Neuroscience and Biobehavioral Reviews 149 (2023) 105157

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

An updated account of overgeneral autobiographical memory in depression

Samuel Weiss-Cowie [a,b,*], Paul Verhaeghen [a], Audrey Duarte [c]

[a] School of Psychology, Georgia Institute of Technology, Atlanta, GA, USA
[b] MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK
[c] Department of Psychology, University of Texas at Austin, Austin, TX, USA

ARTICLE INFO

Keywords:
Major depression
Subthreshold depression
Overgeneral autobiographical memory
Meta-analysis

ABSTRACT

Previous meta-analyses on Overgeneral Autobiographical Memory (OGM) and depression have emphasised clinically diagnosed current depression, leaving questions about subthreshold and remitted depression. Further, numerous studies of OGM remain unconsidered due to a focus on one testing paradigm, the Autobiographical Memory Test (AMT). We conducted a meta-analysis on OGM in depression including remitted, subthreshold, and currently depressed samples and incorporating non-AMT studies. Our novel use of three-level models enabled robust variance analyses with multiple effect sizes from each study while controlling for dependencies across effect sizes. With results from 67 published and unpublished works, ours is the largest meta-analysis to date on OGM in depression. We identified decreased autobiographical memory specificity (Hedges' $g = -0.73$) and increased categoricity (Hedges' $g = 0.77$) for depressed individuals over controls. Moderator analyses suggested more severe OGM in current, clinical MDD than subthreshold and remitted depression, although deficits were still present in the latter groups. Our results highlight the importance of utilising a broader range of testing paradigms and considering non-clinical depression in future work.

1. Introduction

Autobiographical memory involves the storage and retrieval of information from one's past and ranges from broad life periods down to the minute sensory details of a given event (Conway and Pleydell-Pearce, 2000). The successful recollection of events from one's past is critical to identity formation, problem-solving, and future goal direction (Gamble et al., 2019; Liu et al., 2013). The tendency to recall elements from autobiographical memory with lower specificity is associated with a range of poor outcomes, such as decreased social problem solving (Goddard et al., 1997) and worse depression prognosis (Hallford et al., 2021; Sumner et al., 2010). One of the most highly researched phenomena in this field is that of overgeneral autobiographical memory (OGM) in major depressive disorder (MDD).

OGM refers to the tendency for individuals prompted to recall a specific personal event to instead provide a *categoric(al)* or *extended* response, defined as, respectively, a memory for a more general event that has occurred multiple times or a memory for events lasting longer than a day. Depressed individuals are especially prone towards the categoric response style (Williams et al., 2007). For instance, a depressed individual asked to report a memory related to the word "school" might give the categorical response "taking exams," rather than a specific response like "taking my organic chemistry final" or the extended response "my freshman year of college."

The instrument most commonly used to elicit autobiographical memories is the Autobiographical Memory Test (AMT; Williams and Broadbent, 1986). In the standard AMT, an experimenter presents the participant with a series of positively, negatively, or neutrally valenced cue words and asks for a specific memory in response to those cues. The responses are typically categorised as specific, categoric, or extended. While the core cuing paradigm of the AMT does not vary across studies, there are several methodological variables that can differ (e.g., number of cues presented, time limit for memory response, cue valence, etc.). Results from the AMT have formed the basis of all meta-analyses to date on OGM in depression (Chiu et al., 2018; Hallford et al., 2021; Liu et al., 2013; Ono et al., 2016; Sumner et al., 2010; Van Vreeswijk and De Wilde, 2004). The stated argument to exclude non-AMT paradigms from meta-analyses typically cites heterogeneity of effect sizes introduced by including other testing paradigms. However, two lines of evidence suggest that this may not be as useful a restriction as previously thought.

First, it seems unnecessary to exclude non-AMT paradigms in the name of homogeneity when variations within the AMT task itself

* Corresponding author at: School of Psychology, Georgia Institute of Technology, Atlanta, GA, USA.
E-mail address: Sammy.Weiss-Cowie@mrc-cbu.cam.ac.uk (S. Weiss-Cowie).

https://doi.org/10.1016/j.neubiorev.2023.105157
Received 25 July 2022; Received in revised form 21 March 2023; Accepted 2 April 2023
Available online 6 April 2023

provide a source of considerable heterogeneity. For instance, Liu et al. (2013) found that depressed individuals provided fewer specific and more categoric responses compared to healthy controls when they were given 60 s to generate a memory versus 30 s or 120 s. These results suggest that the way the AMT is presented can influence the responses depressed individuals provide. Excluding non-AMT studies cannot control for effect size variability caused by AMT task parameters, leaving significant task heterogeneity in the final sample. Meta-analysis is, of course, well equipped to investigate the effect of such task variations, including differences between AMT and non-AMT paradigms. It is not clear that non-AMT studies are worth excluding whereas AMT studies with different testing parameters are not.

Second, the AMT has been found to have reduced sensitivity for detecting OGM in individuals who have elevated symptoms that do not meet the threshold for clinical MDD (i.e., subthreshold depression) (Raes et al., 2007). The standard AMT includes multiple practice trials and reminders to provide specific memories, and it is likely that the low rate of categorical responses in subthreshold (compared to clinical) depression might be due to the former group's ability to use this scaffolding to keep the task instructions in mind (Debeer et al., 2009; Raes et al., 2007). Clinically depressed individuals may have a comparably difficult time taking advantage of such support structures due to impaired executive function (Dotson et al., 2020). In support of this notion, a "minimal instructions" version of the AMT that removes these aids has been shown to elicit a higher rate of categorical responses in samples with subthreshold depression than the standard AMT (Debeer et al., 2009). Alternative measures of autobiographical memory specificity, such as the Sentence Completion for Events from the Past Test (SCEPT; Raes et al., 2007) have shown similar promise for detecting OGM in subthreshold depression (Anderson et al., 2016). Thus, the present work includes non-AMT paradigms to better assess the robustness of OGM across measures of autobiographical memory performance in individuals with subthreshold depression.

Researchers have also utilised meta-analyses to test the predictions of theories behind OGM. One well established framework is the "CaR-FA-X" model (Williams et al., 2007). This theory explains OGM in terms of three interconnected processes: capture and rumination (CaR), functional avoidance (FA), and reduced executive function (X). The foundational idea is that intentional memory search tends to begin at the categorical level and simultaneously proceed towards more specific and more general elements (Conway and Pleydell-Pearce, 2000). Failure to take the search to its conclusion may be caused by any one of the CaR-FA-X factors on its own or by a combination of them. For instance, early search termination could be: a consequence of overdeveloped self-schema at the categorical level (i.e., CaR), a mechanism for avoiding highly disturbing memories (i.e., FA), or a result of not having enough executive resources to find a specific memory (i.e., X). Research assessing the predictions of the CaR-FA-X model has since been met with mixed results (Chiu et al., 2018; Sumner et al., 2011; Sumner et al., 2014).

We seek to address a critical question related to the CaR-FA-X model, as well as to alternate theories of OGM. Namely, whether OGM is a temporary state-marker associated with depressed mood, or a lasting trait-marker that presents a risk factor for depression (Brittlebank et al., 1993). The CaR-FA-X model presents OGM as a product of cognitive styles that can develop absent depression, thus favoring a trait-marker account (Williams et al., 2007). One way to address the trait- versus state-marker debate is by assessing how strongly autobiographical memory specificity can predict later depression severity. The two existing meta-analyses on the topic bring support to the trait-marker theory by demonstrating OGM to be a significant predictor of symptoms at follow up (Hallford et al., 2021; Sumner et al., 2010). Another element of the trait-marker theory that *has not* been considered in previous meta-analytic work is the persistence of OGM into remission. Although there exists a sizable empirical literature suggesting that OGM remains into the euthymic phase of depression (Mackinger et al., 2000; Spinhoven et al., 2006; Young et al., 2014), some studies have failed to detect an association between remitted depression and OGM (Wessel et al., 2001). It is possible that this inconsistency is a consequence of differing AMT parameters or subject variables (e.g., age, clinical diagnosis, comorbid diagnoses) between studies. Considering remitted (versus current) depressed status as a moderator variable across a large number of studies would provide a way to cut through individual differences among research projects. The present work will be the first meta-analysis to consider the issue of remission, with the inclusion of non-AMT paradigms boosting our ability to detect remission effects, should they exist.

Finally, valence has also been a topic of interest in past work. The original AMT contained 5 positively valenced and 5 negatively valenced cue words (Williams and Broadbent, 1986). Subsequent studies have typically maintained this binary distinction, or alternatively added a third category of neutral words. Despite the longstanding use of valenced cue words, conclusions regarding valence effects on OGM remain elusive. There are some indications that OGM occurs to the same extent regardless of cue or memory valence (Griffith et al., 2009; Wessel et al., 2001). Other studies have identified depressed samples as having more specific memory for negative cues over positive cues (Lemogne et al., 2006; Nandrino et al., 2002) and the reverse trend of more specific memory for positive cues over negative ones (Mansell and Lam, 2004). Similar discrepancies regarding valence have come up in meta-analyses and review articles as well (Van Vreeswijk and De Wilde, 2004; Williams et al., 2007).

One possible explanation for the mixed results is related to the method used to consider valence as a moderator. That is, previous meta-analyses have employed one of two methods when dealing with valence: combining positive and negative effect sizes into a single measure and omitting valence analyses (e.g., Sumner et al., 2010, Hallford et al., 2021), or conducting entirely separate analyses by valence (e.g., Van Vreeswijk and De wilde, 2004). The former option controls for issues that come with including multiple effect sizes from a single study, such as similar standard errors and inflating the relative importance of studies with a greater number of effect sizes. However, this comes at the cost of blurring across potentially meaningful distinctions between valence categories. By contrast, splitting positive and negative effect sizes creates concerns about dependent effect sizes if researchers do not control for the relatedness of multiple effect sizes from the same study. Using a three-level meta-analysis—which nests effect sizes within studies and thus allows for the inclusion of multiple effect sizes from each study—avoids the pitfalls of both these approaches. Three-level meta-analyses not only account for the individual sampling variance and between-study variance considered in traditional meta-analyses, but also the variance between effect sizes within a study (Cheung, 2014). Because many autobiographical memory studies report several scores that vary by valence or test (e.g., AMT versus Autobiographical Memory Interview; (Kopelman et al., 1989)), a method allowing for the safe inclusion of more than one effect size per study is necessary to capture these effects. The current work will be the first to apply a three-level meta-analysis to the study of OGM in depression.

1.1. Present study

The current meta-analysis is a comprehensive synthesis of the published and unpublished literature regarding OGM in depression. As done in past work, we consider autobiographical memory specificity and categoricity in separate analyses (Hallford et al., 2021; Liu et al., 2013; Sumner et al., 2010; Van Vreeswijk and De Wilde, 2004). We will expand upon previous research in the following three ways: (a) The inclusion of studies using paradigms besides the AMT will speak to the robustness of OGM across different memory specificity measures while mitigating concerns about the AMT being insensitive to OGM in subclinical or remitted depression; (b) the addition of several participant moderator variables—most notably current depression status—will

further our understanding of how OGM patterns in different depressed populations; and (c) the use of a three-level meta-analysis will enable the simultaneous consideration of OGM severity as a whole and separated by valence without raising concerns about dependent effect sizes (Cheung, 2014).

2. Methods

We conducted a meta-analysis following the guidelines set forth in the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA; Moher et al., 2009). All the relevant data and analysis files can be found on the project's page on the Open Science Framework (https://osf.io/56bsj/?view_only=12d37beae3a444149a0c65bddb0c9043).

2.1. Study selection

The first author and a research assistant independently conducted a thorough database search using PubMed, PsychInfo, and Proquest Dissertations and Theses. Proquest Dissertations and Theses was utilised to find unpublished manuscripts, which were evaluated using the same criteria as published works. The search results ranged from February 2021 back to 1980, the year "major depressive disorder" first appeared in the *Diagnostic and Statistical Manual of Mental Disorders,* third edition (DSM-III; American Psychiatric Association, 1980). Filters were applied to limit the studies to human populations with subjects aged 18 and above. The following search terms were used to sift through articles based on the titles and abstracts, where the asterisk allows for different endings for a given word: ('depr*' OR 'MDD' OR 'dyshtymi*' OR 'dysphor*') AND ('autobiographical').

Articles were assessed for eligibility by two raters based on titles, then abstracts, then full texts. The inclusion criteria required: (a) the presence of a sample with either diagnosed MDD (current or remitted) or subthreshold depression based on symptomatology levels, (b) the presence of a healthy control group, and (c) the inclusion of a measure of autobiographical memory specificity or categoricity—such as the AMT or the Autobiographical Memory Interview—with sufficient information to extract effect sizes (i.e., means and standard deviations, *F* statistics, t statistics, or *p*-values). Both published articles and unpublished manuscripts were considered. To avoid confounding effects from comorbid conditions that affect memory, we excluded studies where either sample exhibited dementia, Alzheimer's disease, mild cognitive impairment, brain injury, Parkinson's disease, history of stroke, HIV, cancer, diabetes, multiple sclerosis, attention deficit hyperactivity disorder (ADHD), schizophrenia, alcohol or substance abuse, personality disorders, PTSD, bipolar disorder, psychotic symptoms of depression, or postpartum depression. Due to the high rates of comorbidity of depression and anxiety disorders, comorbid anxiety was not considered exclusionary. Studies were excluded if any of the participants overlapped with individuals in other studies, with precedence given to studies with larger sample sizes and then to those published more recently. Data was included from a range of community, university, outpatient, and inpatient settings. If a study focused on electroconvulsive therapy (ECT) or other treatments, only pretreatment data was included. Where the two raters disagreed over study inclusion, an agreement was reached after discussing the particulars of the article in question.

The initial search returned 938 non-duplicate records. After removing studies based on their titles and abstracts, the remaining 294 records were screened based on full texts. When full texts were unavailable or more information was required, the original authors were contacted and 26 out of 47 such requests were met with positive responses. Ultimately, 67 records met all criteria and were included in the specificity and/or categoricity analyses. Fig. 1 depicts the overall search process and reasons for article exclusion.

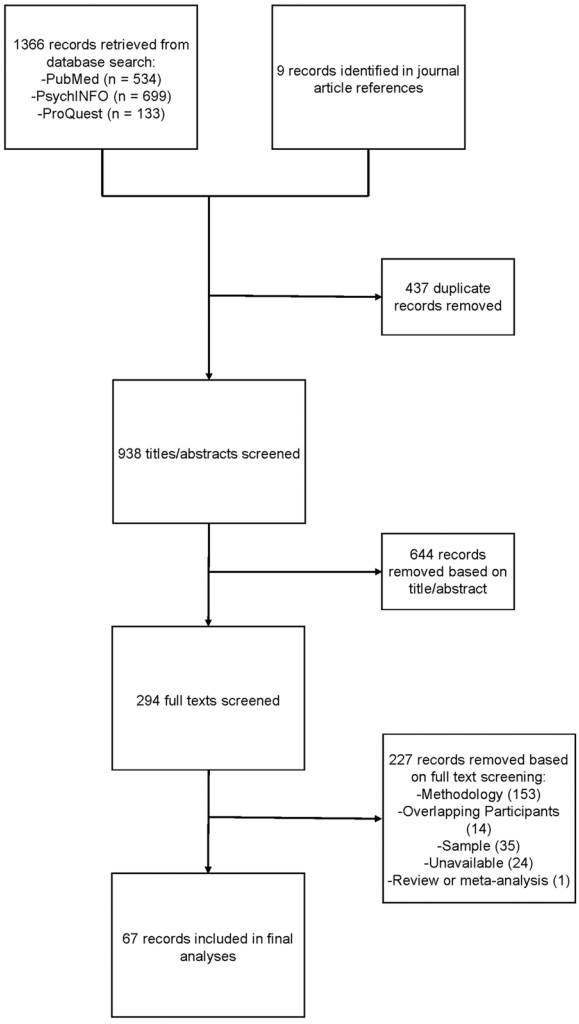

Fig. 1. Search Process.

2.2. Data extraction and variable coding

The following variables were recorded where possible: publication status, diagnostic status of the depressed sample (i.e., proportion with a clinical MDD diagnosis), current depression status of the depressed sample (i.e., proportion in a current MDD episode versus proportion remitted), sample size of the depressed and healthy control groups, average age in the depressed and healthy control groups, sex composition of each group (i.e., proportion females in each group), average years of education in the depressed and healthy control groups, whether the depressed and healthy control groups were matched for age and/or years of education, average IQ in the depressed and healthy control groups, first-episode versus recurrent depression, onset age of depression, time since initial depression diagnosis, proportion of the depressed sample on antidepressant medication, recruitment site (i.e., a clinical setting such as an inpatient ward or outpatient clinic versus a community or university setting), depression severity (as measured by a self-report form), whether the test used in the study was the AMT, the valence of the memory cues (i.e., positive, negative, or neutral), and whether an effect size corresponded to memory specificity or categoricity. In cases where only age of depression onset was reported, time since diagnosis was calculated based on the mean age of the depressed sample. When only time since diagnosis was reported, age of depression onset was estimated using the mean age of the depressed sample. A standardised measure of depression was calculated by converting scores to the Beck Depression Inventory (BDI) scale, which was the most widely

reported measure in our sample. Values were converted from the Hamilton Depression Rating Scale (HDRS) using the formula 'BDI = (HDRS – 0.65)/0.67′ (Vittengl et al., 2005) and from the Montgomery-Åsberg Depression Rating Scale (MADRS) using the formula 'BDI = 1.28 *MADRS – 3.33′ (Heo et al., 2007).

2.3. Meta-analytic approach

Effect sizes were initially converted to Cohen's *d* based on the formula $(M_1 - M_2)/SD_{pooled}$, with M_1 referring to the performance of the depressed group, M_2 to the performance of the healthy control group, and SD_{pooled} to the pooled standard deviation of the two groups. As a result, positive effect sizes indicate a higher level of specificity or categoricity in the depressed group for measures of specific memories and categoric memories, respectively. In cases where means and standard deviations were not available, *F* statistics, *t* statistics, or *p*-values were used to estimate effect sizes based on the equations used by (Coles et al., 2019). Hedges' *g* correction for small sample bias (Hedges and Olkin, 1985) was used to convert effect sizes before conducting analyses. The following equation was used: $g = d[1 - (3/4\,df - 1)]$, with *df* representing the combined degrees of freedom for the depressed and healthy control groups. In cases where multiple relevant measures were reported in a single study (e.g., number of specific and categoric memories separated by valence), a separate effect size was calculated for each.

These effect sizes were then analyzed via three-level meta-analysis in R, using the *metafor* package to implement a three-level random effects model (Viechtbauer, 2010), the *clubSandwich* package for cluster-robust variance estimation (Pustejovsky and Tipton, 2021), and *dplyr* for data manipulation (Wickham et al., 2021). Separate analyses were conducted for measures of specificity and categoricity. The mean effect sizes were calculated, in addition to 95% confidence intervals. Effect size heterogeneity was assessed using both the *Q* statistic and τ^2. In a three-level meta-analysis, τ^2 refers to two distinct measures: $\tau^2_{(2)}$ and $\tau^2_{(3)}$. While the former deals with heterogeneity between effect sizes from the same study, the latter corresponds to effect size heterogeneity after controlling for differences between measures at the study level.

2.4. Moderator analyses

Moderator analyses were conducted using the maximum likelihood estimate method. The following moderator variables were considered: age, sex, years of education, premorbid IQ, diagnostic status (i.e., proportion with a clinical MDD diagnosis), depression status (i.e., proportion in a current MDD episode versus proportion remitted), matched age, matched education, antidepressant use, use of the AMT, symptom severity (using scores on the Beck Depression Interview (BDI) and others converted to the BDI scale), first-episode or recurrent depression, time since diagnosis, age of depression onset, publication status, cue/memory valence, and recruitment site.

3. Results

3.1. Outliers and publication bias

Two effect sizes were identified as being over 3 SD away from the average effect size in the specificity analyses (Gupta and Kar, 2012; memory specificity for positive cues in the remitted and currently depressed groups). These effect sizes were excluded from primary and moderator analyses. Publication bias was assessed via regression analysis on the funnel plots, once using all effect sizes (which preserves all data but violates the assumption of independence inherent in these analyses), and once using effect sizes averaged within studies (to meet the independence assumption; Viechtbauer, 2010). Analysis using all effect sizes identified significant evidence of publication bias in both the specificity ($z = -8.03$, $p < 0.0001$) and categoricity ($z = 2.60$, $p = 0.009$) analyses. The same was true for analyses using averaged effect sizes ($z = -6.29$, $p < 0.001$; $z = 2.07$, $p = 0.04$ for the specificity and categoricity analyses, respectively). The trim-and-fill method (Duval and Tweedie, 2000), however, did not identify any missing studies for categoricity analyses. Trim-and-fill analyses identified 5 missing studies in the specificity analyses using the R_0-estimator method on all effect sizes, but not for the effect sizes averaged within study. Funnel plots corresponding to the memory specificity and categoricity results using all effect sizes are presented in Fig. 2 and Fig. 3, respectively.

3.2. Overall effect sizes

3.2.1. Autobiographical memory specificity

A three-level model considering autobiographical memory specificity across 140 effect sizes, including 83 unique comparisons from 67 records, resulted in an overall effect size of $g = -0.73$ (95% CI [−0.90, −0.57]; $p < 0.001$). This represents a medium-to-large effect corresponding to lower memory specificity in the depressed groups. The effect size remained of similar magnitude following trim-and-fill analyses regarding publication bias ($g = -0.70$, 95% CI [−0.83, −0.56]; $g = -0.75$, 95% CI [−0.91, −0.58] for analyses with all effect sizes and with averaged effect sizes, respectively). The distribution of effect sizes across studies can be seen via forest plot in Fig. 4. Moderator analyses were warranted, given the presence of significant heterogeneity ($Q(139) = 912.39$, $p = 0.00$; $\tau^2_{(2)} = 0.23$, $p = .0004$; $\tau^2_{(3)} = 0.36$, $p = 0.0006$; $I^2_{(2)} = 0.35$ and $I^2_{(3)} = 0.53$). See Table 1 for descriptive statistics of the studies included in the specificity analyses.

3.2.2. Autobiographical memory categoricity

A three-level model concerning effect sizes related to categorical memories across 41 effect sizes, including 28 unique comparisons from 22 records resulted in an overall effect size of $g = 0.77$ (95% CI [0.49, 1.05]; $p < 0.001$). This is a medium-to-large effect size pointing to a higher degree of overgeneral memory in the depressed samples. The effect size was of similar magnitude even after conducting trim-and-fill analyses regarding publication bias ($g = 0.77$, 95% CI [0.47, 1.07]; $g = 0.80$, 95% CI [0.47, 1.14] for analyses with all effect sizes and with averaged effect sizes, respectively). A forest plot with all the relevant effect sizes is shown in Fig. 5. Significant heterogeneity was found ($Q(40) = 362.1$, $p < 0.0001$; $\tau^2_{(2)} = 0.84$, $p = .012$; $\tau^2_{(3)} = 0.00$, $p = 1.00$; $I^2_{(2)} = 0.91$ and $I^2_{(3)} = 0.00$), indicating the need for moderator analyses. See Table 2 for descriptive statistics of the studies included in the specificity analyses.

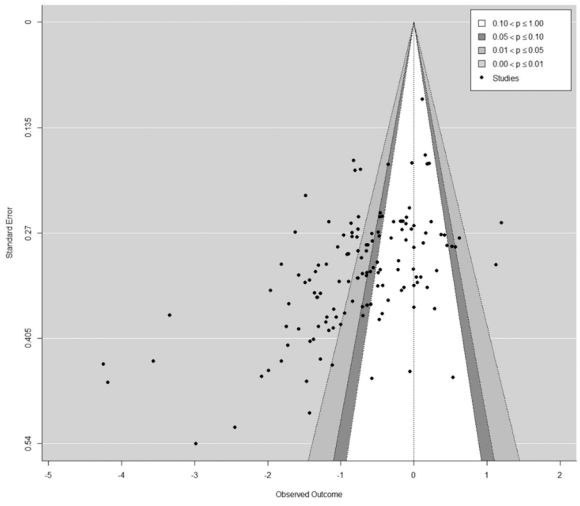

Fig. 2. : Funnel Plot for Memory Specificity Results. *Each dot represents a single included effect size placed according to its value and corresponding standard error.*

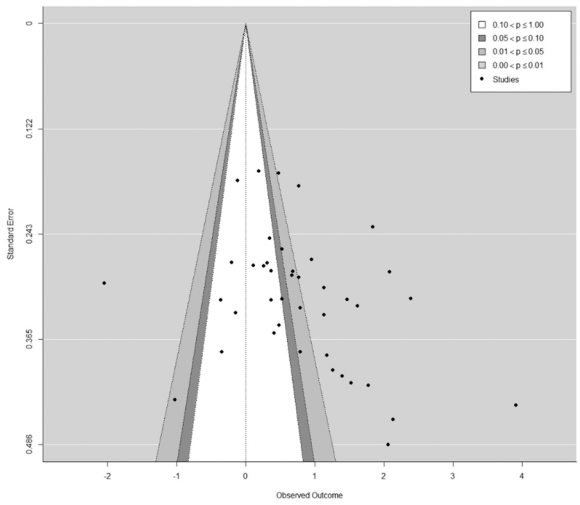

Fig. 3. : Funnel Plot for Memory Categoricity Results. *Each dot represents a single included effect size placed according to its value and corresponding standard error.*

3.3. Moderator analyses

Moderator analyses were conducted for multiple participant and study variables, the results of which are reported in Table 3 and Table 4.

As can be seen in Tables 3 and 4, the only significant moderators for memory specificity were current depressed status (with current depression being associated with lower memory specificity than remitted depression), recruitment site (with worse performance amongst depressed samples recruited from clinical settings compared to those recruited from university or community environments), and years of education (increasing education level was associated with higher levels of specificity). Pairwise comparisons regarding valence did not reveal and significant effects.

As shown in Tables 3 and 4, the only significant moderators for memory categoricity were depression diagnosis status (a diagnosis of MDD was associated with higher levels of memory categoricity), use of the AMT (with higher levels of categoricity in studies using the AMT versus studies that did not use the AMT), and age-matching (with age-matched samples associated with a higher degree of categoricity versus samples not matched for age). However, there was only one sample not matched for age, so that result should be interpreted with extreme caution. Pairwise comparisons for valence yielded no significant effects.

4. Discussion

The present work is the largest meta-analysis to date on OGM in depression and includes results from both the published and unpublished literature. Our results are in line with past reviews indicating reduced autobiographical memory specificity and heightened autobiographical memory categoricity in depression (Liu et al., 2013; Van Vreeswijk and De Wilde, 2004; Williams et al., 2007). Analysis of 140 effect sizes from 67 records revealed a medium-to-large effect size of − 0.73, indicating reduced memory specificity in depressed samples. In addition, the analysis of 41 effect sizes from 22 records resulted in a medium-to-large effect size of 0.77, suggesting heightened categoricity in the depressed samples versus healthy controls. These results echo past findings regarding OGM in MDD while also providing novel evidence that the OGM phenomenon is evident even when paradigms besides the AMT are used and samples with subthreshold or remitted depression are included. The implications of our moderator analyses are discussed below.

4.1. Does OGM severity differ in clinical versus subthreshold or remitted depression?

Results regarding diagnostic status and OGM were mixed. Diagnosed depression was associated with higher levels of memory categoricity than was subthreshold depression, although categoricity was significantly elevated in both groups compared to controls. At the same time, we failed to uncover an effect of diagnostic status on memory specificity. Relatedly, moderator analyses revealed decreased specificity for participants in a current depressive episode compared to those remitted from depression. Self-reported depression severity did not moderate specificity or categoricity, which fits with empirical work arguing against a direct association between symptom severity and OGM (Young et al., 2014).

Taken as a whole, the moderator analyses imply that OGM is a phenomenon more severe for people in a current episode of clinical MDD than with remitted or subthreshold depression. This idea fits with the finding that depressed samples recruited in inpatient or outpatient settings—where individuals are more likely to have clinically diagnosed, current depression-had significantly lower specificity than depressed samples drawn from community or university samples. Nonetheless, individuals experiencing subthreshold or remitted depression still exhibit OGM compared to the general population.

These results can be interpreted in several ways relating to the trait-marker theory of OGM, which argues that OGM is not a direct consequence of being in a depressive episode but is instead a consistent cognitive style that makes individuals more vulnerable to depression (Brittlebank et al., 1993; Hallford et al., 2021; Liu et al., 2013; Sumner et al., 2010; Young et al., 2016). On the one hand, it appears that OGM is a phenomenon that occurs in people who have not yet reached a severe level of depression or are no longer in a depressive episode—albeit at a lower severity than in clinical MDD. Thus it could be the case that certain individuals may have an underlying tendency for retrieving less specific memories regardless of whether they are currently depressed. The lack of an effect of depressive symptom severity on memory specificity or categoricity further supports this notion. On the other hand, we did find that current depression was associated lower specificity than remitted depression and that clinical depression was associated with higher categoricity than subthreshold depression. Although OGM may be present before depression onset and after remission, it seems to be exacerbated once depression begins.

The present findings suggest that OGM becomes more pronounced after individuals meet the criteria for clinical depression and that it largely abates after other depression symptoms have decreased. This raises a critical question of causality: Does increasing autobiographical memory specificity contribute to remission from depression, is OGM ameliorated only once remission has occurred, or do improvements in memory specificity coincide with depression remission? Some evidence suggests that interventions targeted at improving autobiographical memory specificity can also reduce depression symptoms (Barry et al., 2019). In this sort of treatment, patients undergo several weeks of training to help them focus on the more specific parts of their memories (Raes et al., 2009). Barry et al. (2019) conducted a meta-analysis looking at the impact of Memory Specificity Training (MeST) on memory specificity, depression symptoms, and several measures of cognitive function. They found that individuals who participated in MeST had greater memory specificity and lower depression severity post-intervention compared to individuals who did not take part in MeST. This provides evidence that OGM may play an active role in depression remission and maintenance. Consistent with this notion, other work has shown that OGM is predictive of later depression symptoms (Hallford et al., 2021; Sumner et al., 2010).

Another potential explanation for the relationship between OGM and depression remission is that general depression symptoms abate before patients recoup their autobiographical memory specificity. Existing work regarding this issue has not found an association between

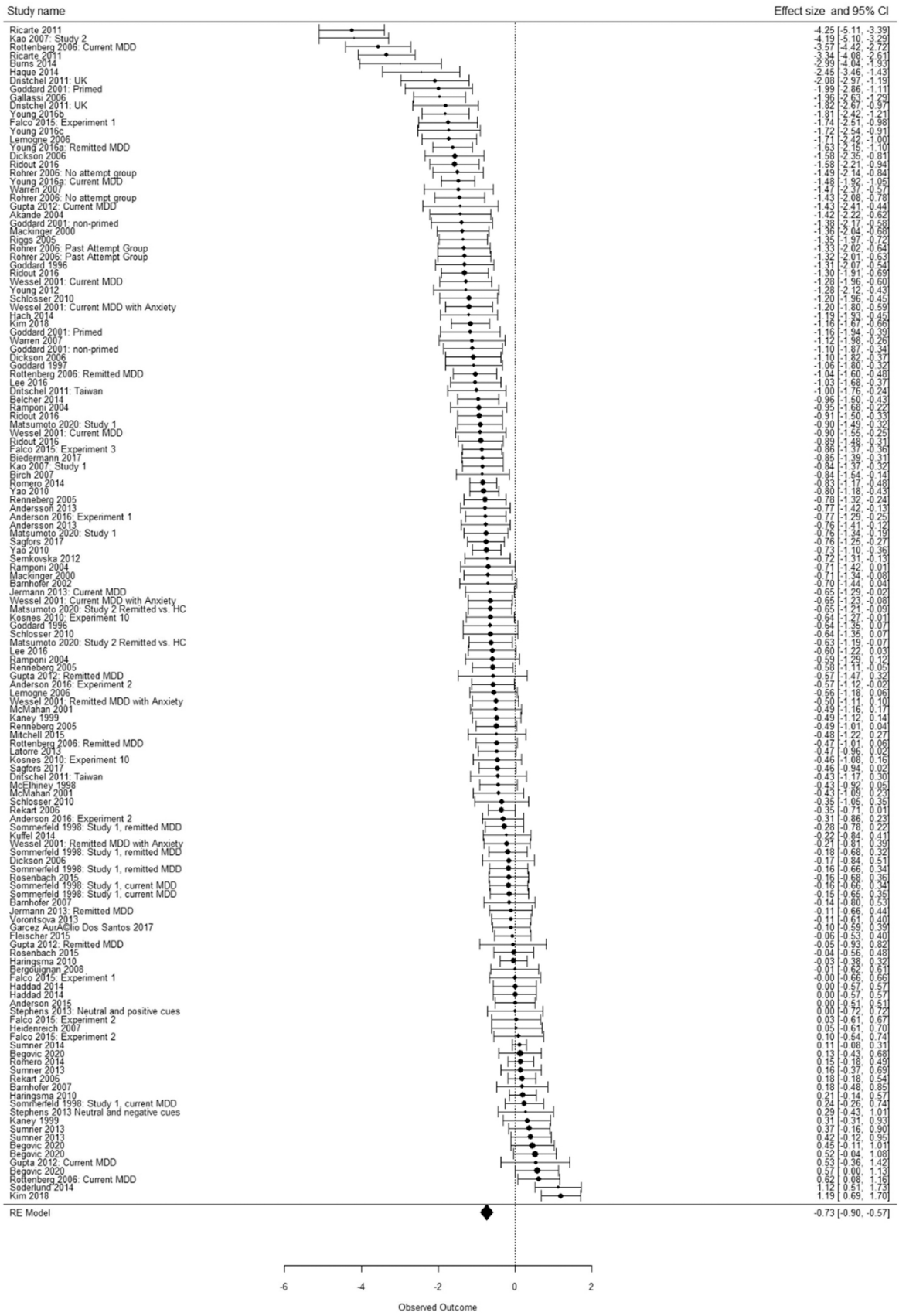

Fig. 4. : Forest Plot of Effect Sizes for Memory Specificity. *Effect sizes are individually represented as dots, with error bars indicating 95% confidence intervals. The diamond at the bottom represents the overall effect size. The dotted vertical line indicates an effect size of zero.*

Table 1
Sample and Moderator Characteristics for Studies Included in the Specificity Meta-Analysis.

Number of effect sizes (*k*)	140		
Number of samples (*j*)	83		
Mean publication year (SD)	2010.03 (6.05)		
Number of depressed individuals	2175		
Number of control individuals	2499		
Continuous Moderators	Mean	Range	*k* (%)
Age	38.30	20.1–74.59	113 (80.7)
Sex (female proportion)	0.685	0.25–1	114 (81.4)
Education (years)	12.65	7.15–16.5	47 (33.6)
Premorbid IQ	104.70	91.1–120	26 (18.6)
Diagnostic Status	0.867	0–1	114 (81.4)
Depression Status	0.819	0–1	140 (100)
Medication Status	0.502	0–1	47 (33.6)
Symptom Severity (BDI scale)	22.4	2.5–44.6	99 (70.7)
Episode (recurrent proportion)	0.638	0.5–0.86	7 (5.0)
Time Since Diagnosis (months)	125.5	30–204	13 (9.3)
Age of Depression Onset	30.8	17.29–54.79	13 (9.3)
Categorical Moderators	*k* (%)		
Test Type			
AMT	117 (83.6)		
Non-AMT	23 (16.4)		
Valence			
Neutral	9 (6.4)		
Positive	41 (29.3)		
Negative	44 (31.4)		
NA	46 (32.9)		
Publication Status			
Published	116 (82.9)		
Unpublished	24 (17.1)		
Age Matching			
Age Matched	104 (74.3)		
Not Age Matched	9 (6.4)		
NA	27 (19.3)		
Education Matching			
Education Matched	65 (46.4)		
Not Education Matched	10 (7.1)		
NA	65 (46.4)		
Recruitment Site			
Clinical	72 (51.4)		
Community/University	60 (42.9)		
NA	8 (5.7)		

Note. j = number of unique depression-control group comparisons. k = number of effect sizes. All continuous moderators are based on information from depression groups. Diagnostic Status refers to the proportion of a sample with a clinical MDD diagnosis. Depression Status refers to the proportion in a current (rather than remitted) depressive episode. Episode represents the proportion with recurrent episodes of depression (as opposed to first-episode depression). Medication represents the proportion taking medication for depression. If a study only reported Age of Depression Onset, Time Since Diagnosis was calculated using the mean age (and vice versa).

remission duration and memory specificity (Spinhoven et al., 2006). Thus, it seems unlikely that depression remission precedes autobiographical memory improvements. Due to the cross-sectional nature of our meta-analysis, the present work cannot necessarily arbitrate between different theories about OGM and depression remission. Longitudinal work looking at MeST as a preventative measure in subthreshold and remitted depression would be worthwhile.

4.2. Are there factors that can mitigate OGM severity?

The 'X' element in the 'CaR-FA-X' model argues that executive dysfunction is a key factor in causing and maintaining OGM in psychopathology (Williams et al., 2007). Although there were not enough explicit executive function measures reported in the extracted studies to enable a direct analysis, we did consider two measures commonly associated with executive function: education level (Opdebeeck et al., 2016) and IQ (Arffa, 2007; Leeson et al., 2010). Higher educational attainment was associated with increased memory specificity, with each additional year of education in the depressed group corresponding to a 0.15 SD smaller gap between the depressed and control sample. It should be noted that the intercept for this analysis was −2.70, meaning that roughly 18 years of education would be needed to completely eliminate the memory specificity effect.

A similar positive association between education level and autobiographical memory specificity has been reported in past work on MDD (Wessel et al., 2001). These findings are fitting with the idea of education imparting increased cognitive reserve (Le Carret et al., 2003). The idea of cognitive reserve is most commonly used to explain a resistance to cognitive difficulties brought on by neurodegenerative disorders (Barulli and Stern, 2013). However, it can also be applied to OGM when framed in the context of broader cognitive deficits like impaired executive function and source memory (Dalgleish et al., 2007; Raes et al., 2006).

One possibility that requires future study is that depressed individuals with greater executive functioning (i.e., higher education/IQ) may be better able to keep the task instructions in mind when searching for a specific memory and thus successfully retrieve specific memories at a greater rate (Debeer et al., 2009; Raes et al., 2007). Improving participants' ability to maintain the goal of retrieving a specific memory is a hallmark of the MeST paradigms described above. Another goal of such interventions is to improve individuals' metacognitive awareness of when they are using a categoric retrieval style, especially in stressful situations that sap already limited cognitive control resources (Barry et al., 2019). Rumination is a common phenomenon in depression (Chiu et al., 2018) that also has the potential to leech executive resources away from memory retrieval. Individuals with higher baseline executive functioning may be quicker to recognise that their attention has been captured by ruminative thoughts and tune out that internal dialogue (Williams et al., 2007).

A somewhat surprising finding was the lack of a moderating effect of age on either memory specificity or categoricity. Note, however, that the effect of age on specificity was in the expected direction and would have been significant if a one-tailed test was conducted. Recent meta-analytic work from our lab has suggested that episodic memory deficits seen in depression are exacerbated as people age (James et al., In press). Thus, we predicted higher age would be associated with decreased autobiographical memory specificity and/or increased categoricity. Past meta-analyses have shown that there is a stronger negative correlation between memory specificity and later depression symptoms with increasing age (Hallford et al., 2021; Sumner et al., 2010), suggesting age may have a moderating influence between OGM and depression. The lack of a moderating effect of age on autobiographical memory in our work may be explained by the uniquely personal nature of autobiographical memories. Although older adults with depression may have more difficulty retrieving episodic memories than younger adults with depression, the time provided by age to gain new memories and rehearse old ones may enable equivalent performance on autobiographical memory measures (Luchetti and Sutin, 2018). Broader episodic memory measures frequently utilise stimuli that are not as self-relevant, such as word lists (e.g., California Verbal Learning Test) or pre-written narratives (e.g., Wechsler Memory Scale Logical Memory subtest). The differences between the constructs of episodic memory and autobiographical may illuminate why age did not moderate the relationship between OGM and depression in our analyses.

4.3. Does OGM severity differ across autobiographical memory measures?

The present work includes effect sizes from non-AMT measures, a first among meta-analyses regarding OGM in depression. Because effect sizes have not previously been combined across autobiographical memory instruments, we conducted moderator analyses to see if there were significant differences between AMT and non-AMT effect sizes. While AMT use did not moderate autobiographical memory specificity, it did significantly moderate categoricity such that AMT effect sizes were

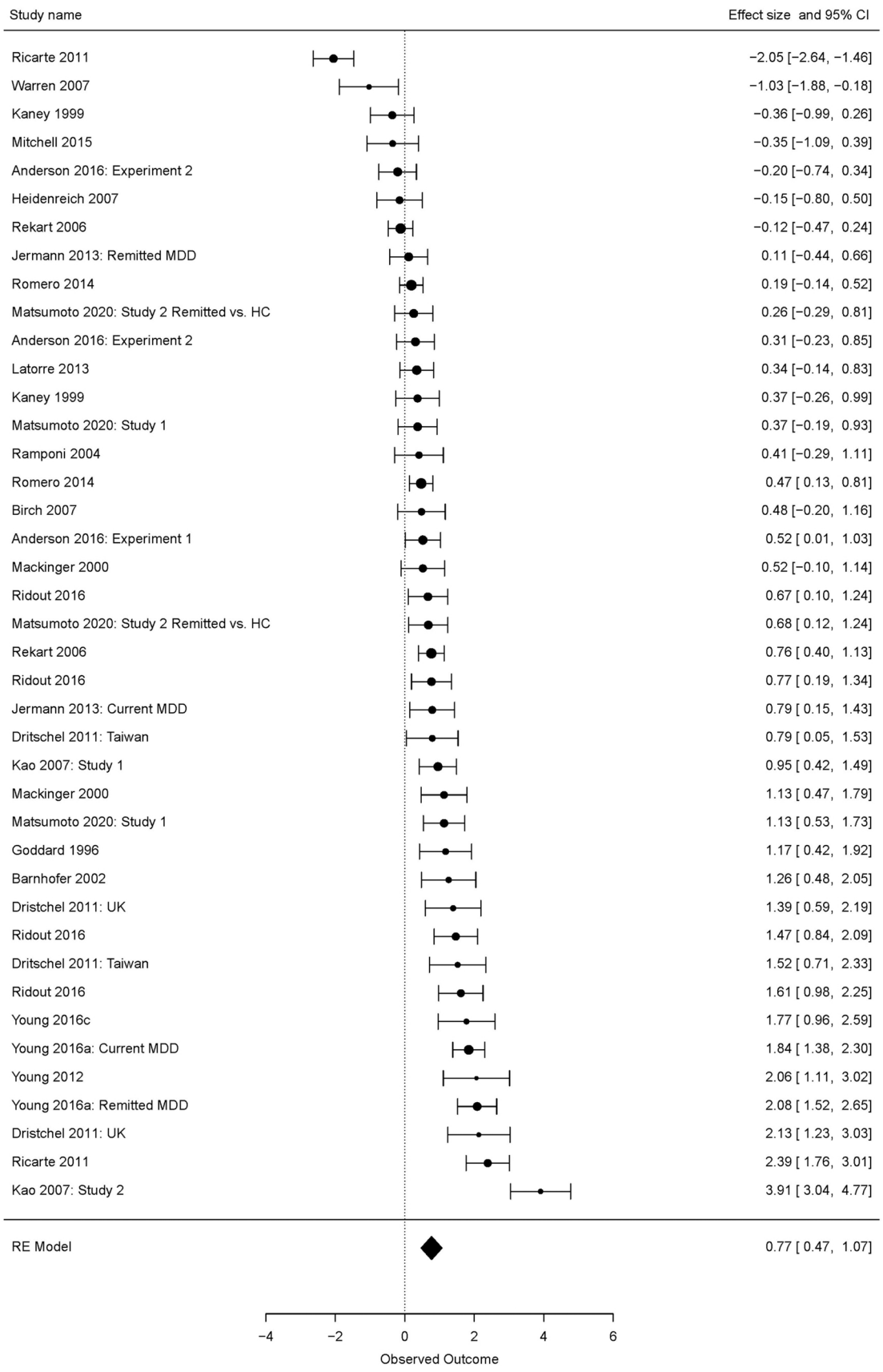

Fig. 5. : Forest Plot of Effect Sizes for Memory Categoricity. *Effect sizes are individually represented as dots, with error bars indicating 95% confidence intervals. The diamond at the bottom represents the overall effect size. The dotted vertical line indicates an effect size of zero.*

Table 2
Sample and Moderator Characteristics for Studies Included in the Categoricity Meta-Analysis.

Number of effect sizes (*k*)	41		
Number of samples (*j*)	28		
Mean publication year (SD)	2009.46 (6.44)		
Number of depressed individuals	715		
Number of control individuals	778		
Continuous Moderators	Mean	Range	*k* (%)
Age	39.03	20.1–74.59	37 (90.2)
Sex (female proportion)	0.662	0.25–1	39 (95.1)
Education (years)	10.97	7.8–15.4	12 (29.3)
Premorbid IQ	106.68	96.3–120	14 (34.1)
Diagnostic Status	0.824	0–1	34 (82.9)
Depression Status	0.829	0–1	41 (100)
Medication Status	0.539	0–1	18 (43.9)
Symptom Severity (BDI scale)	22.84	6.3–34.7	29 (70.7)
Episode (recurrent proportion)	0.500	0.5–0.5	2 (4.9)
Categorical Moderators	*k* (%)		
Test Type			
AMT	35 (85.4)		
Non-AMT	6 (14.6)		
Valence			
Neutral	1 (2.4)		
Positive	12 (29.3)		
Negative	12 (29.3)		
NA	16 (39.0)		
Publication Status			
Published	37 (90.2)		
Unpublished	4 (9.8)		
Age Matching			
Age Matched	34 (82.9)		
Not Age Matched	1 (2.4)		
NA	6 (14.6)		
Education Matching			
Education Matched	17 (41.5)		
Not Education Matched	3 (7.3)		
NA	21 (51.2)		
Recruitment Site			
Clinical	22 (53.7)		
Community/University	18 (43.9)		
NA	1 (2.4)		

Note. j = number of unique depression-control group comparisons. k = number of effect sizes. All continuous moderators are based on information from depression groups. Diagnostic Status refers to the proportion of a sample with a clinical MDD diagnosis. Depression Status refers to the proportion in a current (rather than remitted) depressive episode. Episode represents the proportion with recurrent episodes of depression (as opposed to first-episode depression). Medication represents the proportion taking medication for depression.

associated with higher categoricity than non-AMT tasks. However, heightened categoricity was still evident when non-AMT tasks were used, and non-AMT effect sizes came from relatively fewer samples in the categoricity analysis than in the specificity analysis (4 versus 18). The lack of an effect in the specificity analysis is encouraging, as it suggests consistent results across different measures of autobiographical memory specificity. Limiting meta-analyses to AMT studies has usually been done under the assumption that introducing new measures would add excessive heterogeneity to the data. This choice has necessarily excluded a significant part of the literature (23 non-AMT effect sizes were included in the present specificity analyses).

4.4. Does OGM severity vary by valence?

The finding of a null effect for valence was not entirely expected, given past meta-analytic work suggesting more severe OGM for positive memories than negative ones (Van Vreeswijk and De Wilde, 2004) and empirical research regarding mood-congruent memory in depression (Marchetti et al., 2018). There are multiple potential explanations for the lack of a valence effect. One possibility is that most studies measure valence based on cue valence rather than memory valence. That is, even though the cues suggest a particular valence, it does not necessarily mean that a participant will generate a memory of the matching valence. For instance, Young et al. (2012) looked at the valence of autobiographical memories provided by depressed and control individuals in response to positive, negative, and neutral cues. After combining across memory categories (e.g., specific, categoric, etc.), they found that depressed individuals provided fewer total positive memories in response to positive and neutral cues than controls (Young et al., 2012). If depressed individuals have a bias towards recalling more negative memories and/or fewer positive ones in response to all cues, it would not be surprising to see minimal effect of cue valence on memory specificity or categoricity. It may be beneficial for future work to focus on memory valence rather than cue valence during data collection and analysis.

Alternatively, there may have been no moderating effect of valence because depressed adults have equivalently reduced memory specificity regardless of valence. Williams et al. (2007) argue that OGM may begin as a means of avoiding the most harrowing details of traumatic memories but grow into a more general style of memory retrieval over time. This idea is consistent with longitudinal work in depressed adolescents showing that OGM for negative cues can predict later depression onset and symptom severity better than memory specificity for non-negative cues (Rawal and Rice, 2012; Warne et al., 2020). Whether these results can be replicated in adult samples—like those included in the present meta-analysis—remains to be seen. Longitudinal work is needed to determine whether depressed individuals' relative specificity for positive and negative memories changes over time. Analyses regarding memory valence and time since diagnosis could provide clues regarding this question, but unfortunately depression duration was not reported frequently enough in the present set of studies to allow such an analysis.

5. Limitations and future directions

The present work necessarily represents a snapshot of the samples from the included studies. To get a clearer picture of how variables such as depression diagnosis, symptom severity, and remission from depression impact autobiographical memory specificity, it is ideal to conduct longitudinal work tracking changes in the same participants over time. Although our results suggest that OGM severity may decrease after a depressive episode has ended, longitudinal work is needed to confirm this finding.

Another limitation of the present work is that it cannot draw a direct causal link between the variables considered herein. For instance, although we found that remitted depression was associated with lower OGM severity than current depression, it was not possible for us to assess whether the remission was driven by improved autobiographical memory or vice versa. Work testing the efficacy of depression treatments targeting autobiographical memory specificity will be useful for teasing apart the relationship between OGM and remitted, subthreshold, and current clinical depression.

The current meta-analysis focuses on the relationship between OGM and adult depression. Studies assessing OGM in non-adult populations were excluded due to evidence of differences in the aetiology and presentation of depression in children, adolescents, and adults—separate from the questions under study here (Karlsson et al., 2007; Thompson, 2012). Investigating the nature of the OGM phenomenon in childhood and adolescent depression would itself be a valuable avenue of enquiry, especially given the increasing prevalence of depression in these populations (Shorey et al., 2022).

Finally, our results may have been impacted by publication bias. Although we made a special effort to include unpublished manuscripts in our work, the publication biases we conducted did suggest the presence of publication bias.

6. Conclusion

As the largest meta-analysis to date on autobiographical memory in depression, our work provides robust evidence of overgeneral

Table 3
Moderator Analyses for Categorical Variables.

Moderator (and *sublevels*)	*j*	*k*	*g*	β_1	95% CI	*p*	$R^2_{(2)}$	$R^2_{(3)}$	*Q*	$\tau^2_{(2)}$	$\tau^2_{(3)}$
Memory Specificity											
Diagnostic Status	66	114	—	-0.40	[−0.83, 0.02]	.10	.00	.04	816.5***	.25***	.40**
Diagnosed	58	98	-0.78	—	[−1.20, −0.36]	—	—	—	751.1***	—	—
Subthreshold	7	15	-0.38	—	[−0.74, −0.02]	—	—	—	48.1***	—	—
Depression Status	83	140	—	-0.43**	[−0.75, −0.12]	.01	.00	.08	912.4***	.24***	.33***
Current	67	113	-0.81	—	[−1.12, −0.50]	—	—	—	783.7	—	—
Remitted	13	23	-0.38	—	[−0.622, −0.14]	—	—	—	62.8*	—	—
Medication Status	28	47	—	-0.32	[−1.42, 0.78]	.57	.00	.03	356.5***	.02	.57**
Medicated	2	5	-1.11	—	[−2.20, −0.02]	—	—	—	77.4***	—	—
Unmedicated	9	10	-0.79	—	[−1.31, −0.28]	—	—	—	71.7***	—	—
Matched Age	68	113	—	-0.09	[−0.48, 0.31]	.69	.00	.00	811.6***	.26**	.40**
Matched	62	104	-0.78	—	[−1.17, −0.38]	—	—	—	798.1***	—	—
Not Matched	6	9	-0.69	—	[−1.03, −0.35]	—	—	—	12.0	—	—
Matched Education	46	75	—	-0.06	[−0.47, 0.36]	.80	.00	.00	494.3***	.34**	.26
Matched	40	65	-0.70	—	[−0.47, 0.36]	—	—	—	476.3***	—	—
Not Matched	6	10	-0.65	—	[−0.99, −0.31]	—	—	—	16.8	—	—
AMT	83	140	—	-0.16	[−0.48, 0.16]	.34	.01	.01	912.4***	.23***	.35***
AMT	67	117	-0.76	—	[−1.08, −0.44]	—	—	—	736.7***	—	—
Non-AMT	18	23	-0.60	—	[−0.88, −0.32]	—	—	—	174.5***	—	—
Published	83	140	—	-0.07	[−0.65, 0.51]	.82	.00	.02	912.***	.24***	.35***
Published	70	116	-0.74	—	[−1.32, −0.16]	—	—	—	736.2***	—	—
Not Published	13	24	-0.68	—	[−1.23, −0.12]	—	—	—	153.0***	—	—
Valence	43	94	—	—	—	.21	—	—	3.1	—	—
Positive	39	41	-0.84	—	[−1.17, −0.51]	—	—	—	312.0***	—	—
Negative	41	44	-0.63	—	[−0.95, −0.30]	—	—	—	223.4***	—	—
Neutral	9	9	-0.74	—	[−1.03, −0.45]	—	—	—	12.6	—	—
Episode (First vs. Recurrent)	4	7	—	-0.56	[−7.77, 6.65]	.89	.00	.00	38.0***	.00	.55
Recurrent	—	—	-0.78	—	[−7.99, 6.44]	—	—	—	—	—	—
First Episode	—	—	-0.22	—	[−4.14, 3.70]	—	—	—	—	—	—
Sex	70	114	—	0.20	[−0.77, 1.17]	.69	.00	.00	730.4***	.25***	.26**
Female	7	11	-0.65	—	[−1.62, 0.33]	—	—	—	30.1***	—	—
Male	—	—	-0.84	—	[−1.46, −0.23]	—	—	—	—	—	—
Recruitment Site	78	132	—	-0.55***	[−0.86, −0.23]	.001	.00	.19	833.9***	.12**	.38***
Clinical	44	72	-0.96	—	[−1.28, −0.65]	—	—	—	409.3***	—	—
Community/University	34	60	-0.41	—	[−0.59, −0.24]	—	—	—	288.3***	—	—
Memory Categoricity											
Diagnostic Status	23	34	—	0.69*	[0.26, 1.12]	.04	.07	.00	339.3***	.91	.00
Diagnosed	19	28	0.97	—	[0.54, 1.40]	—	—	—	295.4***	—	—
Subthreshold	4	6	0.28	—	[0.07, 0.49]	—	—	—	15.6**	—	—
Depression Status	28	41	—	0.16	[−0.50, 0.82]	.66	.00	.11	362.1***	.84*	.00
Current	23	34	0.80	—	[0.14, 1.46]	—	—	—	322.2***	—	—
Remitted	5	7	0.64	—	[0.06, 1.21]	—	—	—	39.8***	—	—
Medication Status	12	18	—	-1.06	[−2.10, −0.02]	.09	.14	.00	216.9***	.97*	.00
Medicated	1	2	0.33	—	[−0.71, 1.37]	—	—	—	103.4***	—	—
Unmedicated	5	5	1.39	—	[0.78, 2.00]	—	—	—	29.7***	—	—
Matched Age	23	35	—	0.96***	[0.64, 1.28]	.000	.03	.00	325.2***	.88*	.00
Matched	22	34	0.81	—	[0.49, 1.13]	—	—	—	319.0***	—	—
Not Matched	1	1	-0.15	—	[−0.15, −0.15]	—	—	—	0.0	—	—
Matched Education	13	20		0.68	[−1.68, 2.74]	.63	.03	.05	133.9***	.08	.88*
Matched	11	17	0.81	—	[−1.41, 3.02]	—	—	—	114.7***	—	—
Not Matched	2	3	0.13	—	[−2.09, 2.34]	—	—	—	19.0***	—	—
AMT	28	41	—	0.71*	[0.30, 1.12]	.03	.08	.11	362.1***	.78*	.00
AMT	24	35	0.88	—	[0.47, 1.29]	—	—	—	331.1***	—	—
Non-AMT	4	6	0.17	—	[−0.10, 0.43]	—	—	—	8.1	—	—
Published	28	41	—	-0.56	[−2.52, 1.40]	.63	.03	.11	362.1***	.82**	.00
Published	25	37	0.71	—	[−1.25, 2.67]	—	—	—	287.7***	—	—
Not Published	3	4	1.28	—	[−0.50, 3.38]	—	—	—	74.3***	—	—
Valence	12	25	—	—	—	.56	—	—	1.2	—	—
Positive	11	12	0.47	—	[−0.31, 1.25]	—	—	—	111.1***	—	—
Negative	11	12	0.83	—	[−0.33, 1.27]	—	—	—	71.7***	—	—
Neutral	1	1	0.52	—	[0.52, 0.52]	—	—	—	0.0	—	—
Sex	26	39	—	0.0009	[−1.58, 1.58]	.99	.00	.00	300.0***	.64**	.00
Female	1	2	0.72	—	[−0.86, 2.30]	—	—	—	1.7	—	—
Male	—	—	0.72	—	[−0.45, 1.89]	—	—	—	—	—	—
Recruitment Site	27	40	—	0.21	[−0.34, 0.76]	.46	.01	.00	353.5***	.82*	.00
Clinical	14	22	0.84	—	[0.29, 1.39]	—	—	—	240.8***	—	—
Community/University	13	18	0.63	—	[0.31, 0.95]	—	—	—	109.4***	—	—

Note. *j* = number of samples. *k* = number of effect sizes. *$p < 0.05$, **$p < 0.01$, *** $p < 0.001$. β_1 coefficients are from meta-regression analyses where categorical moderators with two levels (or three in the case of valence) were dummy coded and entered in the models as predictors. 95% CI for the moderators corresponds to the β_1 coefficient. *g* refers to the predicted average effect size in each subcategory of a given moderator. Corresponding 95% CI for each effect size is shown alongside. *p* corresponds to the β_1 coefficient for moderators, except for the valence analyses, where they correspond to the results of an omnibus test with neutral effect sizes as a baseline and positive and negative effect sizes coded as dummy variables. $R^2_{(2)}$ = proportion of estimated heterogeneity explained by the predictors at level 2. $R^2_{(3)}$ = proportion of estimated heterogeneity explained by the predictors at level 3. $Q = Q$ statistic on the homogeneity of effect sizes. $\tau^2_{(2)}$ = heterogeneity of effects due to differences between measures. $\tau^2_{(3)}$ = heterogeneity of effect sizes across studies after controlling for the different types of measures at level 2.

Table 4
Moderator Analyses for Continuous Variables.

Moderator	*j*	*k*	β_1	95% CI	*p*	$R^2_{(2)}$	$R^2_{(3)}$	*Q*	$\tau^2_{(2)}$	$\tau^2_{(3)}$
Memory Specificity										
Age	69	113	-0.02	[−0.04, −0.0002]	.09	.00	.27	786.8***	.30**	.26*
Education	28	47	0.15*	[0.04, 0.26]	.03	.02	.50	338.8***	.58*	.12
Premorbid IQ	16	26	0.01	[−0.02, 0.03]	.52	.00	.14	69.2***	.07	.10
Symptom Severity – HDRS	16	22	-0.002	[−0.04, 0.03]	.93	.00	.05	75.5***	.12	.11
Symptom Severity – BDI	47	85	-0.02	[−0.04, 0.004]	.13	.02	.08	469.6***	.32***	.15
Symptom Severity (BDI scale)	58	99	-0.02	[−0.03, 0.002]	.09	.01	.10	557.2***	.33***	.14
Time Since Diagnosis	8	13	-0.001	[−0.005, 0.002]	.47	.01	.00	74.4***	.57	.00
Age of Depression Onset	8	13	-0.02	[−0.08, 0.05]	.66	.04	.00	74.4***	.55	.00
Memory Categoricity										
Age	25	37	-0.003	[−0.02, 0.01]	.76	.00	.00	334.9***	.89*	.00
Education	7	12	-0.15	[−0.34, 0.05]	.25	.00	.19	71.9***	.15	.74
Premorbid IQ	9	14	0.06	[0.02, 0.10]	.06	.00	.40	72.0***	.14	.20
Symptom Severity – HDRS	5	6	0.031	[−0.05, 0.11]	.52	.00	.36	18.4**	.04	.11
Symptom Severity – BDI	15	25	0.04	[−0.003, 0.09]	.11	.00	.29	129.3***	.12	.38
Symptom Severity (BDI scale)	19	29	0.03	[−0.02, 0.08]	.29	.00	.13	183.9***	.11	.51

Note. j = number of samples. k = number of effect sizes. *$p < 0.05$, **$p < 0.0$, 1***$p < 0.001$. β_1 coefficients are from meta-regression analyses where continuous moderators were entered in the models as predictors. 95% CI corresponds to the β_1 coefficient for moderators. p corresponds to the β_1 coefficient for moderators. $R^2_{(2)}$ = proportion of estimated heterogeneity explained by the predictors at level 2. $R^2_{(3)}$ = proportion of estimated heterogeneity explained by the predictors at level 3. Q = Q statistic on the homogeneity of effect sizes. $\tau^2_{(2)}$ = heterogeneity of effects due to differences between measures. $\tau^2_{(3)}$ = heterogeneity of effect sizes across studies after controlling for the different types of measures at level 2.

autobiographical memory in depressed individuals compared to their non-depressed counterparts. Samples of depressed individuals simultaneously exhibited less specific and more categoric memories than controls. We present the first meta-analysis on OGM in depression to consider effect sizes from non-AMT measures, validating their inclusion alongside the AMT. OGM effects were more pronounced in individuals with clinical rather than subthreshold depression, and in current over remitted depression. Moderator analyses also suggested that higher educational attainment may mitigate the severity of OGM among depressed individuals. Use of a three-level meta-analysis allowed us to include multiple effect sizes from a given study while controlling for dependencies across those effect sizes. As a result, we could conduct robust variance analyses which failed to identify any effects of valence on memory specificity or categoricity. Our work thus adds further evidence of OGM being a symptom of depression that cuts across positive, neutral, and negative memories. The current meta-analysis highlights the importance of considering additional participant moderator variables and testing parameters when discussing OGM in depression.

Funding

This work was supported by the Georgia Institute of Technology [grant number DE00017885].

Data availability

We have included a link to the OSF page containing our data in the manuscript.

Acknowledgements

We are grateful to Sanya Khan for her assistance in conducting and cross-checking the results of the database search.

References

American Psychiatric Association, 1980. Diagnostic and Statistical Manual of Mental Disorders (3rd ed.). American Psychiatric Publishing,.

Anderson, R.J., Boland, J., Garner, S.R., 2016. Overgeneral past and future thinking in dysphoria: the role of emotional cues and cueing methodology. Memory 24 (5), 708–719. https://doi.org/10.1080/09658211.2015.1046134.

Arffa, S., 2007. The relationship of intelligence to executive function and non-executive function measures in a sample of average above average and gifted youth. Arch. Clin. Neuropsychol. 22 (8), 969–978. https://doi.org/10.1016/j.acn.2007.08.001.

Barry, T.J., Sze, W.Y., Raes, F., 2019. A meta-analysis and systematic review of memory specificity training (MeST) in the treatment of emotional disorders. Behav. Res. Ther. 116, 36–51. https://doi.org/10.1016/j.brat.2019.02.001.

Barulli, D., Stern, Y., 2013. Efficiency capacity compensation maintenance plasticity: emerging concepts in cognitive reserve. Trends Cogn. Sci. 17 (10), 502–509. https://doi.org/10.1016/j.tics.2013.08.012.

Brittlebank, A.D., Scott, J., Mark, J., Williams, G., Ferrier, I.N., 1993. Autobiographical memory in depression: state or trait marker. Br. J. Psychiatry 162 (1), 118–121. https://doi.org/10.1192/bjp.162.1.118.

Cheung, M.W.-L., 2014. Modeling dependent effect sizes with three-level meta-analyses: a structura; equation modeling approach. Psychol. Methods 19 (2), 211–229. https://doi.org/10.1037/a0032968.

Chiu, C.P.Y., Griffith, J.W., Lenaert, B., Raes, F., Hermans, D., Barry, T.J., 2018. Meta-analysis of the association between rumination and reduced autobiographical memory specificity. Memory 26 (10), 1323–1334. https://doi.org/10.1080/09658211.2018.1474928.

Coles, N.A., Larsen, J.T., Lench, H.C., 2019. A meta-analysis of the facial feedback literature: effects of facial feedback on emotional experience are small and variable (Jun)). Psychol. Bull. 145 (6), 610–651. https://doi.org/10.1037/bul0000194.

Conway, M.A., Pleydell-Pearce, C.W., 2000. The construction of autobiographical memories in the self-memory system. Psychol. Rev. 107 (2), 261–288. https://doi.org/10.1037/0033-295x.107.2.261.

Dalgleish, T., Williams, J.M.G., Golden, A.-M.J., Perkins, N., Barrett, L.F., Barnard, P.J., Au Yeung, C., Murphy, V., Elward, R., Tchanturia, K., Watkins, E., 2007. Reduced specificity of autobiographical memory and depression: the role of executive control. J. Exp. Psychol. Gen. 136 (1), 23–42. https://doi.org/10.1037/0096-3445.136.1.23.

Debeer, E., Hermans, D., Raes, F., 2009. Associations between components of rumination and autobiographical memory specificity as measured by a minimal instructions autobiographical memory test. Memory 17 (8), 892–903. https://doi.org/10.1080/09658210903376243.

Dotson, V.M., McClintock, S.M., Verhaeghen, P., Kim, J.U., Draheim, A.A., Syzmkowicz, S.M., Gradone, A.M., Bogoian, H.R., Wit, L.D., 2020. Depression and cognitive control across the lifespan: a systematic review and meta-analysis. Neuropsychol. Rev. 30 (4), 461–476. https://doi.org/10.1007/s11065-020-09436-6.

Duval, S., Tweedie, R., 2000. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56 (2), 455–463. https://doi.org/10.1111/j.0006-341x.2000.00455.x.

Gamble, B., Moreau, D., Tippett, L.J., Addis, D.R., 2019. Specificity of future thinking in depression: a meta-analysis. Perspect. Psychol. Sci. 14 (5), 816–834. https://doi.org/10.1177/1745691619851784.

Goddard, L., Dritschel, B., Burton, A., 1997. Social problem solving and autobiographical memory in non-clinical depression. Br. J. Clin. Psychol. 36 (3), 449–451. https://doi.org/10.1111/j.2044-8260.1997.tb01252.x.

Griffith, J.W., Sumner, J.A., Debeer, E., Raes, F., Hermans, D., Mineka, S., Zinbarg, R.E., Craske, M.G., 2009. An item response theory/confirmatory factor analysis of the autobiographical memory test. Memory 17 (6), 609–623. https://doi.org/10.1080/09658210902939348.

Gupta, R., Kar, B.R., 2012. Attention and memory biases as stable abnormalities among currently depressed and currently remitted individuals with unipolar depression. Front. Psychiatry *3*. https://doi.org/10.3389/fpsyt.2012.00099.

Hallford, D.J., Rusanov, D., Yeow, J.J.E., Barry, T.J., 2021. Overgeneral and specific autobiographical memory predict the course of depression: an updated meta-analysis. Psychol. Med. 51 (6), 909–926. https://doi.org/10.1017/s0033291721001343.

Hedges, L.V., Olkin, I., 1985. Statistical methods for meta-analysis. Academic Press,.

Heo, M., Murphy, C.F., Meyers, B.S., 2007. Relationship between the hamilton depression rating scale and the montgomery-åsberg depression rating scale in depressed elderly: a meta-analysis, 2007/10/01/). Am. J. Geriatr. Psychiatry. https://doi.org/10.1097/JGP.0b013e318098614e.

James, T., Weiss-Cowie, S., Hopton, Z., Verhaeghen, P., Dotson, V.M., Duarte, A., (In press). Depression and Episodic Memory Across the Adult Lifespan: A Meta-Analysis. Psychological Bulletin.

Karlsson, L., Pelkonen, M., Heilä, H., Holi, M., Kiviruusu, O., Tuisku, V., Ruuttu, T., Marttunen, M., 2007. Differences in the clinical characteristics of adolescent depressive disorders. Depress. Anxiety 24 (6), 421–432. https://doi.org/10.1002/da.20233.

Kopelman, M.D., Wilson, B.A., Baddeley, A.D., 1989. The autobiographical memory interview: a new assessment of autobiographical and personal semantic memory in amnesic patients. J. Clin. Exp. Neuropsychol. 11 (5), 724–744. https://doi.org/10.1080/01688638908400928.

Le Carret, N., Lafont, S., Letenneur, L., Dartigues, J.-F., Mayo, W., Fabrigoule, C., 2003. The effect of education on cognitive performances and its implication for the constitution of the cognitive reserve. Dev. Neuropsychol. 23 (3), 317–337. https://doi.org/10.1207/s15326942dn2303_1.

Leeson, V.C., Barnes, T.R.E., Harrison, M., Matheson, E., Harrison, I., Mutsatsa, S.H., Ron, M.A., Joyce, E.M., 2010. The relationship between iq, memory, executive function, and processing speed in recent-onset psychosis: 1-year stability and clinical outcome. Schizophr. Bull. 36 (2), 400–409. https://doi.org/10.1093/schbul/sbn100.

Lemogne, C., Piolino, P., Friszer, S., Claret, A., Girault, N., Jouvent, R., Allilaire, J.-F., Fossati, P., 2006. Episodic autobiographical memory in depression: specificity autonoetic consciousness and self-perspective. Conscious. Cogn. 15 (2), 258–268. https://doi.org/10.1016/j.concog.2005.07.005.

Liu, X., Li, L., Xiao, J., Yang, J., Jiang, X., 2013. Abnormalities of autobiographical memory of patients with depressive disorders: a meta-analysis. Psychol. Psychother. Theory Res. Pract. 86 (4), 353–373. https://doi.org/10.1111/j.2044-8341.2012.02077.x.

Luchetti, M., Sutin, A.R., 2018. Age differences in autobiographical memory across the adult lifespan: older adults report stronger phenomenology, 2018/01/02). Memory 26 (1), 117–130. https://doi.org/10.1080/09658211.2017.1335326.

Mackinger, H.F., Pachinger, M.M., Leibetseder, M.M., Fartacek, R.R., 2000. Autobiographical memories in women remitted from major depression. J. Abnorm. Psychol. 109 (2), 331–334. https://doi.org/10.1037/0021-843X.109.2.331.

Mansell, W., Lam, D., 2004. A preliminary study of autobiographical memory in remitted bipolar and unipolar depression and the role of imagery in the specificity of memory. Memory 12 (4), 437–446. https://doi.org/10.1080/09658210444000052.

Marchetti, I., Everaert, J., Dainer-Best, J., Loeys, T., Beevers, C.G., Koster, E.H.W., 2018. Specificity and overlap of attention and memory biases in depression. J. Affect. Disord. 225, 404–412. https://doi.org/10.1016/j.jad.2017.08.037.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. b2535-b2535 BMJ 339 (jul21 1). https://doi.org/10.1136/bmj.b2535.

Nandrino, J.L., Pezard, L., Poste, A., Reveillere, C., Beaune, D., 2002. Autobiographical memory in major depression: a comparison between first-episode and recurrent patients (Nov-Dec). Psychopathology 35 (6), 335–340. https://doi.org/10.1159/000068591.

Ono, M., Devilly, G.J., Shum, D.H.K., 2016. A meta-analytic review of overgeneral memory: the role of trauma history mood and the presence of posttraumatic stress disorder. Psychol. Trauma. Theory Res. Pract. Policy 8 (2), 157–164. https://doi.org/10.1037/tra0000027.

Opdebeeck, C., Martyr, A., Clare, L., 2016. Cognitive reserve and cognitive function in healthy older people: a meta-analysis. Aging Neuropsychol. Cogn. 23 (1), 40–60. https://doi.org/10.1080/13825585.2015.1041450.

Pustejovsky, J.E., Tipton, E., 2021. Meta-analysis with robust variance estimation: expanding the range of working models. Prev. Sci. https://doi.org/10.1007/s11121-021-01246-3.

Raes, F., Hermans, D., Williams, J.M.G., Demyttenaere, K., Sabbe, B., Pieters, G., Eelen, P., 2006. Is overgeneral autobiographical memory an isolated memory phenomenon in major depression. Memory 14 (5), 584–594. https://doi.org/10.1080/09658210600624614.

Raes, F., Hermans, D., Williams, J.M.G., Eelen, P., 2007. A sentence completion procedure as an alternative to the autobiographical memory test for assessing overgeneral memory in non-clinical populations. Memory 15 (5), 495–507. https://doi.org/10.1080/09658210701390982.

Raes, F., Williams, J.M.G., Hermans, D., 2009. Reducing cognitive vulnerability to depression: a preliminary investigation of memory specificity training (MEST) in inpatients with depressive symptomatology. J. Behav. Ther. Exp. Psychiatry 40 (1), 24–38. https://doi.org/10.1016/j.jbtep.2008.03.001.

Rawal, A., Rice, F., 2012. Examining overgeneral autobiographical memory as a risk factor for adolescent depression. J. Am. Acad. Child Adolesc. Psychiatry 51 (5), 518–527.

Shorey, S., Ng, E.D., Wong, C.H.J., 2022. Global prevalence of depression and elevated depressive symptoms among adolescents: a systematic review and meta-analysis. Br. J. Clin. Psychol. 61 (2), 287–305. https://doi.org/10.1111/bjc.12333.

Spinhoven, P., Bockting, C.L.H., Schene, A.H., Koeter, M.W.J., Wekking, E.M., Williams, J.M.G., 2006. Autobiographical memory in the euthymic phase of recurrent depression. J. Abnorm. Psychol. 115 (3), 590–600. https://doi.org/10.1037/0021-843x.115.3.590.

Sumner, J.A., Griffith, J.W., Mineka, S., 2010. Overgeneral autobiographical memory as a predictor of the course of depression: a meta-analysis. Behav. Res. Ther. 48 (7), 614–625. https://doi.org/10.1016/j.brat.2010.03.013.

Sumner, J.A., Griffith, J.W., Mineka, S., 2011. Examining the mechanisms of overgeneral autobiographical memory: capture and rumination and impaired executive control (Feb)). Memory 19 (2), 169–183. https://doi.org/10.1080/09658211.2010.541467.

Sumner, J.A., Mineka, S., Adam, E.K., Craske, M.G., Vrshek-Schallhorn, S., Wolitzky-Taylor, K., Zinbarg, R.E., 2014. Testing the CaR–FA–X model: investigating the mechanisms underlying reduced autobiographical memory specificity in individuals with and without a history of depression. J. Abnorm. Psychol. 123 (3), 471–486. https://doi.org/10.1037/a0037271.

Thompson, A.H., 2012. Childhood depression revisited: indicators normative tests and clinical course (Feb)). J. Can. Acad. Child Adolesc. Psychiatry 21 (1), 5–8.

Van Vreeswijk, M.F., De Wilde, E.J., 2004. Autobiographical memory specificity psychopathology depressed mood and the use of the autobiographical memory test: a meta-analysis. Behav. Res. Ther. 42 (6), 731–743. https://doi.org/10.1016/s0005-7967(03)00194-3.

Viechtbauer, W., 2010. Conducting meta-analyses in r with the metafor package. J. Stat. Softw. 36 (3), 1–48. https://doi.org/10.18637/jss.v036.i03.

Vittengl, J.R., Clark, L.A., Kraft, D., Jarrett, R.B., 2005. Multiple measures methods and moments: a factor-analytic investigation of change in depressive symptoms during acute-phase cognitive therapy for depression. Psychol. Med. 35 (5), 693–704. https://doi.org/10.1017/s0033291704004143.

Warne, N., Caseras, X., Rice, F., 2020. The cross-sectional and longitudinal relationship between overgeneral autobiographical memory and adolescent depression in a UK population-based cohort. J. Affect. Disord. 266, 621–625. https://doi.org/10.1016/j.jad.2020.02.011.

Wessel, I., Meeren, M., Peeters, F., Arntz, A., Merckelbach, H., 2001. Correlates of autobiographical memory specificity: the role of depression anxiety and childhood trauma. Behav. Res. Ther. 39 (4), 409–421. https://doi.org/10.1016/s0005-7967(00)00011-5.

Wickham, H., François, R., Henry, L., Müller, K., 2021. dplyr: A Grammar of Data Manipulation. In (Version 1.0.7) [R package]. https://CRAN.R-project.org/package=dplyr.

Williams, J.M., Broadbent, K., 1986. Autobiographical memory in suicide attempters (May)). J. Abnorm Psychol. 95 (2), 144–149. https://doi.org/10.1037//0021-843x.95.2.144.

Williams, J.M., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E., Dalgleish, T., 2007. Autobiographical memory specificity and emotional disorder (Jan)). Psychol. Bull. 133 (1), 122–148. https://doi.org/10.1037/0033-2909.133.1.122.

Young, K.D., Erickson, K., Drevets, W.C., 2012. Match between cue and memory valence during autobiographical memory recall in depression. Psychol. Rep. 111 (1), 129–148. https://doi.org/10.2466/09.02.15.pr0.111.4.129-148.

Young, K.D., Bellgowan, P.S.F., Bodurka, J., Drevets, W.C., 2014. Neurophysiological correlates of autobiographical memory deficits in currently and formerly depressed subjects. Psychol. Med. 44 (14), 2951–2963. https://doi.org/10.1017/S0033291714000464.

Young, K.D., Siegle, G.J., Bodurka, J., Drevets, W.C., 2016. Amygdala activity during autobiographical memory recall in depressed and vulnerable individuals: association with symptom severity and autobiographical overgenerality. Am. J. Psychiatry 173 (1), 78–89. https://doi.org/10.1176/appi.ajp.2015.15010119.