A meta-analytic investigation into the factors moderating the number of errors groups make when recalling shared information.

[redacted]

University of Adelaide

This thesis is submitted in partial fulfillment of the Honours degree of Bachelor of Psychological Science.

Submitted on 24 September 2023

Word Count: 6996

Table of Contents

List of Figures 4

List of Tables 4

Abstract 5

Declaration 6

Contributor roles 7

A Meta-Analytic Investigation into the Factors Moderating the Number of Errors Groups Make When Recalling Shared Information. 10

- Recall Groups 10
- Potential Moderators of Group Errors 13
  - Error Type 13
  - Retrieval Method 14
  - Consensus Requirement 14
  - Group Size 15
  - Group Member Relationship 16
  - Group Member Age 16
- Present Meta-Analyses Overview 17
  - Hypothesis 1: Error type. 17
  - Hypothesis 2: Retrieval method. 17
  - Hypothesis 3: Consensus requirement. 18
  - Hypothesis 4: Group size. 18
  - Exploratory analysis 1: Group member relationship. 19
  - Exploratory analysis 2: Group member age. 19

Method 20

- Study Search and Inclusion Criteria 20

Calculation of Effect Sizes 23
Factor Coding 24
Factor Reliability Coding 26
Data Analysis 26
Meta-Analysis 1 Results 27
Summary effect size 27
Meta-Analysis 2 Results 29
Summary effect size 29
Moderator analysis 30
Discussion 31
Meta-analysis 1 32
Meta-analysis 2 32
Applied implications. 35
Limitations 36
Future directions 36
Conclusion 37
Appendix A 47
Appendix B 50

List of Figures

Figure 1. Collaborative vs. Nominal Group Encoding and Retrieval………………………. 12

Figure 2. PRISMA 2020 V2 Flowchart………………………………………………..……. 22

List of Tables

Table 1. Results from two meta-analyses examining whether six factors moderate collaboration's impact on group memory errors…………………………………… 28

Abstract

Collaborative remembering occurs in many everyday settings, including social gatherings (for example, a family recalling childhood memories) and courts (for example, deliberating jurors recalling trial evidence). When groups recall shared experiences, they can make mistakes. Researchers have examined whether collaborative groups make more errors than an equivalent number of lone individuals whose recall is pooled (i.e., a nominal group). This is studied by having collaborative and nominal groups study and recall simple stimuli, such as word lists. Collaborative groups typically make fewer errors than nominal groups, as collaborative group members prune each other's errors. Occasionally, however, collaborative groups make more mistakes than nominal groups, suggesting collaboration sometimes harms accuracy. Here, two meta-analyses examined whether nine factors moderate the number of errors collaborative groups make, relative to nominal groups. Three factors moderated the number of errors made. First, when studies deliberately induced errors, collaborative groups made as many errors as nominal groups. However, when studies focussed on spontaneous errors, collaborative groups made fewer errors than nominal groups. Second, the retrieval method groups engage in moderated the number of errors made. Specifically, if collaborative group members took turns to recall items, they made as many errors as nominal groups. When their members recalled in a 'free-for-all' manner, they made fewer errors than nominal groups. Third, whilst collaborative groups generally made fewer errors than nominal groups; this effect was accentuated when they could only add items to their recall that all members agree on. The applied implications of these findings are considered.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University, and, to the best of my knowledge, this thesis contains no material previously published except where due reference is made. I give permission for the digital version of this thesis to be made available on the web, via the University of Adelaide's digital thesis repository, the Library Search and through web search engines, unless permission has been granted by the School to restrict access for a period of time.

Contributor roles

Role	Role Descriptions	Student	Supervisor 1	Supervisor 2
Conceptualisation	Ideas; formulation or evolution of overarching research goals and aims.		**X**	
Methodology	Development or design of methodology; creation of models.		**X**	
Project Administration	Management and coordination responsibility for the research activity planning and execution.	**X**	**X**	
Supervision	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.		**X**	
Resources	Provision of study materials, laboratory samples, instrumentation, computing resources, or other analysis tools.	**X**	**X**	

Software	Programming; software development; designing computer code and supporting algorithms; testing of existing code.			
Investigation	Conducting research – specifically performing experiments, or data/evidence collection.	**X**		
Validation	Verification of the overall replication/reproducibility of results/experiments.	**X**	**X**	
Data Curation	Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later re-use.	**X**	**X**	
Formal Analysis	Application of statistical, mathematical, computational, or other formal techniques to	**X**	**X**	

	analyse or synthesise study data.			
Visualisation	Visualisation/data presentation of the results.	**X**		
Writing – Original Draft	Specifically writing the initial draft.	**X**		
Writing – Review & Editing	Critical review, commentary, or revision of the original draft.	**X**	**X**	

A Meta-Analytic Investigation into the Factors Moderating the Number of Errors Groups Make When Recalling Shared Information.

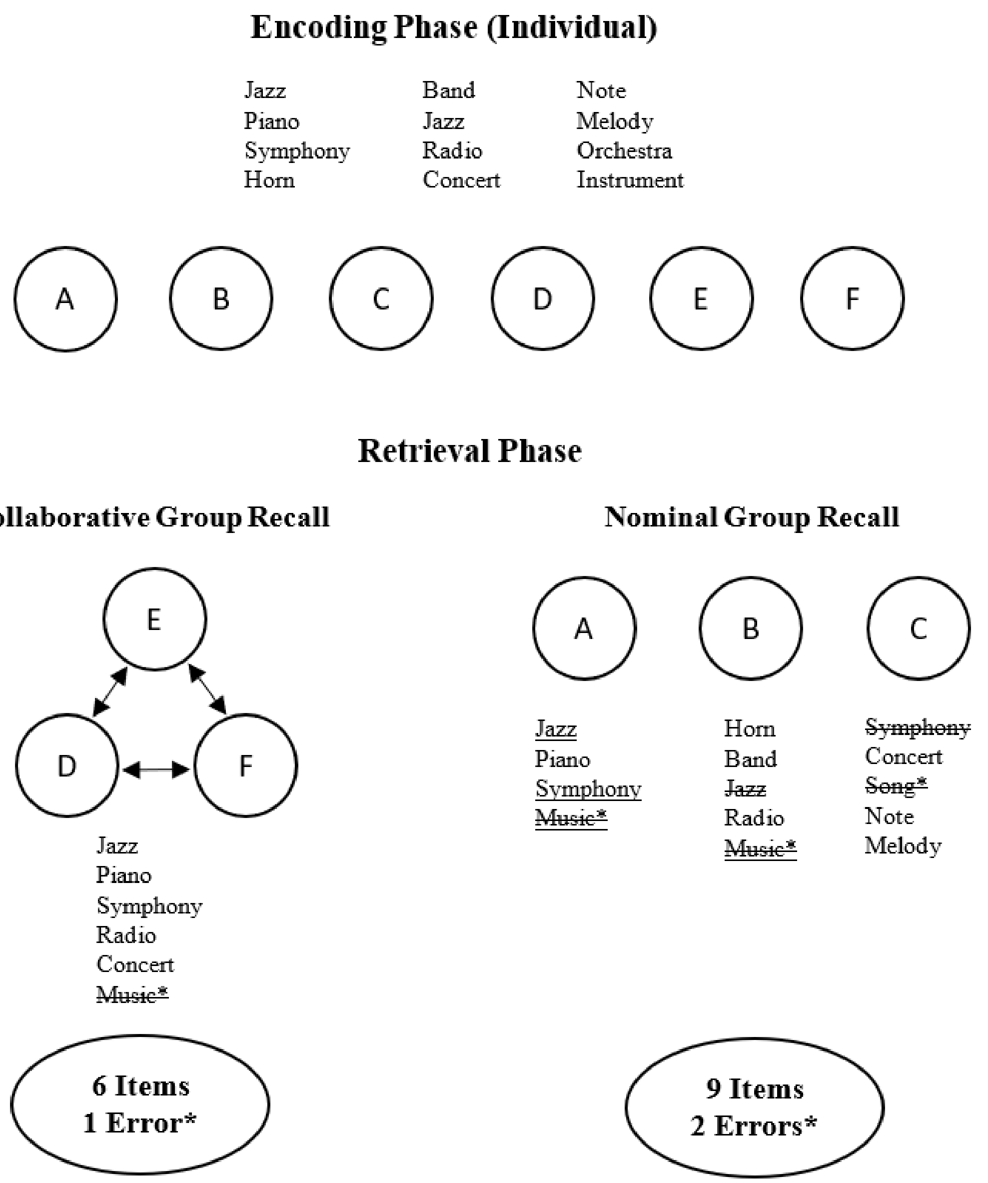
Memory is often social in nature, and people often collaborate to recall shared experiences. For example, a family may collaborate to recall childhood memories, or a jury may collaborate to recall trial evidence to reach a final verdict. When engaging in collaborative memory, group members can make errors (Maswood & Rajaram, 2019). In some settings, such as a jury collaborating to recall trial evidence, recall errors can lead to harmful consequences. Fortunately, a benefit of collaborative group recall is that group members have the ability to correct one another's mistakes. Generally, it has been found that collaborative groups are less likely to produce errors than the pooled answers of an equivalent number of lone individuals (i.e., a nominal group). This is due to the ability of collaborative group members to prune one another's errors (Rossi-Arnaud et al., 2019). This is, however, not always the case, with some studies showing collaborative groups make more errors than nominal groups (see Basden et al., 1998). However, the reasons behind this discrepancy are not fully understood. Here, two meta-analyses examine whether six factors moderate the number of errors collaborative groups make, relative to nominal groups.

Recall Groups

Intuitively, collaborative groups recalling studied material should be able to produce more information than a lone individual, given the increased resources within a group setting. Early research has demonstrated that this is the case (see Clark & Stephenson, 1990); however, comparing collaborative groups to individual recall does not identify whether a lone individual performs similarly to an individual in a group setting (Marion & Thorley, 2016). To address this, modern research studies compare collaborative groups to nominal groups. A collaborative group comprises of a group of individuals that recall the studied information together. A nominal group comprises of the pooled results of two or more individuals who recalled the study information alone. Comparing collaborative groups to nominal groups allows researchers to identify if the process of collaborating with others impacts group recall, by comparing collaborative group recall to the pooled performance of individual recall (Marion & Thorley, 2016). Figure 1 shows the procedure in a typical collaborative remembering study. First, during the encoding phase, all participants study simple materials, like word lists, alone. Then, during the retrieval phase, some participants recall the studied materials in a collaborative group, whilst others recall them in a nominal group. The items recalled by nominal group members are then combined, however, in instances where multiple individuals provide the same item, the item is only counted once. Given the nature of nominal groups' individual recall, nominal group members are unable to discuss their recall and eliminate one another's errors. Collaborative groups are therefore advantageous over nominal groups, as group members can deliberate and eliminate errors from the group output (Harris et al., 2012). The collaborative group monitoring of responses is known as error-pruning (Nie et al., 2019).

Figure 1

Collaborative vs. Nominal Group Encoding and Retrieval



Note. Stimuli are studied individually during the encoding phase. During the retrieval phase, collaborative group members will recall stimuli in the presence of other group members, and nominal group members will recall stimuli individually, and have their answers pooled. Errors recalled by groups during the retrieval phase are denoted by an asterisk and are crossed out. Stimuli recalled by multiple nominal group members are underlined and only contribute to the nominal group's final score once.

Some early studies reported that collaborative groups make more errors than nominal groups (see Basden et al., 1998), while others reported no significant differences between collaborative and nominal group errors (see Weldon & Bellinger, 1997). Despite these early findings, literature has suggested that collaborative groups have the potential to report fewer errors than nominal groups, due to the ability of collaborative groups to engage in error-pruning. Consistent with this idea, studies from 2007 onwards have generally found that collaborative groups make fewer, or sometimes equal, errors compared to nominal groups (Rajaram, 2011). The reasons behind this discrepancy are not well understood; however, it has been suggested that there could be several potential moderators of the number of errors collaborative groups make, such as whether collaborative group members respond in *turn-taking* or *free-for-all* setting. The six potential moderators will now be discussed in turn.

Potential Moderators of Group Errors

Error Type

Errors that are reported during the recall phase can either be induced or spontaneous. Induced errors can occur when the experimenter purposefully attempts to elicit a specific error from the participants (Takahashi, 2007). Errors can be induced by the experimenter in several ways, such as exposing participants to post-event misinformation prior to the recall phase or by asking leading questions during the recall phase. Collaborative groups appear to find it more difficult to prune errors when they are induced. For example, Basden et al. (1998) found that group members produced a higher number of errors when these errors were induced, compared to nominal groups (see also, Thorley & Dewhurst, 2007). However, other studies have found that collaborative and nominal groups make an equal number of errors when errors are induced (see Maswood et al., 2022). Conversely, spontaneous errors occur when the experimenter does not attempt to elicit a specific error and errors are produced naturally by group members. Previous literature has generally reported that collaborative groups make fewer spontaneous errors than nominal groups (see Wissman & Rawson, 2015). It can be assumed that it would be generally more difficult for groups to error-prune when errors are induced; however, this has not been tested directly and therefore further testing is required.

Most studies on collaborative group errors focus on spontaneous errors. Therefore, all subsequent potential moderators considered in Meta-Analysis 2 are in relation to spontaneous errors only.

Retrieval Method

Studies have examined the effect of the retrieval method on collaborative memory. Typically, previous studies have considered one of two main methods for collaboration: *turn-taking* or *free-for-all* (see Thorley & Dewhurst, 2007). In turn-taking collaboration, group members will sequentially provide a recalled study item without deliberation. In a turn-taking environment, the interaction between group members is restricted and therefore, there are no opportunities for error-pruning and increased social pressure to produce a recalled studied item due to the expectation to retrieve on each turn (Maswood & Rajaram, 2019). Previous studies have shown that in turn-taking environments, collaborative groups make as many errors as nominal groups (see Saraiva et al., 2023), or more errors than nominal groups (Thorley & Dewhurst, 2007). Conversely, free-for-all group members can discuss studied items to produce their final recollection. Due to the interactive nature of the free-for-all environment, group members have the opportunity to error-prune (Maswood & Rajaram, 2019). As a result, fewer false memories are likely to be reported in the final group recall report when compared to nominal groups (Maswood & Rajaram, 2019).

Consensus Requirement

A small number of studies require groups to only report on items that all collaborative group members remember being included in the stimuli. Therefore, these studies require a “true consensus” (see Webster, 2017). A study conducted by Harris et al. (2012) found that collaborative groups that were required to reach a true consensus produced fewer errors than their equivalent nominal groups. Collaborative groups that required a true consensus demonstrated an error-pruning process, whereby group members removed one another’s errors during group discussion (Harris et al., 2012). Nominal groups cannot engage in error-pruning and a true consensus cannot be required given that group members are tested individually; therefore, it is expected that nominal groups will make more errors than collaborative groups. Collaborative groups that are not required to reach a true consensus are not forced to engage in error-pruning, and studies have shown that these collaborative groups tend to make less or equal errors when compared to nominal groups (see Hinds & Payne, 2016, and Finlay et al., 2000 for respective examples).

Group Size

Group size refers to the number of people that comprise of the collaborative and nominal groups. Intuitively, one would believe that as group size increases, there is a greater opportunity to error-prune due to the presence of more group members. However, as nominal groups increase in size, there is a larger number of individual people potentially recalling false information with no ability to prune one another’s errors, resulting in a higher number of errors (Wessel et al., 2015). Previous studies have generally found that collaborative pairs make fewer or equal errors compared to nominal groups (see Nie & Deng, 2023, and Saraiva et al., 2023 for respective examples). However, there have been inconsistent findings when larger groups are considered. For example, Harris et al. (2012) report that collaborative trios make the same number of errors as nominal groups while Bärthel et al. (2016) report that collaborative trios make fewer errors than nominal groups. Contrary to intuition, Basden et al. (1997) report that collaborative trios make more errors than nominal groups. Therefore, further exploration is required to identify whether group size moderates collaborative recall.

Group Member Relationship

Collaborative remembering studies often denote whether group members have a prior relationship with one another. Group member relationships can include friendships and romantic partners, or alternatively, group members may be unacquainted. Generally, studies have found that acquainted collaborative group members make fewer errors compared to nominal groups (see Harris et al., 2013). Unacquainted collaborative groups also tend to make fewer errors compared to nominal groups (see Nie et al., 2022). It has been suggested that collaborative group members who are acquainted are more likely to accept recalled stimuli information that their group member provides (Maswood & Rajaram, 2019). This increased acceptance of information is likely due to perceived reliability (Maswood & Rajaram, 2019). Conversely, Ross et al. (2008) report that collaborative group errors were reduced for married couples compared to nominal groups, as partners were more likely to respond to incorrect information with doubt or rejection. At present it is unclear whether collaborative group member relationships moderate group errors.

Group Member Age

There has been little research exploring the effect of age on collaborative remembering errors, and results across different age brackets have been mixed. For example, Hinds and Payne (2016) found that collaborative groups with a mean age of 21.7 years made fewer errors than nominal groups, while Finlay et al. (2000) found that collaborative groups with a mean age of 27 years made equal errors to nominal groups. Comparatively, Barnier et al. (2018) conducted a study of adults with a mean age of 74.7 years, and it was found that collaborative groups make fewer errors than nominal groups. However, Meade and Roediger (2009) note that older adult study participants are often older couples who have been married for many years, and therefore, it can also be suggested that the group member relationship has a moderating effect on collaborative group errors. Due to the limited research on age and collaborative recall, further exploration is required to determine if it is a moderating factor.

Present Meta-Analyses Overview

The present meta-analyses examine whether six factors moderate the number of errors collaborative groups make, relative to nominal groups. The first meta-analysis focuses on collaborative groups when errors are induced by the experimenters versus when errors are spontaneous. The second meta-analysis will consider five other potential moderators when errors are spontaneous. These potential moderators include retrieval method, consensus requirement, group size, group member relationship, and group member age. Based on the previously discussed literature, it is possible to form several hypotheses and exploratory analyses.

Hypothesis 1: Error Type.

As mentioned, in studies where researchers deliberately induce errors, collaborative groups can make as many induced errors as nominal groups (Maswood et al., 2022), or more than nominal groups (Thorley & Dewhurst, 2007). Conversely, in studies where researchers do not deliberately induce errors, collaborative groups consistently make fewer spontaneous errors than nominal groups (see Wissman & Rawson, 2015). Based upon this past research, it is hypothesised that error type will moderate collaborative groups error rates, with collaborative groups making a similar number of errors (or more errors) than nominal groups when induced errors are studied, but collaborative groups also making fewer errors than nominal groups when spontaneous errors are studied.

Hypothesis 2: Retrieval Method.

Previous literature has shown collaborative groups who engage in turn-taking recall sometimes make as many errors as nominal groups (see Saraiva et al., 2023) or more errors than nominal groups (Thorley & Dewhurst, 2007) as group members are unable to engage in error-pruning. Previous literature has also consistently demonstrated that collaborative groups in a free-for-all setting make fewer errors than nominal groups (see Congleton & Rajaram, 2011). Therefore, it is predicted that collaborative groups in a turn-taking setting will make equal, or more errors than nominal groups. Additionally, it is hypothesised that collaborative groups in a free-for-all condition will produce fewer errors than nominal groups.

Hypothesis 3: Consensus Requirement.

Previous literature has reported that collaborative groups who are directed to reach a true consensus are more likely to engage in error-pruning, and therefore will produce fewer errors than nominal groups (see Webster, 2017). Therefore, it is hypothesised that collaborative groups who are required to reach a true consensus will make fewer errors than nominal groups. It is also hypothesised that collaborative groups that are not required to reach a true consensus will also produce fewer errors than nominal groups. It is expected that there will be a larger effect size for collaborative groups that are required to reach a true consensus vs. nominal groups.

Hypothesis 4: Group Size.

Previous research has generally suggested that as collaborative group size increases, the number of errors decreases, however, there have been some inconsistencies in this research. For example, early research conducted by Basden et al. (1997) reported that collaborative trios make more errors than nominal groups, while more recent research reports that collaborative trios make the same or fewer errors than nominal groups (see Harris et al., 2012, and Bärthel et al., 2016 for respective examples). It is expected that as group size increases, there will be a greater opportunity for collaborative group members to engage with error-pruning. Therefore, it is tentatively predicted that larger collaborative groups (i.e., trios or greater) will make fewer errors than nominal groups. It is also predicted that collaborative pairs will also make fewer errors than nominal groups; however, the effect size is predicted to be smaller than when compared to collaborative trios or greater.

The remaining potential moderators are exploratory in nature, due to the conflicting previous research.

Exploratory analysis 1: Group Member Relationship.

Generally, previous literature has shown that collaborative groups make fewer errors than nominal groups for both acquainted (see Barnier et al., 2016) and unacquainted (see Nie & Deng, 2023) group members. There is some discrepancy in the literature. For example, Harris et al. (2012) reported that collaborative groups make equal errors to nominal groups when group members are unacquainted; however, it is unclear whether this can be attributed to the group member relationship as it is also noted that the recall occurred in a turn-taking setting. Additionally, Vredeveldt et al. (2019) reported that collaborative groups make an equal number of errors to nominal groups when group members are acquainted. The moderating effect of group member relationships requires further exploration.

Exploratory analysis 2: Group Member Age.

Previous literature typically studies group members of similar ages, although previous studies have not focussed on age as a moderating factor of collaborative group members. Previous studies have found that older adults typically produce a lower number of errors than nominal groups (see Barnier et al., 2018). Studies using younger participants have found somewhat mixed results, with some results reporting that collaborative groups provide either less (see Nie & Deng, 2023) or equal (see Rossi-Arnaud et al., 2020) errors to nominal groups. Age as a moderating factor in collaborative group errors is exploratory and requires further investigation.

Method

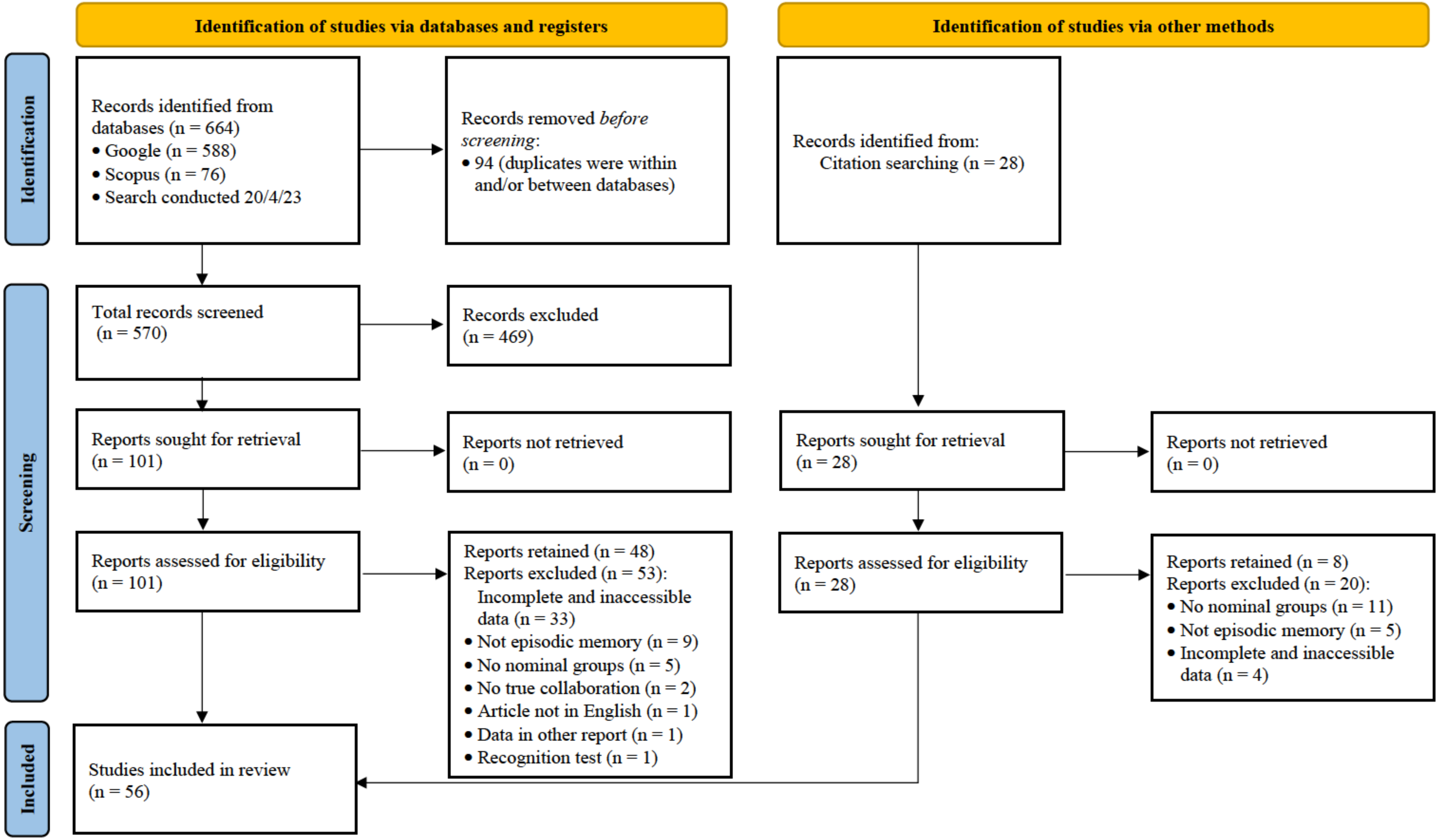
Study Search and Inclusion Criteria

Two bibliographic databases (Google Scholar and Scopus) for studies published between database inception and April 20, 2023. Both databases were searched using the terms *'collaborative memory', OR 'collaborative remembering', OR 'group memory' OR 'group remembering' AND 'nominal group',* as per Boolean logic. Additionally, studies were manually retrieved from collaborative recall literature reviews, a previous collaborative recall meta-analysis, and by examining the reference sections of the publications selected through online searches. The Preferred Reporting Item for Systematic reviews and Meta-Analyses (PRISMA) 2020 V2 flowchart was used to outline included and excluded studies from the meta-analyses (see Figure 2). After the initial sample of 664 publications, the data was screened for duplicates, leaving 570 studies. A pre-screening was conducted to filter out articles that did not focus on collaborative remembering, leaving 101 articles from the databases and 28 articles from other sources. Inclusion and exclusion criteria were applied which resulted in a final sample of 56 studies (48 from databases, 8 from other sources). Inclusion criteria included; (a) studies must have a focus on collaborative memory with collaborative and nominal groups, which may include either a between-subjects or within-subjects design), (b) all study participants must have studied and recalled the same stimuli, (c) all collaborative and nominal groups must have consisted entirely of authentic participants (i.e. no confederates), (d) studies must not have assessed episodic memory, or rather, memories of life events and experiences, (e) the study authors must have included enough information to determine an effect size representing the standardised mean difference between collaborative and nominal group recall error rates, (f) the study must only include objective measures of group errors, and (g) nominal group error rates must have been calculated by pooling together each member's errors and scoring identical errors once.

Articles were also filtered to remove any that were not available in English, as a translator was not available. Additionally, studies that included a recognition test were also excluded, as recognition tests only require the participants to *recognise* a studied item which involves different memory processes to recall (Stadler et al., 1999). After applying this criteria, 56 publications (55 research articles, one doctoral thesis) were identified as containing studies that were eligible for inclusion in the meta-analysis. These 56 publications are marked with an asterisk in the References section. Several of these publications contained more than one eligible study (for example, Hyman et al., 2013, contained three). Overall, the 56 publications contained 74 studies that met the inclusion criteria for the meta-analysis. Those 74 studies contain a total of 2697 collaborative groups and 2640 nominal groups, which produced 129 effect sizes that were eligible to be included in the meta-analyses.

Figure 2

PRISMA 2020 V2 Flowchart



Calculation of Effect Sizes

Effect sizes were calculated using the standardised Mean difference between collaborative group errors and nominal group errors. Comprehensive Meta-Analysis V3 software was used to calculate effect sizes, with Hedges' *g* as the effect size measure. The values of *g* can be interpreted as; 0.2 = small effect size, 0.5 = medium effect size, 0.8 = large effect size. A negative effect size indicates that collaborative groups make fewer errors than nominal groups, and a positive effect size indicates collaborative groups make more errors than nominal groups. A score of zero indicates no difference in the number of errors made between collaborative and nominal groups. When data to calculate effect sizes was not available, but the study provided the results of a t-test or ANOVA from analyses comparing the collaborative and nominal groups, a t statistic or F was used. In cases where this information was not available, authors of papers written in the last ten years were contacted for this information. In cases where this was supplied, the paper was included in the meta-analysis (for example, Nie et al., 2023). Some studies (for example, Basden et al., (1998), experiment 2) only provided approximate numbers of collaborative and nominal groups. In these cases, the approximation provided was used to calculate the effect size. In cases where the required information to calculate the effect size was unable to be obtained, it was excluded from the meta-analysis.

Care was taken to ensure that collaborative and nominal groups only contributed to one effect size. Collaborative and nominal groups contributing to multiple effect sizes violate the assumption of independence, as it inflates the sample size and biases the results. In studies where a collaborative or nominal group contributed to multiple effect sizes, the first effect size was always selected for inclusion (for example, Vredevelt et al., 2019)). Similarly, some studies focused on induced and non-induced errors and only one of these effect sizes featured in the meta-analysis. Additionally, in studies where there was ambiguity around the number of collaborative and nominal groups, the approximation provided was used (see Basden et al., 1998).

Factor Coding

Error type. The type of error was coded as either (a) induced or (b) spontaneous. Studies where experimenters did not attempt to induce any specific errors were coded as spontaneous. Studies were considered induced in instances where the experimenters purposefully try to elicit a specific error. Experimenters are able to induce errors through methods such as asking leading questions or attempting to induce a false memory of a specific word known as a critical lure. Critical lures are typically found in the Deese, Roediger and McDermott (DRM) task, which is a false memory paradigm in which participants are presented with semantically related lists of words to encode (Roediger & McDermott, 1995). After a delay, participants are then asked to recall the list of words (Roediger & McDermott, 1995). Typically, there is a critical word that is not mentioned in the originally presented list, but is recalled with high probability and confidence (Pardilla-Delgado & Payne, 2017). An example of this may be presenting a participant with a list of words such as *nurse*, *hospital,* and *medicine,* but they report the critical lure, *doctor*, despite *doctor* not being presented in the study list. The DRM task is considered an example of an induced error study (Pardilla-Delgado & Payne, 2017).

Retrieval method. The interaction of collaborative group members during recall was coded as (a) *turn-taking* or (b) *free-for-all.* Studies were coded as using a turn-taking procedure when collaborative group members took turns recalling items. In turn-taking groups, each group member can only make one contribution at a time and cannot interact with each other. Alternatively, studies coded as having a free-for-all procedure allowed for unrestricted interaction between group members; so that answers could be discussed and corrected amongst group members. In the present meta-analysis, there were 65 free-for-all and 14 turn-taking effect sizes eligible for inclusion.

Consensus requirement. The requirement of a group consensus was coded as (a) yes or (b) no. For studies that required the group to reach a true consensus, the collaborative groups were instructed to only report items that all group members remember. For studies that did not require a group consensus, collaborative groups were able to report items that only one group member recalled. Studies were classed as mixed where some collaborative groups were required to reach a consensus, but others were not. Studies that did not report whether a group consensus was required were classed as unclear. In the present meta-analysis, 15 true consensus required effect sizes and 55 no consensus required effect sizes were eligible for inclusion.

Group size. The number of participants making up the nominal and collaborative groups was coded for each effect size. Collaborative and nominal groups were always of equal size. Group size was coded as either (a) pairs or (b) trios or greater. Group sizes larger than three people were not coded separately due to the limited number of studies that utilised collaborative groups larger than trios. In the present meta-analysis, 47 paired collaborative group effect sizes and 32 effect sizes of collaborative groups comprised of trios or greater were eligible to be included.

Group member relationship. Collaborative group members were coded as being either (a) acquainted or (b) unacquainted. Acquainted group members included spouses, friends, and classmates. For studies that did not specify the group member relationship but stated that groups were randomly assigned, it was assumed that the group members were unacquainted. For studies that stated that collaborative group members signed up for the study together, it was assumed that the group members were acquainted. For studies where the relationship between group members was ambiguous or varied between groups, the study was not coded. Nominal group relationships were not coded as these participants did not interact with one another and were not aware that their recollection scores would be combined with other participants' scores. In the present meta-analysis, 12 acquainted collaborative group effect sizes and 28 unacquainted collaborative group effect sizes were eligible for inclusion.

Group member age. The mean age of individuals participating in the study was coded into either (a) older participants or (b) younger participants. For means that were under 35.0 years, the study was coded as younger participants. For means that were over 35.0 years old, the study was coded as older participants. These age groups were modelled on studies showing ageing effects on memory (see Old & Naveh-Benjamin, 2008; Fraundorf et al., 2019). In the present study, there were 46 effect sizes of younger adult groups, and six older adult groups effect sizes that were eligible to be included.

Factor Reliability Coding

All articles were initially coded by the first author and a random selection of 12 articles (21.43%) then had their factors independently coded by the supervisor to assess inter-rater reliability. For the eight categorical factors, agreement ranged from 75.00% (group member relationship) to 100% (error type; collaboration method), so was substantial. For the one continuous factor (retention interval), agreement was assessed via an intraclass correlation. r was .84 so, again, there was substantial agreement. Disagreements were reconciled by referring back to the articles and agreeing on the correct code.

Data Analysis

Two meta-analyses were conducted. The first analysis considered the effect of induced versus spontaneous errors on collaborative group errors. The second analysis considered five potential moderators of spontaneous errors only. In each analysis, collaborative group errors were compared to nominal groups comprised of an equal number of individuals. The data was screened for outliers, and none were observed. Hedge's g was used to calculate effect size. A negative g value indicates that collaborative groups have produced fewer errors than nominal groups, and a positive g value indicates that collaborative groups produced more errors than nominal groups. Not all eligible effect sizes were featured in the analyses, as only one effect size was used in circumstances where a collaborative or nominal group contributed two or more effect sizes. A random effects model was used as heterogeneity is expected for both meta-analyses.

Meta-Analysis 1 Results

Summary Effect Size

Meta-analysis 1 considers the moderating effect of induced vs. spontaneous errors on collaborative group errors. The analysis featured 99 effect sizes from 74 studies, with 2231 collaborative groups and 2171 nominal groups. The overall effect size ($g=-0.59$, $p<.001$) shows that collaborative groups make fewer errors than nominal groups, suggesting that collaborative groups engage in error-pruning. See Appendix A for individual effect sizes of individual studies.

The analysis found evidence of high heterogeneity ($Q=456.71$, $p<.001$, $I^2=81\%$ [-0.72, -0.46]). Low heterogeneity is considered to be between 10% to 39%, moderate heterogeneity is considered to be between 40% to 59%, high heterogeneity is considered to be 60% to 89%, and very high heterogeneity is considered to be between 90% to 100% (Schober et al., 2021). Publication bias was not checked for as statistical procedures for checking publication bias can produce invalid results when heterogeneity is high, as indicated by an I^2 value of .75 or above (Harrer et al., 2021).

As a high level of heterogeneity was found, a sub-group analysis was conducted to see if the heterogeneity was caused by two different effect sizes (one for induced errors and one for spontaneous errors). Table 1 shows relevant statistics for the moderator analysis and subsequent analyses. Collaboration reduced the number of errors produced when studies analysed spontaneous errors in comparison to nominal groups (g=-0.80, p<.001). However, there was no significant difference between collaborative and nominal group errors in studies which induced errors (g = -0.26, p=.09). Furthermore, the difference between effect sizes was significant (W=10.29, p=.001).

Table 1

Results from two meta-analyses examining whether six factors moderate collaboration's impact on group memory errors.

Factor	Q-test	95% Pred Int	k	*g*	95% CI
		Meta-Analysis 1			
Error Type	10.29, *p*=.001				
Induced Errors			27	-0.26	-0.56, 0.04
Spontaneous Errors			72	-0.80	-0.94, -0.66*
		Meta-Analysis 2			
Retrieval Method	23.99, *p*<.001				
Turn-Taking			14	0.33	-0.15, 0.82
Free-For-All			68	-0.91	-1.04, -0.79*
Consensus Requirement	4.65, *p* = .031				
Yes			15	-0.93	-1.14, -0.72*
No			58	-0.62	-0.81, -0.42*
Group Size	0.06, *p*= .81				
Pairs			48	-0.73	-0.89, -0.58*

Trios or Greater		34	-0.69	-1.02, -0.36*
Relationship	0.20, *p* = .65			
Acquainted		12	-0.96	-1.21, -0.72*
Unacquainted		29	-1.04	-1.23, -0.85*
Group Member Age	0.48, *p* = .49			
Younger Adults		47	-0.86	-1.01, -0.72*
Older Adults		6	-1.01	-1.38, -0.62*

Note. Meta-Analysis 1 examined recall errors when studies analysed induced or spontaneous errors. Meta-Analysis 2 examined recall errors when studies analysed spontaneous errors only. Negative effect sizes indicate that collaboration has reduced group errors. Positive effect sizes indicate that collaboration increased group errors. k = the number of effect sizes in the analysis. *p<.001, **p = .03

Overall, the results show that collaborative groups make fewer errors than nominal groups when errors are spontaneous. However, collaborative and nominal groups make the same number of errors when errors are induced. Therefore, it can be inferred that less error-pruning occurs in collaborative groups when errors are induced, resulting in a larger number of errors.

Meta-Analysis 2 Results

Summary Effect Size

Meta-analysis 2 considers studies where groups produced spontaneous errors only. There were 82 eligible effect sizes from 59 studies, with 1830 collaborative groups and 1773 nominal groups. The overall effect size (g=-0.71, p<.001) suggests that error-pruning typically occurs amongst collaborative groups and therefore, collaborative groups produce fewer errors than nominal groups. The analysis found evidence of high heterogeneity ($Q=429.28$, $p<0.001$, $I^2= 81.13\%$ [-0.88, -0.56]). As a result, publication bias was not checked for as there is no current method providing acceptable results when the between-study heterogeneity is high. Furthermore, a sub-group analysis was conducted to see if the heterogeneity was being caused by multiple different effect sizes. Five moderators were examined in relation to collaborative group spontaneous errors. Full results are shown in Table 1. See Appendix B for individual effect sizes of individual studies.

Moderator analysis

Retrieval method. The first moderator examined was the *retrieval type*. There were two subgroups, turn-taking, and free-for-all. As expected, it was found that collaborative groups that engaged in free-for-all recall made fewer errors than nominal groups ($g=-0.91$, $p<.001$), while there was no significant difference between turn-taking collaborative group errors and nominal group errors ($g=-0.33$, $p=.18$). These two effect sizes were significantly different ($Q=23.99$, $p<.001$), suggesting that retrieval method is a moderator of collaborative group errors.

Consensus requirement. The second moderator variable examined was *consensus requirement.* There were two subgroups, yes (true consensus required) and no (no consensus required). Collaborative groups that were required to reach a true consensus were shown to make fewer errors than nominal groups ($g=-0.93$, $p<.001$). This was also true for collaborative groups that were not required to reach a true consensus ($g=-0.62$, $p<.001$), suggesting that error-pruning occurs whether or not collaborative groups were required to reach a true consensus. However, collaborative groups that were required to reach a true consensus made fewer errors than non-true consensus collaborative groups, when compared to their respective nominal groups. The two effect sizes were significantly different ($Q=4.65$,

$p=.03$), suggesting that a true consensus requirement is a moderator of collaborative group errors.

Group size. The third moderator examined was *group size*. There were two subgroups, pairs and trios or greater. As expected, paired collaborative groups made fewer errors than their comparative nominal groups ($g=-0.73$, $p<.001$). Collaborative trios or greater also produced fewer errors than their comparative nominal groups ($g=-0.69$, $p<.001$). However, contrary to the hypothesis, the two effect sizes were not significantly different ($Q=0.06$, $p=.81$), suggesting that group size is not a moderator of collaborative group errors.

Group member relationship. The fourth moderator examined was *group member relationship.* There were two subgroups, acquainted and unacquainted. It was found that acquainted collaborative groups made fewer errors than nominal groups ($g=-0.964$, $p<.001$). This was also true for unacquainted collaborative groups ($g=-1.04$, $p<.001$). The two effect sizes were not significantly different ($Q=0.20$, $p=.65$), suggesting that the group member relationship is not a moderator of collaborative group errors.

Group member age. The seventh moderator variable examined was *group member age*. There were two subgroups, younger adults (mean age <35 years), and older adults (mean age > 35 years). Collaborative groups comprised of younger adults produced fewer errors than nominal groups ($g=-0.86$, $p<.001$). Collaborative groups comprised of older adults also produced fewer errors than nominal groups (-1.006, $p<.001$) The two effect sizes were not significantly different ($Q=0.48$, $p=.49$), suggesting that group member age is not a moderator of collaborative group errors.

Discussion

The present meta-analyses aimed to explore six potential moderators of collaborative group errors. The first meta-analysis focused on collaborative groups when errors were induced by the experimenters versus when errors were spontaneous. The second meta-analysis considered five other potential moderators for spontaneous errors only. The potential moderators included retrieval method, consensus requirement, group size, group member relationship, and group member age. Each finding related to collaborative group errors will be discussed in turn, followed by applied implications and future directions.

Meta-Analysis 1

The first meta-analysis revealed that error type moderated collaborative group errors. Collaborative groups made as many errors as nominal groups when errors were induced, however, made fewer errors than nominal groups when errors were spontaneous. These results suggest that inducing errors makes it more difficult for collaborative group members to engage in error-pruning. These findings align with previous research conducted by Maswood et al. (2022) and Thorley and Dewhurst (2007), which employed DRM word lists to induce errors and found comparable error rates between collaborative and nominal groups. Additionally, Wissman and Rawson (2015) provided study participants with related word lists for encoding and found that collaborative groups make fewer spontaneous errors than nominal groups (Wissman and Rawson, 2015). Overall, the results of the present meta-analysis clarify the moderating effect of inducing errors on collaborative group recall.

Meta-Analysis 2

The second meta-analysis considered five potential moderators of spontaneous collaborative group errors, which revealed two moderators of collaborative group errors.

First, it was found that the retrieval method moderated collaborative group errors. Collaborative groups who engaged in turn-taking recall made as many errors as nominal groups (see also, Saraiva et al., 2023). Collaborative groups that engaged in free-for-all recall made fewer errors than nominal groups (see also, Congleton & Rajaram, 2011). Maswood and Rajaram (2019) report that due to the interactive nature of the free-for-all collaborative group environment, group members can actively engage in error-pruning. However, turn-taking collaborative groups are unable to actively engage in the error-pruning process, and therefore, are just as likely to make errors as nominal groups (Maswood & Rajaram, 2019). Overall, the findings of the present meta-analysis have been reported consistently throughout previous literature, and therefore, provide further clarification that the retrieval method has a moderating effect on collaborative groups' ability to error-prune.

Second, it was shown that collaborative groups produced fewer errors than their corresponding nominal groups, whether or not they were required to reach a true consensus. However, it is noted that the effect size was larger among groups that were required to reach a true consensus, suggesting that the consensus requirement plays a role in moderating collaborative group errors. This finding aligns with previous literature. For example, Webster (2017) reported that collaborative groups are less likely to produce errors compared to nominal groups, however, the effect size is greater for those required to reach a true consensus. It can be deduced that collaborative groups that are required to reach a true consensus are forced to engage in the error-pruning process, and therefore, are less likely to produce errors than their corresponding nominal groups. Overall, it has been found that the consensus requirement is a moderator of collaborative group errors.

Third, the results demonstrate that group size did not moderate collaborative group errors. This finding is inconsistent with previous literature, as it has generally been found collaborative group errors will decrease with larger group sizes, due to the increased number of group members that are available to error-prune. For example, previous studies have mainly reported that collaborative groups make fewer errors in pairs than nominal groups (see Nie et al., 2019); however, some studies have reported that collaborative and nominal group pairs make equal errors (see Saraiva et al., 2023). Additionally, most studies for trios or greater tend to show collaborative groups make fewer errors (see Abel & Bauml, 2017). The present findings, however, suggest that the group size does not affect the number of errors collaborative groups make. It noted that most existing literature only examines collaborative groups comprised of pairs or trios, and therefore, it is reasonable to expect that the differences in collaborative group size may not vary enough to observe a significant effect size.

Fourth, it was shown that both acquainted and unacquainted collaborative groups make fewer errors than their respective nominal groups. This is consistent with previous literature, which generally shows that collaborative groups make fewer errors than nominal groups for both acquainted and unacquainted group members. For example, Barnier et al. (2016) reported that collaborative romantic couples make fewer errors than nominal groups. Additionally, Nie and Deng (2023) tested unacquainted collaborative pairs and reported that collaborative groups engaged in error-pruning, and therefore made fewer errors than nominal groups. However, there is some discrepancy in the literature. For example, Vredeveldt et al. (2019) reported that collaborative groups make an equal number of errors to nominal groups when group members are acquainted. Maswood and Rajaram (2019) have suggested that collaborative group members are more likely to accept recalled stimuli information from acquainted group members, due to perceived reliability, therefore reducing error-pruning. Ross et al. (2008) provide an alternative perspective, in that married couples are more likely to respond to incorrect information with doubt or rejection, therefore increasing error pruning. It is reasonable to suggest that the type of relationship (specifically, whether group members are friends, university peers, or married couples) may contribute to the conflicting research. Overall, the present findings suggest that group member relationship is not a moderator of collaborative group errors, however, this may be due to the lack of specificity in the type of group member relationship.

Fifth, it was found that both younger and older adult collaborative groups made fewer errors than their comparative nominal groups. Group member age is an exploratory analysis as previous literature has varied. Typically, it has been found that older adults typically produce a lower number of errors than nominal groups (see Barnier et al., 2018). Studies using younger participants have found somewhat mixed results. For example, Nie and Deng (2023) tested young adults aged 18 to 25 years and found that collaborative groups make fewer errors than nominal groups, whereas a study conducted by Rossi-Arnaud et al. (2020) tested young adults aged 16 to 31 years of age and found collaborative groups made the same number of errors as nominal groups. Overall, the results suggest that group member age is not a moderator of collaborative group errors.

Applied Implications

Collaborative remembering occurs frequently in day-to-day contexts, including casual social settings, such as family recalling childhood memories. Collaborative memory is not as pervasive in clinical psychology contexts, where memory is typically investigated at the individual level (Wessel & Moulds, 2008). However, collaborative memory and group recall errors offer important insight in clinical settings, particularly in the context of family trauma therapy (Rajaram, 2011). Collaborative remembering can be either beneficial or harmful in these settings, particularly when considering erroneous memories. The core element of trauma therapy often involves prolonged, repeated reliving of the traumatic memory through providing a verbal narrative, allowing the space for the client to “emotionally process” the memory (Wessel & Moulds, 2008). In collaborative trauma therapy, including the input of others who have memories of the traumatic event may assist a client to gather evidence and prune errors in their own memory (Wessel & Moulds, 2008). Specifically, the findings of the present research suggest that a free-for-all collaborative remembering setting where group members reach a true consensus on past events, can encourage error-pruning, challenge distorted cognitions, offer alternative details that may reduce the traumatic impact of a memory (see also, Rajaram, 2011).

Limitations

One limitation is the exclusion of several eligible studies that, while potentially valuable, did not provide sufficient information to calculate an effect size. The omission of such studies brings limitations to the comprehensiveness of the analysis, as it may exclude relevant insights from the broader body of research. Additionally, it is recognised that there may be other factors influencing the number of errors collaborative groups make, and therefore, the results cannot be overgeneralised. The moderators outlined in the present research do not constitute an exhaustive list of potential moderators that could be explored, however, highlight the potential for future research in relation to collaborative remembering and recall errors. Finally, several potential moderators that were not found to influence collaborative group errors were limited by the factor coding in Meta-Analysis 2. For example, group member relationship was coded as (a) acquainted or (b) unacquainted. However, there are several types of relationships found within the included studies, such as university peers and married couples. It is possible that the type of relationship may have a moderating effect on collaborative group errors, which was not explored in the present meta-analysis. Additionally, group size was coded as (a) pairs or (b) trios or greater. Given that previous research has suggested that larger groups may have a greater opportunity to error-prune, it is possible that by including more group sizes in the factor coding (for example, quartets and quintets), a moderating effect may have been found. However, only a limited number of studies include groups that are larger than triads, and therefore, it was not possible to explore group size as a potential moderator in such detail in the present meta-analysis.

Future directions

To continue developing our understanding of collaborative group recall errors, future studies should first consider examining potential moderators in more detail. Specifically, research may examine whether the type of group member relationship (such as married couples vs. friends) influences collaborative group errors. Additionally, studies may further explore the impact of group size, by utilising groups larger than triads and examining whether there is a significant effect on collaborative group errors. Second, future research should investigate the role of the potential moderators examined in the present study in relation to induced errors. It is reasonable to expect a different pattern of results when considering induced errors in relation to potential moderators, given that the first meta-analysis showed that it is more difficult for collaborative groups to prune errors when they are induced. For example, Thorley and Dewhurst (2007) report that collaborative group errors increased in proportion to group size when errors were induced, however, the present analysis found that group size is not a moderator of spontaneous group errors. Examining moderators of collaborative group errors in the context of induced errors is particularly important in the context of trauma therapy applications, given the possibility of false memories being reported in clinical sessions, therefore, affecting the number of errors made by the group (Rajaram, 2011).

Conclusion

Memory is often a shared experience among groups of people, and collaborative remembering is a pervasive aspect of everyday life. The extent that collaborative group recall impacts on errors is not well-researched or understood. Here, six potential moderators were examined across two meta-analyses in relation to the number of errors collaborative groups make when recalling studied material. The results of these meta-analyses suggest that error-inducing, retrieval methods, and consensus requirements are moderators of collaborative group errors. It was also revealed that group size, group member relationship, and group member age are not moderators of collaborative group errors. This research demonstrates clear benefits of collaborative group remembering and holds potential practical applications in clinical group therapy.

References

References marked with an asterisk indicate studies included in the meta-analyses.

*Abel, M., & Bäuml, K.-H. T. (2017). Collaborative remembering revisited: Study context access modulates collaborative inhibition and later benefits for individual memory. *Memory & Cognition*, *45*(8), 1319–1334. https://doi.org/10.3758/s13421-017-0737-9

*Barber, S.J., Castrellon, J. J., Opitz, P., & Mather, M. (2017). Younger and older adults' collaborative recall of shared and unshared emotional pictures. *Memory & Cognition*, *45*(5), 716–730. https://doi.org/10.3758/s13421-017-0694-3

*Barber, S.J., & Rajaram, S. (2011). Collaborative memory and part-set cueing impairments: The role of executive depletion in modulating retrieval disruption. *Memory, 19*(4), 378–397. https://doi.org/10.1080/09658211.2011.575787

*Barber, S.J., & Rajaram, S. (2011). Exploring the relationship between retrieval disruption from collaboration and recall. *Memory*, *19*(5), 462–469. https://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*(3), 255–264. https://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*(1), 121–132. https://doi.org/10.1521/soco.2012.30.1.121

*Barnier, A.J., Harris, C. B., Morris, T., & Savage, G. (2018). Collaborative facilitation in older couples: Successful joint remembering across memory tasks. *Frontiers in Psychology*, *9*, 2385–2385. https://doi.org/10.3389/fpsyg.2018.02385

*Bärthel, G.A., Wessel, I., Huntjens, R. J. C., & Verwoerd, J. (2016). Collaboration enhances later individual memory for emotional material. *Memory*, *25*(5), 636–646. https://doi.org/10.1080/09658211.2016.1208248

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

*Basden, B.H., Basden, D. R., Thomas III, R. L., & Souphasith, S. (1998). Memory distortion in group recall. *Current Psychology*, *16*(3-4), 225–246. https://doi.org/10.1007/s12144-997-1001-4

Clark, N.K., Stephenson, G. M., & Kniveton, B. H. (1990). Social remembering: Quantitative aspects individual and collaborative remembering police officers and students. *The British Journal of Psychology*, *81*(1), 73–94. https://doi.org/10.1111/j.2044-8295.1990.tb02347.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*(4), 535–551. https://doi.org/10.1037/a0024308

*Ekeocha, J.O., & Brennan, S. E. (2008). Collaborative recall in face-to-face and electronic groups. *Memory*, *16*(3), 245–261. https://doi.org/10.1080/09658210701807480

*Finlay, F., Hitch, G., & Meudell, P. (2000). Mutual inhibition in collaborative recall: Evidence for a retrieval-based account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1556–1567. https://doi.org/10.1037//0278-7393.26.6.1556

Fraundorf, S.H. Hourihan, K. L., Peters, R. A., & Benjamin, A. S. (2019). Aging and Recognition Memory: A Meta-Analysis. *Psychological Bulletin*, *145*(4), 339–371. https://doi.org/10.1037/bul0000185

Harrer, M., Cuijpers, P., Furukawa, T. A., Ebert, D. D. (2021). Doing Meta-Analysis with R: A Hands-On Guide. United States: CRC Press.

*Harris, C.B., Barnier, A. J., & Sutton, J. (2012). Consensus collaboration enhances group and individual recall accuracy. *Quarterly Journal of Experimental Psychology (2006)*, *65*(1), 179–194. https://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*(1), 183–195. https://doi.org/10.1037/a002890

*Harris, C.B., Barnier, A. J., Sutton, J., Keil, P. G., & Dixon, R. A. (2017). “Going episodic”: collaborative inhibition and facilitation when long-married couples remember together. *Memory*, *25*(8), 1148–1159. https://doi.org/10.1080/09658211.2016.1274405

*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*(4), 267–303. https://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*(3), 532–545. https://doi.org/10.1037/a0023106

*Hinds, J.M., & Payne, S. J. (2016). Collaborative Inhibition and Semantic Recall: Improving Collaboration Through Computer-mediated Communication. *Applied Cognitive Psychology*, *30*(4), 554–565. https://doi.org/10.1002/acp.3228

*Hood, A.V., Whillock, S. R., Meade, M. L., & Hutchison, K. A. (2023). Does Collaboration Help or Hurt Recall? The Answer Depends on Working Memory Capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(3), 350–370. https://doi.org/10.1037/xlm0001155

*Hyman, I.E., Cardwell, B. A., & Roy, R. A. (2013). Multiple causes of collaborative inhibition in memory for categorised word lists. *Memory*, *21*(7), 875–890. https://doi.org/10.1080/09658211.2013.769058

*Maki, R.H., Weigold, A., & Arellano, A. (2008). False memory for associated word lists in individuals and collaborating groups. *Memory & Cognition*, *36*(3), 598–603. https://doi.org/10.3758/MC.36.3.598

Marion, S., & Thorley, C. (2016). A Meta-Analytic Review of Collaborative Inhibition and Post collaborative Memory: Testing the Predictions of the Retrieval Strategy Disruption Hypothesis. *Psychological Bulletin*, *142*(11), 1141–1164. https://doi.org/10.1037/bul0000071

*Maswood, R., Luhmann, C. C., & Rajaram, S. (2022). Persistence of false memories and emergence of collective false memory: collaborative recall of DRM word lists. *Memory*, *30*(4), 465–479. https://doi.org/10.1080/09658211.2021.1928222

Maswood, R., & Rajaram, S. (2019). Social Transmission of False Memory in Small Groups and Large Networks. *Topics in Cognitive Science*, *11*(4), 687–709. https://doi.org/10.1111/tops.12348

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

*Nie, A., & Deng, C. (2023). Detrimental and beneficial effects in ongoing and lasting collaborative memory: Insight from the emotional timeout procedure. *Advances in Cognitive Psychology*, *19*(1), 59–79. https://doi.org/10.5709/acp-0376-4

*Nie, A., & Guo, B. (2023). Benefits and detriments of social collaborative memory in turn-taking and directed forgetting. *Perceptual and Motor Skills*, *130*(3), 1040–1076. https://doi.org/10.1177/00315125231163626

*Nie, A., Ke, C., Guo, B., Li, M., & Xiao, Y. (2023). Collaborative memory for categorized lists: ongoing and lasting effects are sensitive to episodic memory tasks. *Current Psychology (New Brunswick, N.J.)*, *42*(5), 3870–3887. https://doi.org/10.1007/s12144-021-01684-w

*Nie, A., Ke, C., Li, M., & Guo, B. (2019). Disrupters as well as monitors: Roles of others during and after collaborative remembering in the DRM procedure. *Advances in Cognitive Psychology*, *15*(4), 276–289. https://doi.org/10.5709/acp-0275-1

*Nie, A., Li, M., Li, M., Xiao, Y., & Wang, S. (2022). Together we lose or gain: Ongoing and enduring impacts of collaboration in episodic memory of emotional DRM lists. *Current Psychology*. https://doi.org/10.1007/s12144-022-03940-z

Old, S.R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*(1), 104–118. https://doi.org/10.1037/0882-7974.23.1.104

Pardilla-Delgado, E., & Payne, J. D. (2017). The Deese-Roediger-Mcdermott (DRM) task: A simple cognitive paradigm to investigate false memories in the laboratory. *Journal of Visualized Experiments*, *2017*(119). https://doi.org/10.3791/54793

*Pepe, N.W., Wang, Q., & Rajaram, S. (2021). Collaborative remembering in ethnically uniform and diverse group settings. *Journal of Applied Research in Memory and Cognition*, *10*(1), 95–103. https://doi.org/10.1016/j.jarmac.2020.08.001

*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*(6), 968–976. https://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: A Journal of the American Psychological Society*, *20*(2), 76–81. https://doi.org/10.1177/0963721411403251

*Rivardo, M.G., Rutledge, A. T., Chelecki, C., Stayer, B. E., Quarles, M., & Kline, A. (2013). Collaborative recall of eyewitness event increases misinformation effect at 1 week. *North American Journal of Psychology*, *15*(3), 495–512.

Roediger, H.L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

*Ross, M., Spencer, S. J., Blatz, C. W., & Restorick, E. (2008). Collaboration reduces the frequency of false memories in older and younger adults. *Psychology and Aging*, *23*(1), 85–92. https://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*(6), 683–696. https://doi.org/10.1002/acp.1023

*Rossi-Arnaud, C., Spataro, P., Bhatia, D., & Cestari, V. (2019). Collaborative remembering reduces suggestibility: A study with the Gudjonsson Suggestibility Scale. *Memory*, *27*(5), 603–611. https://doi.org/10.1080/09658211.2018.1542004

*Rossi-Arnaud, C., Spataro, P., Bhatia, D., Doricchi, F., Mastroberardino, S., & Cestari, V. (2020). Long-lasting positive effects of collaborative remembering on false assents to misleading questions. *Acta Psychologica*, *203*, 102986–102986. https://doi.org/10.1016/j.actpsy.2019.102986

*Rossi-Arnaud, C., Spataro, P., Mastroberardino, S., Lucaroni, E., Giannini, A. M., & Cestari, V. (2023). Why collaboration reduces suggestibility: The role of source-monitoring processes and retrieval strategies. *Current Psychology*, *42*(8), 6386–6394. https://doi.org/10.1007/s12144-021-01970-7

*Saraiva, M., Albuquerque, P. B., & Arantes, J. (2017). Production of false memories in collaborative memory tasks using the DRM paradigm. *Psicológica*, *38*(2), 209–229.

*Saraiva, M., Albuquerque, P. B., & Garrido, M. V. (2023). Collaborative inhibition effect: the role of memory task and retrieval method. *Psychological Research*. https://doi.org/10.1007/s00426-023-01821-z

*Saraiva, M., Garrido, M. V., & Albuquerque, P. B. (2021). Emergence and transmission of misinformation in the context of social interactions. *Memory & Cognition*, *49*(1), 1–13. https://doi.org/10.3758/s13421-020-01081-x

Schober, P., Mascha, E. J., & Vetter, T. R. (2021). Statistics from a guide to interpreting common measures of association, agreement, diagnostic accuracy, effect size, heterogeneity, and reliability in medical research. *Anesthesia and Analgesia*, *133*(6), 1633–. https://doi.org/10.1213/ANE.0000000000005773

Stadler, M.A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, *27*(3), 494–500. https://doi.org/10.3758/BF03211543

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

*Takahashi, M., & Saito, S. (2004). Does test delay eliminate collaborative inhibition? *Memory*, *12*(6), 722–731. https://doi.org/10.1080/09658210344000521

*Thorley, C. (2007). *Collaborative false remembering* [Unpublished doctoral dissertation]. University of Lancaster.

*Thorley, C., (2018). Enhancing individual and collaborative eyewitness memory with category clustering recall. *Memory*, *26*(8), 1128–1139. https://doi.org/10.1080/09658211.2018.1432058

*Thorley, C., & Dewhurst, S. A. (2007). Collaborative false recall in the DRM procedure: Effects of group size and group pressure. *European Journal of Cognitive Psychology*, *19*(6), 867–881. https://doi.org/10.1080/09541440600872068

Thorley, C., & Dewhurst, S. A. (2009). False and veridical collaborative recognition. *Memory*, *17*(1), 17–25. https://doi.org/10.1080/09658210802484817

*Vredeveldt, A., Groen, R. N., Ampt, J. E., & Koppen, P. J. (2017). When discussion between eyewitnesses helps memory. *Legal and Criminological Psychology*, *22*(2), 242–259. https://doi.org/10.1111/lcrp.12097

*Vredeveldt, A., Hildebrandt, A., & van Koppen, P. J. (2016). Acknowledge, repeat, rephrase, elaborate: Witnesses can help each other remember more. *Memory*, *24*(5), 669–682. https://doi.org/10.1080/09658211.2015.1042884

*Vredeveldt, A., Kesteloo, L., & van Koppen, P. J. (2018). Writing alone or together: Police officers' collaborative reports of an incident. *Criminal Justice and Behavior*, *45*(7), 1071–1092. https://doi.org/10.1177/0093854818771721

*Vredeveldt, A., van Deuren, S., & van Koppen, P. J. (2019). Remembering with a friend or a stranger: comparing acquainted and unacquainted pairs in collaborative eyewitness interviews. *Memory*, *27*(10), 1390–1403. https://doi.org/10.1080/09658211.2019.1662052

*Vredeveldt, A., & van Koppen, P. J. (2018). Recounting a common experience: On the effectiveness of instructing eyewitness Pairs. *Frontiers in Psychology*, *9*(March), 1–11. https://doi.org/10.3389/fpsyg.2018.00284

*Webster, V. (2017). *Collaborative memory: the role of closeness, cognitive need, and strategies* [Unpublished doctoral dissertation]. Macquarie University.

*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*(2), 183–190. https://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*(5), 1160–1175. https://doi.org/10.1037/0278-7393.23.5.1160

Wessel, I., & Moulds, M. L. (2008). Collective memory: A perspective from (experimental) clinical psychology. *Memory (Hove)*, *16*(3), 288–304. https://doi.org/10.1080/09658210701811813

*Wessel, I., Zandstra, A. R. E., Hengeveld, H. M. E., & Moulds, M. L. (2015). Collaborative recall of details of an emotional film. *Memory*, *23*(3), 437–444. https://doi.org/10.1080/09658211.2014.895384

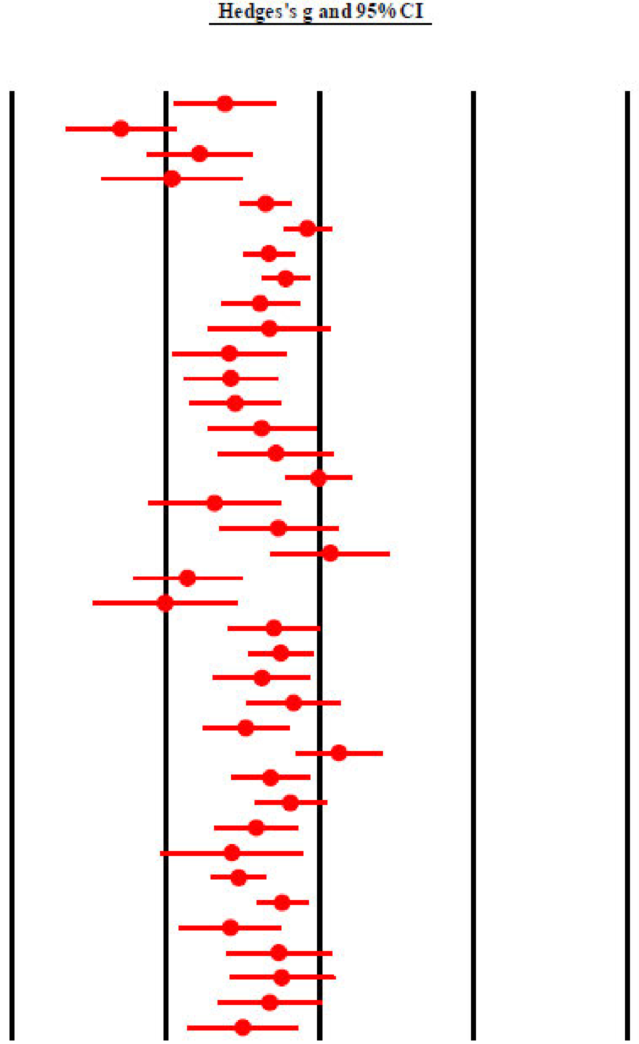
*Whillock, S.R., Meade, M. L., Hutchison, K. A., & Tsosie, M. D. (2020). Collaborative inhibition in same-age and mixed-age dyads. *Psychology and Aging*, *35*(7), 963–973. https://doi.org/10.1037/pag0000490

*Wissman, K.T., & Rawson, K. A. (2015). Why does collaborative retrieval improve memory? Enhanced relational and item-specific processing. *Journal of Memory and Language*, *84*, 75–87. https://doi.org/10.1016/j.jml.2015.05.003

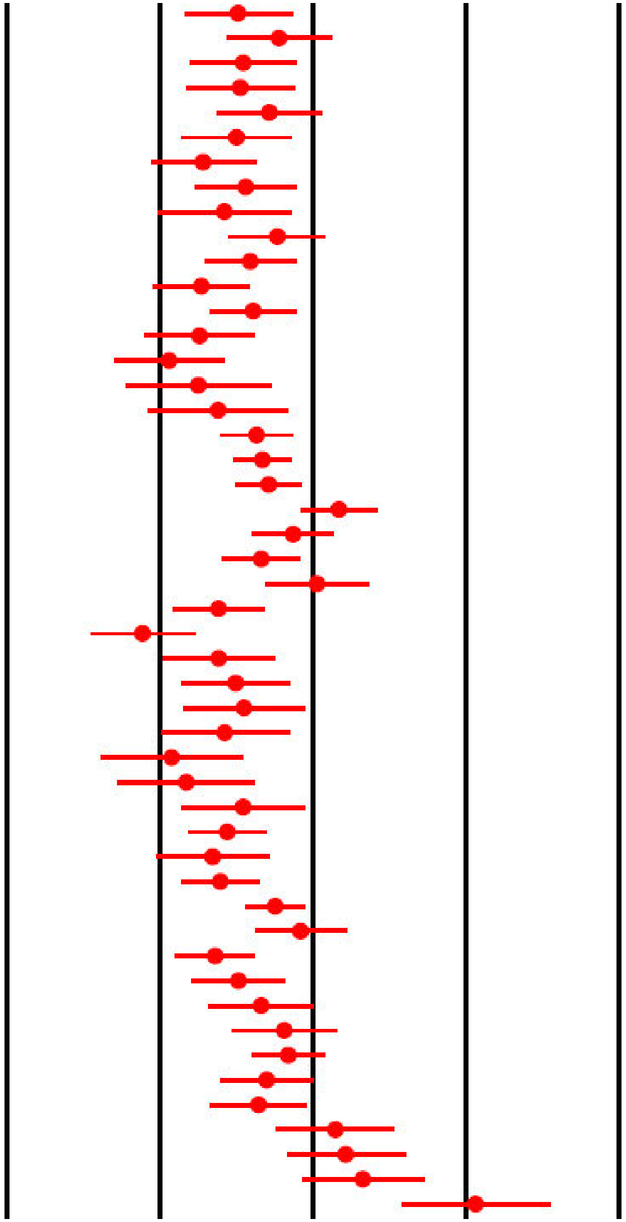
Appendix A

Meta-Analysis 1: Induced vs Spontaneous Collaborative Group Errors Forest Plot

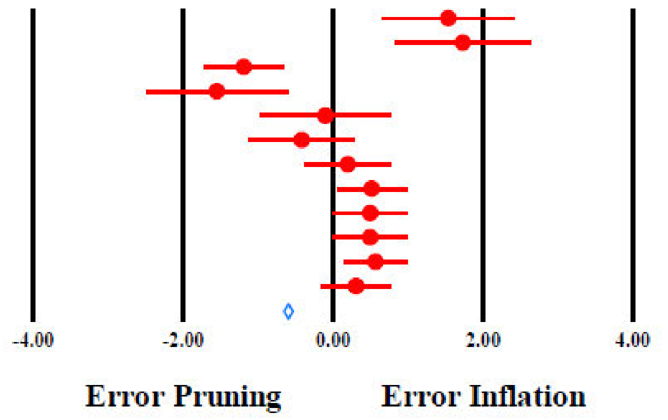
Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Nie & Deng (2023)	-1.225	0.340	0.115	-1.891	-0.560	-3.608	0.000
Nie & Guo (2023)	-2.581	0.369	0.136	-3.303	-1.858	-7.002	0.000
Nie et al. (2023)	-1.558	0.353	0.125	-2.251	-0.866	-4.410	0.000
Rossi et al. (2023)	-1.915	0.474	0.224	-2.843	-0.986	-4.042	0.000
Saraiva et al. (2023) - Exp. 1 free recall	-0.697	0.178	0.032	-1.046	-0.348	-3.918	0.000
Saraiva et al. (2023) - Exp. 1 serial recall	-0.155	0.160	0.026	-0.469	0.158	-0.972	0.331
Saraiva et al. (2023) - Exp. 2 free recall	-0.655	0.176	0.031	-0.999	-0.310	-3.725	0.000
Saraiva et al. (2023) - Exp. 2 serial recall	-0.436	0.167	0.028	-0.762	-0.109	-2.615	0.009
Hood et al. (2022)	-0.770	0.261	0.068	-1.282	-0.258	-2.948	0.003
Maswood et al. (2022) - TT	-0.647	0.410	0.168	-1.450	0.156	-1.579	0.114
Nie et al. (2022) - Critical lures	-1.170	0.378	0.143	-1.910	-0.430	-3.098	0.002
Pepe et al. (2021) - Diverse Groups	-1.151	0.311	0.097	-1.760	-0.541	-3.700	0.000
Pepe et al. (2021) - Uniform Groups	-1.092	0.311	0.097	-1.701	-0.482	-3.510	0.000
Rossi-Arnaud et al. (2020) - Exp. 1	-0.749	0.364	0.133	-1.463	-0.036	-2.058	0.040
Rossi-Arnaud et al. (2020) - Exp. 2	-0.562	0.387	0.150	-1.321	0.198	-1.450	0.147
Saraiva et al. (2020) - Exp. 1	-0.010	0.221	0.049	-0.444	0.424	-0.045	0.964
Whillock et al. (2020) - Younger groups	-1.361	0.442	0.196	-2.229	-0.494	-3.077	0.002
Whillock et al. (2020) - Mixed-age groups	-0.531	0.398	0.158	-1.311	0.249	-1.334	0.182
Whillock et al. (2020) - Older groups	0.145	0.398	0.158	-0.635	0.925	0.364	0.716
Nie et al. (2019) - Critical lures	-1.715	0.367	0.134	-2.434	-0.996	-4.677	0.000
Rossi-Arnaud et al. (2019)	-2.005	0.484	0.234	-2.954	-1.056	-4.140	0.000
Vredeveldt et al. (2019) - Acquainted pairs	-0.589	0.310	0.096	-1.197	0.019	-1.897	0.058
Barnier et al. (2018)	-0.502	0.217	0.047	-0.926	-0.077	-2.317	0.020
Thorley (2018) - Free recall	-0.745	0.325	0.106	-1.382	-0.108	-2.291	0.022
Thorley (2018) - Category clustering recall	-0.333	0.310	0.096	-0.941	0.274	-1.075	0.282
Vredeveldt & van Koppen (2018) - No instructions	-0.955	0.295	0.087	-1.534	-0.376	-3.233	0.001
Vredeveldt et al. (2018)	0.255	0.294	0.087	-0.322	0.832	0.867	0.386
Abel & Bauml (2017) - Exp. 2 remember instructions	-0.632	0.261	0.068	-1.144	-0.120	-2.419	0.016
Barber et al. (2017) - Younger adults	-0.375	0.242	0.059	-0.850	0.099	-1.551	0.121
Barber et al. (2017) - Older adults	-0.819	0.279	0.078	-1.365	-0.272	-2.934	0.003
Harris et al. (2017)	-1.137	0.478	0.228	-2.073	-0.201	-2.380	0.017
Saraiva et al. (2017) - Exp. 1	-1.047	0.188	0.035	-1.414	-0.679	-5.582	0.000
Saraiva et al. (2017) - Exp. 2	-0.483	0.173	0.030	-0.821	-0.144	-2.795	0.005
Vredeveldt et al. (2017)	-1.157	0.340	0.115	-1.822	-0.491	-3.406	0.001
Webster (2017) - Exp. 1 not divided or close	-0.526	0.351	0.124	-1.215	0.162	-1.498	0.134
Webster (2017) - Exp. 1 not divided but close	-0.487	0.351	0.124	-1.176	0.201	-1.387	0.166
Webster (2017) - Exp. 1 divided not close	-0.643	0.351	0.124	-1.332	0.045	-1.831	0.067
Webster (2017) - Exp. 1 divided and close	-0.994	0.365	0.133	-1.709	-0.279	-2.726	0.006



Webster (2017) - Exp. 2 not divided	-0.975	0.365	0.133	-1.690	-0.260	-2.673	0.008
Webster (2017) - Exp. 2 divided	-0.439	0.351	0.124	-1.128	0.250	-1.248	0.212
Webster (2017) - Exp. 3 strategy	-0.907	0.365	0.133	-1.621	-0.192	-2.486	0.013
Webster (2017) - Exp. 3 no strategy	-0.946	0.365	0.133	-1.660	-0.231	-2.592	0.010
Webster (2017) - Exp. 4 no strategy	-0.565	0.351	0.124	-1.254	0.123	-1.609	0.108
Webster (2017) - Exp. 4 strategy	-0.994	0.365	0.133	-1.709	-0.279	-2.726	0.006
Bärthel et al. (2016)	-1.431	0.353	0.125	-2.124	-0.738	-4.049	0.000
Hinds & Payne (2016) - Exp. 1 episodic recall	-0.875	0.339	0.115	-1.540	-0.210	-2.579	0.010
Vredeveldt et al. (2016)	-1.154	0.448	0.201	-2.032	-0.276	-2.575	0.010
Wessel et al. (2015)	-0.460	0.325	0.105	-1.096	0.176	-1.417	0.156
Wissman & Rawson (2015) - Exp. 1	-0.815	0.311	0.097	-1.424	-0.207	-2.625	0.009
Wissman & Rawson (2015) - Exp. 2	-1.455	0.326	0.106	-2.094	-0.816	-4.462	0.000
Wissman & Rawson (2015) - Exp. 3	-0.778	0.295	0.087	-1.356	-0.199	-2.633	0.008
Weigold et al. (2014)	-1.478	0.366	0.134	-2.196	-0.760	-4.036	0.000
Harris et al. (2013) - Exp. 1	-1.875	0.367	0.135	-2.595	-1.155	-5.105	0.000
Harris et al. (2013) - Exp. 2 shared encoding	-1.494	0.488	0.238	-2.451	-0.537	-3.059	0.002
Harris et al. (2013) - Exp. 2 unshared encoding	-1.235	0.469	0.220	-2.155	-0.316	-2.633	0.008
Hyman et al. (2013) - Exp. 1	-0.731	0.242	0.059	-1.206	-0.257	-3.021	0.003
Hyman et al. (2013) - Exp. 2	-0.655	0.199	0.039	-1.045	-0.266	-3.300	0.001
Hyman et al. (2013) - Exp. 3	-0.575	0.222	0.049	-1.009	-0.141	-2.594	0.009
Rivardo et al. (2013) - Misinformation groups	0.345	0.261	0.068	-0.166	0.856	1.323	0.186
Rivardo et al. (2013) - No misinformation groups	-0.256	0.278	0.078	-0.802	0.290	-0.919	0.358
Barber et al. (2012)	-0.671	0.261	0.068	-1.183	-0.159	-2.570	0.010
Harris et al. (2012) - No consensus	0.058	0.351	0.123	-0.629	0.746	0.166	0.868
Barber & Rajaram (2011) - Identical organisation	-1.230	0.311	0.097	-1.840	-0.620	-3.953	0.000
Barber & Rajaram (2011) - Divergent organisation	-2.225	0.355	0.126	-2.920	-1.529	-6.268	0.000
Barber & Rajaram (2011) - Exp. 2 executive depletion	-1.228	0.378	0.143	-1.968	-0.488	-3.253	0.001
Barber & Rajaram (2011) - Exp. 2 no executive depletion	-1.004	0.365	0.133	-1.719	-0.289	-2.753	0.006
Congleton & Rajaram (2011) - Short delay/repeated study	-0.898	0.410	0.168	-1.701	-0.095	-2.192	0.028
Congleton & Rajaram (2011) - Short delay/repeated retrieval	-1.149	0.432	0.186	-1.995	-0.303	-2.661	0.008
Congleton & Rajaram (2011) - Long delay/repeated study	-1.844	0.473	0.224	-2.771	-0.917	-3.899	0.000
Congleton & Rajaram (2011) - Long delay/repeated retrieval	-1.651	0.463	0.214	-2.559	-0.743	-3.566	0.000
Harris et al. (2011)	-0.908	0.413	0.170	-1.717	-0.099	-2.200	0.028
Henkel & Rajaram (2011)	-1.116	0.261	0.068	-1.629	-0.604	-4.271	0.000
Pereira-Pasarin & Rajaram (2011) - Exp. 1	-1.306	0.378	0.143	-2.046	-0.566	-3.460	0.001
Pereira-Pasarin & Rajaram (2011) - Exp. 2	-1.205	0.261	0.068	-1.717	-0.693	-4.611	0.000
Barber et al. (2010) - Exp. 1	-0.487	0.199	0.039	-0.876	-0.097	-2.450	0.014
Meade & Roediger (2009) - Exp. 1	-0.157	0.310	0.096	-0.764	0.451	-0.506	0.613
Ekeocha & Brennan (2008)	-1.274	0.270	0.073	-1.803	-0.746	-4.725	0.000
Maki et al. (2008) - Exp 1	-0.969	0.311	0.097	-1.579	-0.359	-3.114	0.002
Ross et al. (2008) - Young adults critical lures	-0.673	0.351	0.124	-1.361	0.016	-1.914	0.056
Ross et al. (2008) - Older adults critical lures	-0.370	0.351	0.124	-1.059	0.318	-1.054	0.292
Takahashi (2007) - Exp. 1	-0.316	0.242	0.059	-0.791	0.158	-1.306	0.191
Takahashi (2007) - Exp. 2 friends	-0.599	0.311	0.096	-1.208	0.010	-1.929	0.054
Takahashi (2007) - Exp. 3	-0.706	0.325	0.106	-1.343	-0.069	-2.171	0.030
Thorley (2007) - Exp. 2 TT pairs critical lures	0.299	0.398	0.158	-0.481	1.080	0.752	0.452
Thorley (2007) - Exp. 2 TT trios critical lures	0.434	0.398	0.158	-0.346	1.215	1.091	0.275
Thorley (2007) - Exp. 2 TT quartets critical lures	0.657	0.410	0.168	-0.146	1.459	1.603	0.109
Thorley & Dewhurst (2007) - TT pairs critical lures	2.134	0.502	0.252	1.150	3.117	4.253	0.000



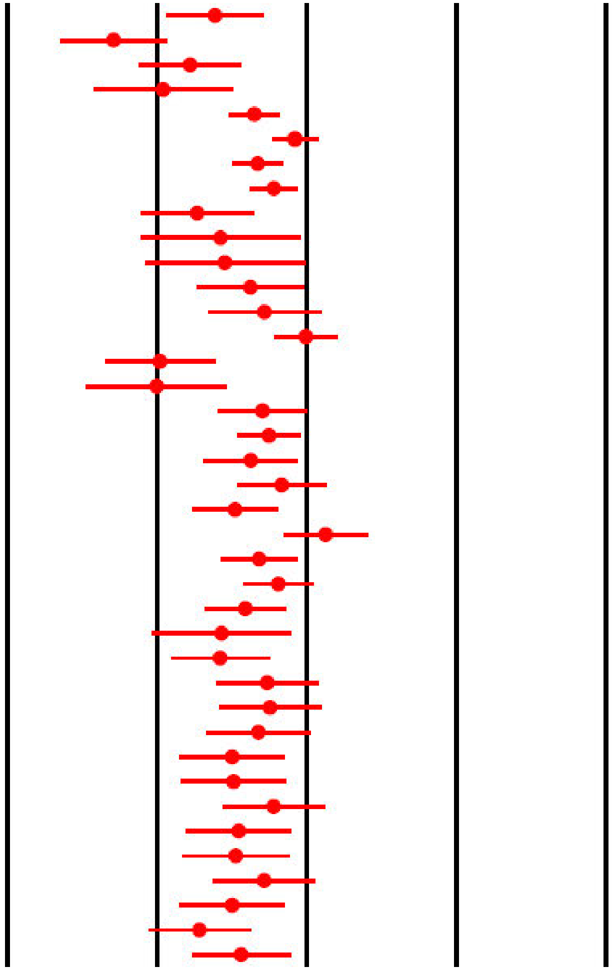
Thorley & Dewhurst (2007) - TT Trios critical lures	1.545	0.453	0.205	0.657	2.432	3.411	0.001
Thorley & Dewhurst (2007) - TT quartets critical lures	1.738	0.463	0.214	0.830	2.645	3.753	0.000
Ross et al. (2004) - Shopping cart false positives	-1.184	0.279	0.078	-1.731	-0.637	-4.243	0.000
Takahashi & Saito (2004) - Exp. 1	-1.542	0.488	0.238	-2.499	-0.585	-3.157	0.002
Takahashi & Saito (2004) - Exp. 2	-0.095	0.447	0.200	-0.971	0.780	-0.213	0.831
Finlay et al. (2000) - Exp. 3	-0.409	0.364	0.133	-1.122	0.305	-1.122	0.262
Basden et al. (1998) - Exp. 1 critical lures	0.206	0.295	0.087	-0.371	0.784	0.700	0.484
Basden et al. (1998) - Exp. 2 critical lures	0.524	0.242	0.059	0.049	0.998	2.164	0.030
Basden et al. (1997) - Exp. 1	0.503	0.261	0.068	-0.008	1.015	1.928	0.054
Basden et al. (1997) - Exp. 2	0.504	0.261	0.068	-0.008	1.016	1.928	0.054
Basden et al. (1997) - Exp. 3	0.575	0.222	0.049	0.140	1.009	2.594	0.009
Weldon & Bellinger (2007) - Exp. 1	0.316	0.242	0.059	-0.158	0.790	1.306	0.191
	-0.598	0.030	0.001	-0.657	-0.540	-19.995	0.000



Appendix B

Meta-Analysis 2: Spontaneous Collaborative Group Errors Forest Plot

Study name	Statistics for each study							Hedges's g and 95% CI
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value	
Nie & Deng (2023)	-1.225	0.340	0.115	-1.891	-0.560	-3.608	0.000	
Nie & Guo (2023)	-2.581	0.369	0.136	-3.303	-1.858	-7.002	0.000	
Nie et al. (2023)	-1.558	0.353	0.125	-2.251	-0.866	-4.410	0.000	
Rossi-Arnaud et al. (2023)	-1.915	0.474	0.224	-2.843	-0.986	-4.042	0.000	
Saraiva et al. (2023) - Exp. 1 free recall	-0.697	0.178	0.032	-1.046	-0.348	-3.918	0.000	
Saraiva et al. (2023) - Exp. 1 serial recall	-0.155	0.160	0.026	-0.469	0.158	-0.972	0.331	
Saraiva et al. (2023) - Exp. 2 free recall	-0.655	0.176	0.031	-0.999	-0.310	-3.725	0.000	
Saraiva et al. (2023) - Exp. 2 serial recall	-0.436	0.167	0.028	-0.762	-0.109	-2.615	0.009	
Nie et al. (2022) - General errors	-1.462	0.390	0.152	-2.226	-0.698	-3.750	0.000	
Pepe et al. (2021) - Diverse Groups	-1.151	0.548	0.300	-2.224	-0.077	-2.101	0.036	
Pepe et al. (2021) - Uniform Groups	-1.092	0.548	0.300	-2.165	-0.018	-1.994	0.046	
Rossi-Arnaud et al. (2020) - Exp. 1	-0.749	0.364	0.133	-1.463	-0.036	-2.058	0.040	
Rossi-Arnaud et al. (2020) - Exp. 2	-0.562	0.387	0.150	-1.321	0.198	-1.450	0.147	
Saraiva et al. (2020) - Exp. 1	-0.010	0.221	0.049	-0.444	0.424	-0.045	0.964	
Nie et al. (2019) - General errors	-1.960	0.380	0.144	-2.704	-1.216	-5.164	0.000	
Rossi-Arnaud et al. (2019)	-2.005	0.484	0.234	-2.954	-1.056	-4.140	0.000	
Vredeveldt et al. (2019) - Acquainted pairs	-0.589	0.310	0.096	-1.197	0.019	-1.897	0.058	
Barnier et al. (2018)	-0.502	0.217	0.047	-0.926	-0.077	-2.317	0.020	
Thorley (2018) - Free recall	-0.745	0.325	0.106	-1.382	-0.108	-2.291	0.022	
Thorley (2018) - Category clustering recall	-0.333	0.310	0.096	-0.941	0.274	-1.075	0.282	
Vredeveldt & van Koppen (2018) - No instructions	-0.955	0.295	0.087	-1.534	-0.376	-3.233	0.001	
Vredeveldt et al. (2018)	0.255	0.294	0.087	-0.322	0.832	0.867	0.386	
Abel & Bauml (2017) - Exp. 2 remember instructions	-0.632	0.261	0.068	-1.144	-0.120	-2.419	0.016	
Barber et al. (2017) - Younger adults	-0.375	0.242	0.059	-0.850	0.099	-1.551	0.121	
Barber et al. (2017) - Older adults	-0.819	0.279	0.078	-1.365	-0.272	-2.934	0.003	
Harris et al. (2017)	-1.137	0.478	0.228	-2.073	-0.201	-2.380	0.017	
Vredeveldt et al. (2017)	-1.157	0.340	0.115	-1.822	-0.491	-3.406	0.001	
Webster (2017) - Exp. 1 not divided or close	-0.526	0.351	0.124	-1.215	0.162	-1.498	0.134	
Webster (2017) - Exp. 1 not divided but close	-0.487	0.351	0.124	-1.176	0.201	-1.387	0.166	
Webster (2017) - Exp. 1 divided not close	-0.643	0.351	0.124	-1.332	0.045	-1.831	0.067	
Webster (2017) - Exp. 1 divided and close	-0.994	0.365	0.133	-1.709	-0.279	-2.726	0.006	
Webster (2017) - Exp. 2 not divided	-0.975	0.365	0.133	-1.690	-0.260	-2.673	0.008	
Webster (2017) - Exp. 2 divided	-0.439	0.351	0.124	-1.128	0.250	-1.248	0.212	
Webster (2017) - Exp. 3 strategy	-0.907	0.365	0.133	-1.621	-0.192	-2.486	0.013	
Webster (2017) - Exp. 3 no strategy	-0.946	0.365	0.133	-1.660	-0.231	-2.592	0.010	
Webster (2017) - Exp. 4 no strategy	-0.565	0.351	0.124	-1.254	0.123	-1.609	0.108	
Webster (2017) - Exp. 4 strategy	-0.994	0.365	0.133	-1.709	-0.279	-2.726	0.006	
Bärthel et al. (2016)	-1.431	0.353	0.125	-2.124	-0.738	-4.049	0.000	
Hinds & Payne (2016) - Exp. 1 episodic recall	-0.875	0.339	0.115	-1.540	-0.210	-2.579	0.010	



Vredeveldt et al. (2016)	-1.154	0.448	0.201	-2.032	-0.276	-2.575	0.010
Wessel et al. (2015)	-0.460	0.325	0.105	-1.096	0.176	-1.417	0.156
Wissman & Rawson (2015) - Exp. 1	-0.815	0.311	0.097	-1.424	-0.207	-2.625	0.009
Wissman & Rawson (2015) - Exp. 2	-1.455	0.326	0.106	-2.094	-0.816	-4.462	0.000
Wissman & Rawson (2015) - Exp. 3	-0.778	0.295	0.087	-1.356	-0.199	-2.633	0.008
Harris et al. (2013) - Exp. 1	-1.875	0.367	0.135	-2.595	-1.155	-5.105	0.000
Harris et al. (2013) - Exp. 2 shared encoding	-1.494	0.488	0.238	-2.451	-0.537	-3.059	0.002
Harris et al. (2013) - Exp. 2 unshared encoding	-1.235	0.469	0.220	-2.155	-0.316	-2.633	0.008
Hyman et al. (2013) - Exp. 1	-0.731	0.242	0.059	-1.206	-0.257	-3.021	0.003
Hyman et al. (2013) - Exp. 2	-0.655	0.199	0.039	-1.045	-0.266	-3.300	0.001
Hyman et al. (2013) - Exp. 3	-0.575	0.222	0.049	-1.009	-0.141	-2.594	0.009
Rivardo et al. (2013) - No misinformation groups	-0.256	0.278	0.078	-0.802	0.290	-0.919	0.358
Barber et al. (2012)	-0.671	0.261	0.068	-1.183	-0.159	-2.570	0.010
Harris et al. (2012) - Consensus	-0.895	0.377	0.142	-1.634	-0.157	-2.375	0.018
Barber & Rajaram (2011) - Identical organisation	-1.230	0.311	0.097	-1.840	-0.620	-3.953	0.000
Barber & Rajaram (2011) - Divergent organisation	-2.225	0.355	0.126	-2.920	-1.529	-6.268	0.000
Barber & Rajaram (2011) - Exp. 2 executive depletion	-1.228	0.378	0.143	-1.968	-0.488	-3.253	0.001
Barber & Rajaram (2011) - Exp. 2 no executive depletion	-1.004	0.365	0.133	-1.719	-0.289	-2.753	0.006
Congleton & Rajaram (2011) - Short delay/repeated study	-0.898	0.410	0.168	-1.701	-0.095	-2.192	0.028
Congleton & Rajaram (2011) - Short delay/repeated retrieval	-1.149	0.432	0.186	-1.995	-0.303	-2.661	0.008
Congleton & Rajaram (2011) - Long delay/repeated study	-1.844	0.473	0.224	-2.771	-0.917	-3.899	0.000
Congleton & Rajaram (2011) - Long delay/repeated retrieval	-1.651	0.463	0.214	-2.559	-0.743	-3.566	0.000
Harris et al. (2011)	-0.908	0.413	0.170	-1.717	-0.099	-2.200	0.028
Pereira-Pasarin & Rajaram (2011) - Exp. 1	-1.306	0.378	0.143	-2.046	-0.566	-3.460	0.001
Pereira-Pasarin & Rajaram (2011) - Exp. 2	-1.205	0.261	0.068	-1.717	-0.693	-4.611	0.000
Barber et al. (2010) - Exp. 1	-0.487	0.199	0.039	-0.876	-0.097	-2.450	0.014
Ekeocha & Brennan (2008)	-1.274	0.270	0.073	-1.803	-0.746	-4.725	0.000
Ross et al. (2008) - Younger adults self-generated errors	-1.014	0.365	0.133	-1.729	-0.299	-2.780	0.005
Ross et al. (2008) - Older adults self-generated errors	-1.930	0.425	0.181	-2.763	-1.097	-4.542	0.000
Thorley (2007) - Exp. 2 TT pairs general errors	0.830	0.410	0.168	0.027	1.633	2.027	0.043
Thorley (2007) - Exp. 2 TT trios general errors	1.361	0.442	0.196	0.494	2.229	3.077	0.002
Thorley (2007) - Exp. 2 TT quartets general errors	1.101	0.421	0.177	0.276	1.926	2.615	0.009
Thorley & Dewhurst (2007) - TT pairs general errors	1.255	0.432	0.186	0.409	2.101	2.907	0.004
Thorley & Dewhurst (2007) - TT Trios general errors	2.008	0.492	0.242	1.043	2.973	4.079	0.000
Thorley & Dewhurst (2007) - TT quartets general errors	1.970	0.483	0.233	1.023	2.916	4.080	0.000
Ross et al. (2004) - Shopping cart false positives	-1.184	0.279	0.078	-1.731	-0.637	-4.243	0.000
Takahashi & Saito (2004) - Exp. 1	-1.542	0.488	0.238	-2.499	-0.585	-3.157	0.002
Takahashi & Saito (2004) - Exp. 2	-0.095	0.447	0.200	-0.971	0.780	-0.213	0.831
Finlay et al. (2000) - Exp. 3	-0.409	0.364	0.133	-1.122	0.305	-1.122	0.262
Basden et al. (1997) - Exp. 1	0.503	0.261	0.068	-0.008	1.015	1.928	0.054
Basden et al. (1997) - Exp. 2	0.504	0.261	0.068	-0.008	1.016	1.928	0.054
Basden et al. (1997) - Exp. 3	0.575	0.222	0.049	0.140	1.009	2.594	0.009
Weldon & Bellinger (2007) - Exp. 1	0.316	0.242	0.059	-0.158	0.790	1.306	0.191
	-0.631	0.034	0.001	-0.698	-0.565	-18.655	0.000