members' retrieval strategies and reduce CI.

Fifth, the social relationship of collaborative group members moderated CI, with the effect being stronger for groups of strangers than close acquaintances. This observation can also be interpreted in terms of the RSDH. In brief, close acquaintances can have a transactive memory system. This is a shared memory system that enables them to implicitly or explicitly divide responsibility for learning and retrieving nonoverlapping information (e.g., Wegner et al., 1991). Similarly, among strangers, CI is not observed when collaborative group members recollect nonoverlapping information (Basden et al., 1997, Exp. 3). This lack of effect is thought to occur as the group members can focus on their own individual retrieval strategies, meaning they are unaffected by hearing others' differing strategies. A moderating effect of social relationship has been reported elsewhere (e.g., Andersson & Rönnberg, 1995) although some studies report no difference between close acquaintances and strangers (e.g., Harris et al., 2013). Those studies with null results failed to establish whether or not the close acquaintances had a transactive memory. If they did not, this could explain their findings. Future research would benefit from establishing this using a validated measure so the role of transactive memories in moderating CI can be established (see Hewitt & Roberts, 2015, for one measure). It is also important to consider the role that communication styles play in moderating CI in close acquaintances, as demonstrated by Harris et al. (2011, see the introduction for more details).

The above results all suggest that retrieval strategy disruption is a cause of CI. However, three factors that have been purported to influence group member's retrieval strategies did not moderate CI. The first factor was category size. Smaller categories are believed to align collaborative group members' retrieval strategies during encoding, reducing retrieval strategy disruption and CI (Rajaram & Pereira-Pasarin, 2010). Repeating an earlier point, it is possible that categorized material lacks a sufficient structure to align collaborative group members' retrieval strategies, meaning they develop individualized strategies irrespective of category size, and succumb to retrieval strategy disruption/CI. The second factor was the number of study/test phases in a study. It has been suggested that multiple encoding and retrieval phases strengthen collaborative group members' retrieval strategies, making them less susceptible to disruption from collaboration (Basden et al., 2000). Some evidence supports this (e.g., Pereira-Pasarin & Rajaram, 2011, Exp. 1) but exceptions exist (e.g., Congleton & Rajaram, 2014). It is possible that repeated encoding and repeated retrieval have different effects on CI (e.g., repeated testing may have a greater protective effect against CI than repeated study, Congleton & Rajaram, 2011), which would not be accounted for in this analysis because both procedures were included in the same coded variable.[2] Future research is needed to clarify whether or not differential effects do occur. The final factor was the encoding task. It has been suggested that retrieval strategies are formed partly based on individuals' expectations of the retrieval context (Basden & Basden, 1995). From this, we assumed that individuals engaging in incidental encoding would be less likely to strategically organize the study material than participants who knew of an impending test, which would lead to a decrease in CI in the former group. Our null finding suggest this assumption may be incorrect and supports the belief that new information is subjectively organized in memory in a similar fashion regardless of whether it is encoded incidentally or intentionally (e.g., Mandler, 1967).

To summarize, the above results suggest the RSDH is a good explanation of CI. This claim is based on evidence showing that five factors which can influence collaborative group members' use of their own retrieval strategies during collaboration also moderate CI. Several results running counter to this theory were, however, observed. For example, reliable CI occurs during forced-order memory tests. Three factors were also presumed to influence individual retrieval strategies did not moderate CI. From the evidence here, it seems unlikely that two of these (category size; encoding task) do actually influence retrieval strategies, whereas the third (the number of study and test phases) most likely does influence them but the analysis was unable to detect this.

Alternative Mechanisms as a Cause of CI

An additional aim of the meta-analyses was to explore the impact of three alternative cognitive mechanisms on CI. The first to be discussed is encoding specificity. To recap, CI is typically found in studies comparing collaborative groups whose members encode material individually and recollect it as a group to nominal groups whose members both encode and recollect material individually. Context-dependent learning predicts improved memory performance when the group composition (i.e., the physical context) at encoding and retrieval is similar, whereas TAP predicts improved memory performance when the level of interaction between group members (i.e., the cognitive context) at encoding and retrieval is similar. Our first meta-analysis examined the impact of different physical and cognitive contexts on CI (as measured by the encoding context variable), offering an insight into whether or not context-dependent learning and TAP moderate CI. All physical and cognitive contexts induced CI, with the effect being strongest in Group 5 studies (those using the standard paradigm described above). If context-dependent memory moderates CI, then the effect should be smaller for Group 4 studies (in which collaborative and nominal group members encoded in noninteractive groups) than Group 5 studies as the matched physical context during encoding and retrieval in Group 4 collaborative groups gives them an advantage over their comparator nominal groups. CI was, however, equivalent for the Group 4 and 5 studies. This suggests context-dependent memory does not moderate CI. Similarly, if TAP moderates CI then collaborative and nominal groups in Group 2 studies (where collaborative group members interacted at encoding and nominal group members worked alone) should have equivalent memory performance since the cognitive contexts are

[2] Unfortunately, separate analyses for repeated study and repeated testing was not performed in this synthesis due to the limited number of moderators that could be tested based on our sample size of included effect sizes.

similar for both. However, CI was again observed among Group 2 studies. It is unlikely that this effect was caused by retrieval strategy disruption as nominal group members could use their individual retrieval strategies and collaborative group members could use the group retrieval strategy developed during encoding (meaning no CI would be anticipated). It is acknowledged, however, that this assumes the collaborative groups developed a shared retrieval strategy during encoding. We recommend further research on this issue before completely ruling out an influence of TAP on CI.

The type of memory test used for the assessment of postcollaborative memory performance was included as a moderator in the second meta-analysis to investigate the potential contribution of retrieval strategy disruption, retrieval inhibition, and retrieval blocking as exploratory mechanisms of CI. Based on the part-list cuing literature, Barber et al. (2015) proposed that comparing the postcollaborative memory of previous collaborative and nominal group members on free recall and recognition tests could reveal the relative contribution of these three mechanisms. Specifically, forgetting effects with the use of both types of tests would suggest that retrieval inhibition has caused CI; rebound effects with the use of both types of tests would suggest that retrieval disruption has caused CI; and a forgetting effect with the use of free recall tests/a rebound effect with the use of recognition tests would suggest that retrieval blocking has caused CI. In our analysis we compared the postcollaborative memory effects between studies that used free recall tests and recognition tests, and found that both types of tests resulted in almost identical and moderately strong postcollaborative advantage effects (i.e., rebound effects). Consistent with Barber et al. (2015), the results clearly demonstrate that retrieval blocking is an unlikely contributor to CI. It is also possible that some forgetting does occur, due to retrieval inhibition, but that the compounding beneficial effect from material reexposure results in a postcollaborative advantage. However, because our obtained mean effect sizes are positive and moderately strong, it is much more likely that they reflect strong rebound effects and, consistent with most of our analyses in the CI meta-analysis, a strong contribution of retrieval strategy disruption.

We must point out, however, that it is still possible that retrieval strategy disruption and retrieval inhibition both contributed to CI in the studies included in the present analyses. Whereas Barber et al. (2015) found evidence both of these mechanisms contribute to CI, their experimental paradigm only allowed a demonstration of the effect of rebound, as they had each group member study and retrieve a different list of words and thus they could not have been reexposed to study material during collaboration. The studies included in our analyses, on the other hand, employed a paradigm in which all group members studied the same material and reexposure during collaboration could therefore serve as a source of relearning for the other group members. Thus, our analysis is limited in that it is impossible to completely separate the effects of rebound and reexposure (and, to some extent, forgetting).

Postcollaborative Memory

Despite the disruptive effect of collaboration on group remembering, the good news is that later individual memory generally benefits from prior collaborative retrieval. The summary effect found in the second meta-analysis suggests that the detrimental effect of collaboration on group retrieval does not carry over to future individual retrieval. In fact, prior collaboration provides a general *advantage* on later individual remembering, an effect that had not always been reported in the literature (e.g., Finlay et al., 2000; Meade & Roediger, 2009; Wright & Klumpp, 2004). A memory benefit experienced by former collaborative group members in comparison to former nominal group members was found at all levels of the potential moderator variables assessed. This robust finding strongly supports the RSDH, which predicts a rebound from any potential retrieval strategy disruption experienced during collaboration on later individual remembering.[3]

In addition, the size of the groups moderated the beneficial effect of collaboration, providing support for the idea that reexposure to study material during collaboration contributes to its beneficial effect for later individual remembering. Significant postcollaborative advantage effects were found for groups of all sizes, but former members of larger collaborative groups experienced an advantage. Because groups of three or four retrieve more of the study material than dyads, the former are reexposed to a greater proportion of the total material to be remembered, giving its members a greater rehearsal opportunity. Thus, despite our findings from Hypothesis 1 that larger collaborative groups experience greater CI, this finding suggests that the greater benefit from reexposure to the study material experienced by members of larger collaborative groups overshadows the obstacle that larger groups pose to the strengthening of members' individual retrieval strategies.

We had also anticipated that factors that are likely to increase the interitem association of study material would translate to increased rebound from retrieval strategy disruption, as similar effects are found in part-list cueing paradigms (e.g., Bäuml & Aslan, 2006). However, the type of interaction at retrieval (with a free-flowing procedure providing a better opportunity to strengthen individualized retrieval strategies) did not moderate postcollaborative memory effects. It is possible that the benefit of reexposure offered by both free-flowing and turn-taking approaches during collaborative interactions overshadows any benefit that a free-flowing interaction may offer by increasing the interitem association of the material. Similarly, study procedures that fostered a greater degree of interitem association between study items did not lead to a greater rebound effect. In fact, we were surprised to find the opposite: A greater benefit of prior collaboration was observed in low/moderate interitem association studies. One possibility may be that, as discussed above, repeated testing and repeated retrieval (which the frequency of both was used to code this variable) have different effects on the strengthening of item associations, thereby biasing our results for this variable. Another possibility, however, may be that when group members already have a very strong retrieval strategy, the benefits of reexposure are less effective.

[3] It is important to note that reminiscence—the number of new items that are recalled during individual postcollaborative remembering that were not recalled by the collaborative group (Basden et al., 2000)—would be a better way to measure rebound, and could help distinguish between items remembered postcollaboratively due to a release from strategy disruption and those due to reexposure during collaboration. Unfortunately, measures of postcollaborative reminiscence are not often reported. Future research on collaborative and postcollaborative memory should consider including these measures to enable future meta-analytic reviews to consider these different effects.

Retrieval strategies that are less rehearsed may be more flexible and thus could be more easily adapted to incorporate items to which an individual is reexposed, leading to higher postcollaboration memory performance. In fact, evidence supporting this idea comes from Bäuml and Aslan (2006) who found a greater beneficial effect of reexposure to cue words on later uncued tests when the list items were low in interitem association than when they were high in interitem association.

Collaborative Memory Errors

The meta-analyses reported here focused on correct retrieval only and did not examine the impact of collaboration upon retrieval errors. The small literature on this issue reveals that collaborative groups make fewer retrieval errors than nominal groups (e.g., Barber et al., 2010; Hyman et al., 2013; Pereira-Pasarin & Rajaram, 2011). This occurs as collaborative group members can correct each other's errors (Ross et al., 2004). The number of errors a collaborative group makes may be influenced by several factors, such as the type of retrieval engaged in and the retrieval instructions received. For example, it has been found that free-flowing retrieval, which permits discussion and error correction, produces fewer errors than turn-taking retrieval, which prohibits discussion and error correction (Basden et al., 1997; Meade & Roediger, 2009; Thorley & Dewhurst, 2007). Similarly, collaborative groups who are required to reach a consensus regarding the accuracy of retrieved information make fewer errors than those not required to reach a consensus (Harris et al., 2012, 2013). Conversely, the memory conformity literature (sometimes known as the social contagion literature) also shows that retrieval errors made by one collaborative group member can contaminate the memory of other group members (see Wright, Memon, Skagerberg, & Gabbert, 2009). Whilst progress is being made with regards to understanding the factors that moderate collaborative retrieval errors, there is much to learn. A meta-analysis that establishes which factors impact retrieval errors would therefore be timely and, combined with the results from the current synthesis, would offer researchers guidance on improving both the completeness and accuracy of collaborative groups' retrieval.

Applied Implications

Collaborative remembering is an everyday activity, used in a variety of contexts including social gatherings (e.g., friends reminiscing about a past holiday), the workplace (e.g., interview panels recalling candidates' performances), schools and universities (e.g., students forming study groups to revise lecture content), and legal settings (e.g., juries recollecting trial evidence during deliberation). Despite its ubiquity, few studies have examined the impact of collaboration upon everyday memory. Instead, researchers have primarily conducted controlled laboratory-based studies addressing theoretical issues relating to recall of basic materials such as word lists. The results of our meta-analyses are therefore largely representative of that literature. The extent to which the findings from these meta-analyses generalize to different everyday memory contexts is unclear. Whilst laboratory findings often generalize to everyday memory contexts, this is not always the case (see Kvavilashvili & Ellis, 2004).

Understanding whether or not the current findings generalize to everyday memory contexts would be beneficial as there is a wealth of potential applied value within collaborative memory research that has been largely overlooked (although see Blumen, Rajaram, & Henkel, 2013, for a discussion of the potential value for both healthy and cognitively impaired older adults, and Rajaram & Pereira-Pasarin, 2010, and Nokes-Malach, Richey, & Gadgil, 2015, for discussions of the potential value in educational contexts). In the remainder of this section we briefly consider how the costs and benefits of collaboration identified in these meta-analyses may be useful in guiding future research into best practice within the context of education. We focus on education as school and college classrooms often employ collaboration as an instructional technique to facilitate student learning. It is hoped readers find this brief discussion useful for highlighting the potential, often neglected, applied value of collaborative memory research.

Collaborative learning activities have been used in schools and higher education systems globally for decades to enhance students' retention of new knowledge (Kollias, Mamalougos, Vamvakoussi, Lakkala, & Vosniadou, 2005; Johnson & Johnson, 2009). Collaborative learning activities differ from collaborative memory tasks as they do not typically require group retrieval. For example, Jigsaw activities require individual group members to learn one element of a larger topic and then teach other group members about this element so that the entire group learns about the entire topic (e.g., Aronson, Blaney, Stephan, Sikes, & Snapp, 1978). As in some collaborative memory studies, educators then ask group members to complete individual postcollaborative tests to assess their learning. Our second meta-analyses demonstrated that collaborative recall enhances postcollaborative individual learning with a medium effect observed ($d = 0.59$). A similar postcollaborative enhancement has been observed following collaborative learning. For example, Johnson, Johnson, and Stanne's (2000) meta-analysis demonstrates that engaging in any one of eight widely used collaborative learning activities can facilitate later individual learning, with effect sizes ranging from small ($d = .09$) to large ($d = .91$). Given that collaborative remembering can be as effective at enhancing individual learning as collaborative learning activities, this raises the question as to whether or not educators could use collaborative memory in classrooms to enhance students learning. It is premature at this juncture to recommend educators do this, especially because collaborative memory studies have primarily been conducted with college-aged or older adults and the generalizability of the findings to younger school-aged students is unclear. Despite this, the findings from our postcollaborative memory meta-analysis can direct future research exploring this possibility. For example, it was found that prior collaboration in groups of three or four produced a greater postcollaborative learning enhancement than prior collaboration in a pair. It would therefore be beneficial to examine whether or not prior collaborative group size influences students' recall of classroom materials in the same way. If so, then educators could encourage students to collaboratively recall newly taught materials in groups of three or four to enhance their learning.

Conclusion

CI is a robust effect and the RSDH provides a strong explanation of the mechanism underlying the effect. CI is increased when collaboration takes place among larger groups, when uncategorized items are retrieved, when remembering takes place in a

free-flowing manner, when a free-order test is used at retrieval, and when the collaborators are strangers. All of these factors are believed to increase disruption to individuals' retrieval strategies during collaborative retrieval but not during individual retrieval. The results of these meta-analyses also suggest minimal contributions of retrieval blocking and encoding specificity on CI, and some potential contribution of retrieval inhibition. Despite the initial costs to memory, however, collaboration can offer long-term benefits to recollection completeness when group members subsequently remember events alone, due to a rebound from the disruption, as well as the reexposure to the study material during collaboration. The field would benefit further from a systematic review that focusses on specifying the effects of collaboration on retrieval errors and an in-depth exploration of how collaboration can help or hinder memory in real-life contexts such as educational settings.

References

*Abbe, A. (2004). *Collaboration and acquaintance in group recall* (Doctoral dissertation). University of California, Riverside. Retrieved from http://search.proquest.com

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1063–1087. http://dx.doi.org/10.1037/0278-7393.20.5.1063

Andersson, J. (2001). Net effect of memory collaboration: How is collaboration affected by factors such as friendship, gender and age? *Scandinavian Journal of Psychology, 42,* 367–375. http://dx.doi.org/10.1111/1467-9450.00248

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. http://dx.doi.org/10.1080/09658210444000557

Andersson, J., & Rönnberg, J. (1995). Recall suffers from collaboration: Joint recall effects of friendship and task complexity. *Applied Cognitive Psychology, 9,* 199–211. http://dx.doi.org/10.1002/acp.2350090303

*Andersson, J., & Rönnberg, J. (1996). Collaboration and memory: Effects of dyadic retrieval on different memory tasks. *Applied Cognitive Psychology, 10,* 171–181. http://dx.doi.org/10.1002/(SICI)1099-0720(199604)10:2<171::AID-ACP385>3.0.CO;2-D

*Andersson, J., & Rönnberg, J. (1997). Cued memory collaboration: Effects of friendship and type of retrieval cue. *European Journal of Cognitive Psychology, 9,* 273–287. http://dx.doi.org/10.1080/713752558

Aronson, E., Blaney, N., Stephan, C., Sikes, J., & Snapp, M. (1978). *The jigsaw classroom*. Beverly Hills, CA: Sage.

Aslan, A., Bäuml, K. H., & Grundgeiger, T. (2007). The role of inhibitory processes in part-list cuing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 335–341. http://dx.doi.org/10.1037/0278-7393.33.2.335

Barber, S. J., Harris, C. B., & Rajaram, S. (2015). Why two heads apart are better than two heads together: Multiple mechanisms underlie the collaborative inhibition effect in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 559–566. http://dx.doi.org/10.1037/xlm0000037

*Barber, S. J., & Rajaram, S. (2011a). Collaborative memory and part-set cueing impairments: The role of executive depletion in modulating retrieval disruption. *Memory, 19,* 378–397. http://dx.doi.org/10.1080/09658211.2011.575787

*Barber, S. J., & Rajaram, S. (2011b). Exploring the relationship between retrieval disruption from collaboration and recall. *Memory, 19,* 462–469. http://dx.doi.org/10.1080/09658211.2011.584389

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

*Barber, S. J., Rajaram, S., & Fox, E. B. (2012). Learning and remembering with others: The key role of retrieval in shaping group recall and collective memory. *Social Cognition, 30,* 121–132. http://dx.doi.org/10.1521/soco.2012.30.1.121

*Basden, B. H., Basden, D. R., Bryner, S., & Thomas, 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. http://dx.doi.org/10.1037/0278-7393.23.5.1176

*Basden, B. H., Basden, D. R., & Henry, S. (2000). Costs and benefits of collaborative remembering. *Applied Cognitive Psychology, 14,* 497–507. http://dx.doi.org/10.1002/1099-0720(200011/12)14:6<497::AID-ACP665>3.0.CO;2-4

Basden, B. H., Basden, D. R., & Stephens, J. P. (2002). Part-set cuing of order information in recall tests. *Journal of Memory and Language, 47,* 517–529. http://dx.doi.org/10.1016/S0749-596X(02)00016-5

*Basden, B. H., Basden, D. R., Thomas, R. L., & Souphasith, S. (1998). Memory distortion in group recall. *Current Psychology: Development, Learning, Personality, Social, 10,* 225–246.

Basden, D. R., & Basden, B. H. (1995). Some tests of the strategy disruption interpretation of part-list cuing inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1656–1669. http://dx.doi.org/10.1037/0278-7393.21.6.1656

Basden, D. R., Basden, B. H., & Galloway, B. C. (1977). Inhibition with part-list cuing: Some tests of the item strength hypothesis. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 100–108. http://dx.doi.org/10.1037/0278-7393.3.1.100

Basden, D. R., & Draper, J. S. (1973). Effect of cuing with list members as a function of number of categories, taxonomic frequency, and presentation order. *Canadian Journal of Psychology, 27,* 327–333.

Bäuml, K. H., & Aslan, A. (2004). Part-list cuing as instructed retrieval inhibition. *Memory & Cognition, 32,* 610–617. http://dx.doi.org/10.3758/BF03195852

Bäuml, K. H., & Aslan, A. (2006). Part-list cuing can be transient and lasting: The role of encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 33–43. http://dx.doi.org/10.1037/0278-7393.32.1.33

*Blumen, H. M., & Rajaram, S. (2008). Influence of re-exposure and retrieval disruption during group collaboration on later individual recall. *Memory, 16,* 231–244. http://dx.doi.org/10.1080/09658210701804495

Blumen, H. M., & Rajaram, S. (2009). Effects of repeated collaborative retrieval on individual memory vary as a function of recall versus recognition tasks. *Memory, 17,* 840–846. http://dx.doi.org/10.1080/09658210903266931

Blumen, H. M., Rajaram, S., & Henkel, L. (2013). The applied value of collaborative memory research in aging: Considerations for broadening the scope. *Journal of Applied Research in Memory & Cognition, 2,* 133–135. http://dx.doi.org/10.1016/j.jarmac.2013.05.004

*Blumen, H. M., & Stern, Y. (2011). Short-term and long-term collaboration benefits on individual recall in younger and older adults. *Memory & Cognition, 39,* 147–154. http://dx.doi.org/10.3758/s13421-010-0023-6

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley, Ltd. http://dx.doi.org/10.1002/9780470743386

Brown, V. R., & Paulus, P. B. (2002). Making group brainstorming more effective: Recommendations from an associative memory perspective. *Current Directions in Psychological Science, 11,* 208–212. http://dx.doi.org/10.1111/1467-8721.00202

Brown, V., Tumeo, M., Larey, T. S., & Paulus, P. B. (1998). Modeling cognitive interactions during group brainstorming. *Small Group Research, 29,* 495–526. http://dx.doi.org/10.1177/1046496498294005

Chan, M. E., & Arvey, R. D. (2012). Meta-analysis and the development of knowledge. *Perspectives on Psychological Science, 7,* 79–92. http://dx.doi.org/10.1177/1745691611429355

*Choi, H.-Y., Blumen, H. M., Congleton, A. R., & Rajaram, S. (2014). The role of group configuration in the social transmission of memory: Evidence from identical and reconfigured groups. *Journal of Cognitive Psychology, 26,* 65–80. http://dx.doi.org/10.1080/20445911.2013.862536

Clark, N. K., & Stephenson, G. M. (1989). *Group remembering*. Hillsdale, NJ: Erlbaum, Inc.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum Associates, Inc.

Cohn, L. D., & Becker, B. J. (2003). How meta-analysis increases statistical power. *Psychological Methods, 8,* 243–253. http://dx.doi.org/10.1037/1082-989X.8.3.243

Collaros, P. A., & Anderson, L. R. (1969). Effect of perceived expertness upon creativity of members of brainstorming groups. *Journal of Applied Psychology, 53,* 159–163. http://dx.doi.org/10.1037/h0027034

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

*Congleton, A. R., & Rajaram, S. (2011). The influence of learning methods on collaboration: Prior repeated retrieval enhances retrieval organization, abolishes collaborative inhibition, and promotes post-collaborative memory. *Journal of Experimental Psychology: General, 140,* 535–551. http://dx.doi.org/10.1037/a0024308

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

Cuc, A., Koppel, J., & Hirst, W. (2007). Silence is not golden: A case for socially shared retrieval-induced forgetting. *Psychological Science, 18,* 727–733. http://dx.doi.org/10.1111/j.1467-9280.2007.01967.x

Dahlström, Ö., Danielsson, H., Emilsson, M., & Andersson, J. (2011). Does retrieval strategy disruption cause general and specific collaborative inhibition? *Memory, 19,* 140–154. http://dx.doi.org/10.1080/09658211.2010.539571

Dickersin, K. (2005). Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. In H. R. Rothstein & A. J. Sutton (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 11–33). Chichester, UK: Wiley, Ltd.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology, 53,* 497–509. http://dx.doi.org/10.1037/0022-3514.53.3.497

Duval, S., & Tweedie, R. (2000). A nonparametric 'trim and fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89–98.

Ekeocha, J. O., & Brennan, S. E. (2008). Collaborative recall in face-to-face and electronic groups. *Memory, 16,* 245–261. http://dx.doi.org/10.1080/09658210701807480

*Finlay, F., Hitch, G. J., & Meudell, P. R. (2000). Mutual inhibition in collaborative recall: Evidence for a retrieval-based account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1556–1567. http://dx.doi.org/10.1037/0278-7393.26.6.1556

*Garcia-Marques, L., Garrido, M. V., Hamilton, D. L., & Ferreira, M. (2012). Effects of correspondence between encoding and retrieval organization in social memory. *Journal of Experimental Social Psychology, 48,* 200–206. http://dx.doi.org/10.1016/j.jesp.2011.06.017

*Garrido, M. V., Garcia-Marques, L., & Hamilton, D. L. (2012). Enhancing the comparability between part-list cueing and collaborative recall: A gradual part-list cueing paradigm. *Experimental Psychology, 59,* 199–205. http://dx.doi.org/10.1027/1618-3169/a000144

Griffith, T. L., Fichman, M., & Moreland, R. L. (1989). Social loafing and social facilitation: An empirical test of the cognitive-motivational model of performance. *Basic and Applied Social Psychology, 10,* 253–271. http://dx.doi.org/10.1207/s15324834basp1003_4

Harkins, S. G., Latane, B., & Williams, K. (1980). Social loafing: Allocating effort or taking it easy? *Journal of Experimental Social Psychology, 16,* 457–465. http://dx.doi.org/10.1016/0022-1031(80)90051-7

Harkins, S. G., & Petty, R. E. (1982). Effects of task difficulty and task uniqueness on social loafing. *Journal of Personality and Social Psychology, 43,* 1214–1229. http://dx.doi.org/10.1037/0022-3514.43.6.1214

*Harris, C. B., Barnier, A. J., & Sutton, J. (2012). Consensus collaboration enhances group and individual recall accuracy. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 65,* 179–194. http://dx.doi.org/10.1080/17470218.2011.608590

*Harris, C. B., Barnier, A. J., & Sutton, J. (2013). Shared encoding and the costs and benefits of collaborative recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 183–195. http://dx.doi.org/10.1037/a0028906

*Harris, C. B., Keil, P. G., Sutton, J., Barnier, A. J., & McIlwain, D. J. F. (2011). We remember, we forget: Collaborative remembering in older couples. *Discourse Processes, 48,* 267–303. http://dx.doi.org/10.1080/0163853X.2010.541854

*Henkel, L. A., & Rajaram, S. (2011). Collaborative remembering in older adults: Age-invariant outcomes in the context of episodic recall deficits. *Psychology and Aging, 26,* 532–545. http://dx.doi.org/10.1037/a0023106

Hewitt, L. Y., & Roberts, L. D. (2015). Transactive memory systems scale for couples: Development and validation. *Frontiers in Psychology, 6,* 516. http://dx.doi.org/10.3389/fpsyg.2015.00516

Hollingshead, A. B. (1998). Retrieval processes in transactive memory systems. *Journal of Personality and Social Psychology, 74,* 659–671.

*Hyman, I. E., Jr., Cardwell, B. A., & Roy, R. A. (2013). Multiple causes of collaborative inhibition in memory for categorised word lists. *Memory, 21,* 875–890. http://dx.doi.org/10.1080/09658211.2013.769058

Ingham, A. G., Levinger, G., Graves, J., & Peckham, V. (1974). The Ringelmann effect: Studies of group size and group performance. *Journal of Experimental Social Psychology, 10,* 371–384. http://dx.doi.org/10.1016/0022-1031(74)90033-X

Johansson, N. O., Andersson, J., & Rönnberg, J. (2005). Compensating strategies in collaborative remembering in very old couples. *Scandinavian Journal of Psychology, 46,* 349–359. http://dx.doi.org/10.1111/j.1467-9450.2005.00465.x

Johansson, O., Andersson, J., & Rönnberg, J. (2000). Do elderly couples have a better prospective memory than other elderly people when they collaborate? *Applied Cognitive Psychology, 14,* 121–133. http://dx.doi.org/10.1002/(SICI)1099-0720(200003/04)14:2<121::AID-ACP626>3.0.CO;2-A

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38,* 365–379. http://dx.doi.org/10.3102/0013189X09339057

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis*. Retrieved from https://www.researchgate.net/publication/220040324_*Cooperative_Learning_Methods_A_Meta-Analysis*

Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology, 65,* 681–706. http://dx.doi.org/10.1037/0022-3514.65.4.681

*Kelley, M. R., Pentz, C., & Reysen, M. B. (2014). The joint influence of collaboration and part-set cueing. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 67,* 1977–1985. http://dx.doi.org/10.1080/17470218.2014.881405

*Kelley, M. R., Reysen, M. B., Ahlstrand, K. M., & Pentz, C. J. (2012). Collaborative inhibition persists following social processing. *Journal of*

Cognitive Psychology, 24, 727–734. http://dx.doi.org/10.1080/20445911.2012.684945

Kollias, V., Mamalougos, N., Vamvakoussi, X., Lakkala, M., & Vosniadou, S. (2005). Teachers' attitudes to and beliefs about web-based collaborative learning environments in the context of an international implementation. *Computers & Education, 45,* 295–315. http://dx.doi.org/10.1016/j.compedu.2005.04.012

Kvavilashvili, L., & Ellis, J. (2004). Ecological validity and the real-life/laboratory controversy in memory research: A critical and historical review. *History & Philosophy of Psychology, 6,* 59–80. Retrieved from http://uhra.herts.ac.uk/bitstream/handle/2299/1971/103125.pdf;jsessionid=3A1D4E3876CCAC308A2145E185CA5850?sequence=1

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174. http://dx.doi.org/10.2307/2529310

Latane, 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. http://dx.doi.org/10.1037/0022-3514.37.6.822

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Mandler, G. (1967). Organization and memory. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 328–372). New York, NY: Academic Press.

*Meade, M. L., Nokes, T. J., & Morrow, D. G. (2009). Expertise promotes facilitation on a collaborative memory task. *Memory, 17,* 39–48. http://dx.doi.org/10.1080/09658210802524240

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

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, 48,* 141–152. http://dx.doi.org/10.1080/14640749508401381

Meudell, P. R., Hitch, G. J., & Kirby, P. (1992). Are two heads better than one? Experimental investigations of the social facilitation of memory. *Applied Cognitive Psychology, 6,* 525–543. http://dx.doi.org/10.1002/acp.2350060606

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16,* 519–533. http://dx.doi.org/10.1016/S0022-5371(77)80016-9

Mullen, B. (1983). Operationalizing the effect of the group on the individual: A self-attention perspective. *Journal of Experimental Social Psychology, 19,* 295–322. http://dx.doi.org/10.1016/00221031(83)90025-2

Mullen, B., & Baumeister, R. F. (1987). Review of personality and social psychology. In C. Hendricks (Ed.), *Group effects on self-attention and performance: Social loafing, social facilitation, and social impairment* (Vol. 9, pp. 189–206). Newbury Park, CA: Sage.

Nickerson, R. S. (1984). Retrieval inhibition from part-set cuing: A persisting enigma in memory research. *Memory & Cognition, 12,* 531–552. http://dx.doi.org/10.3758/BF03213342

Nijstad, B. A., & Stroebe, W. (2006). How the group affects the mind: A cognitive model of idea generation in groups. *Personality and Social Psychology Review, 10,* 186–213. http://dx.doi.org/10.1207/s15327957pspr1003_1

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. (2003). Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology, 39,* 531–548. http://dx.doi.org/10.1016/S0022-1031(03)00040-4

Nokes-Malach, T. J., Richey, J. E., & Gadgil, S. (2015). When is it better to learn together? Insights from research on collaborative learning. *Educational Psychology Review, 27,* 645–656. http://dx.doi.org/10.1007/s10648-015-9312-8

Payne, D. G. (1987). Hypermnesia and reminiscence in recall: A historical and empirical review. *Psychological Bulletin, 101,* 5–27.

*Peker, M., & Tekcan, A. İ. (2009). The role of familiarity among group members in collaborative inhibition and social contagion. *Social Psychology, 40,* 111–118. http://dx.doi.org/10.1027/1864-9335.40.3.111

*Pereira-Pasarin, L. (2007). Divided attention and repetition at encoding: The role of individual subjective organization in group recall. *Dissertation Abstracts International: Section B: The Sciences and Engineering,* (MSTAR_622050502; 2009–99080-006).

*Pereira-Pasarin, L. P., & Rajaram, S. (2011). Study repetition and divided attention: Effects of encoding manipulations on collaborative inhibition in group recall. *Memory & Cognition, 39,* 968–976. http://dx.doi.org/10.3758/s13421-011-0087-y

Rajaram, S. (2011). Collaboration both hurts and helps memory: A cognitive perspective. *Current Directions in Psychological Science, 20,* 76–81. http://dx.doi.org/10.1177/0963721411403251

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

*Reysen, M. B., Talbert, N. G., Dominko, M., Jones, A. N., & Kelley, M. R. (2011). The effects of collaboration on recall of social information. *British Journal of Psychology, 102,* 646–661. http://dx.doi.org/10.1111/j.2044-8295.2011.02035.x

*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. http://dx.doi.org/10.1037/0882-7974.23.1.85

*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. http://dx.doi.org/10.1002/acp.1023

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89,* 63–77. http://dx.doi.org/10.1037/h0031185

Rundus, D. (1973). Negative effects of using list items as recall cues. *Journal of Verbal Learning and Verbal Behavior, 12,* 43–50. http://dx.doi.org/10.1016/S0022-5371(73)80059-3

Smith, E. E., Adams, N., & Schorr, D. (1978). Fact retrieval and the paradox of interference. *Cognitive Psychology, 10,* 438–464. http://dx.doi.org/10.1016/0010-0285(78)90007-5

Stephenson, G. M., Kniveton, B. H., & Wagner, W. (1991). Social influences on remembering: Intellectual, interpersonal and intergroup components. *European Journal of Social Psychology, 21,* 463–475. http://dx.doi.org/10.1002/ejsp.2420210602

*Stephenson, G. M., & Wagner, W. (1989). Origins of the misplaced confidence effect in collaborative recall. *Applied Cognitive Psychology, 3,* 227–236. http://dx.doi.org/10.1002/acp.2350030303

*Takahashi, M. (2007). Does collaborative remembering reduce false memories? *British Journal of Psychology, 98,* 1–13. http://dx.doi.org/10.1348/000712606X101628

*Takahashi, M., & Saito, S. (2004). Does test delay eliminate collaborative inhibition? *Memory, 12,* 722–731. http://dx.doi.org/10.1080/09658210344000521

Taylor, D. W., Berry, P. C., & Block, C. H. (1958). Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly, 3,* 23–47. http://dx.doi.org/10.2307/2390603

*Thorley, C. (2007). *Collaborative false memories* (Doctoral dissertation). Lancaster University, Lancaster, UK. Retrieved from http://search.proquest.com

*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. http://dx.doi.org/10.1080/09541440600872068

Thorley, C., & Dewhurst, S. A. (2009). False and veridical collaborative recognition. *Memory, 17,* 17–25. http://dx.doi.org/10.1080/09658210802484817

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York, NY: Springer-Verlag.

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In. B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4612-4634-3_9

Wegner, D. M., Erber, R., & Raymond, P. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology, 61,* 923–929. http://dx.doi.org/10.1037/0022-3514.61.6.923

Wegner, D. M., Giuliano, T., & Hertel, P. (1985). Cognitive interdependence in close relationships. In W. J. Ickes (Ed.), *Compatible and incompatible relationships* (pp. 253–276). New York, NY: Springer-Verlag. http://dx.doi.org/10.1007/978-1-4612-5044-9_12

*Weigold, A., Russell, E. J., & Natera, S. N. (2014). Correction of false memory for associated word lists by collaborating groups. *The American Journal of Psychology, 127,* 183–190. http://dx.doi.org/10.5406/amerjpsyc.127.2.0183

*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. http://dx.doi.org/10.1037/0278-7393.23.5.1160

*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,* 1568–1577. http://dx.doi.org/10.1037/0278-7393.26.6.1568

*Wessel, I., Zandstra, A. R. E., Hengeveld, H. M. E., & Moulds, M. L. (2015). Collaborative recall of details of an emotional film. *Memory, 23,* 437–444. http://dx.doi.org/10.1080/09658211.2014.895384

Williams, K., Harkins, S. G., & Latané, B. (1981). Identifiability as a deterrent to social loafing: Two cheering experiments. *Journal of Personality and Social Psychology, 40,* 303–311. http://dx.doi.org/10.1037/0022-3514.40.2.303

*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. http://dx.doi.org/10.3758/BF03196740

Wright, D. B., Memon, A., Skagerberg, E. M., & Gabbert, F. (2009). When eyewitnesses talk. *Current Directions in Psychological Science, 18,* 174–178. http://dx.doi.org/10.1111/j.1467-8721.2009.01631.x

Received June 22, 2015
Revision received May 25, 2016
Accepted July 13, 2016 ■