The Effects of Emotion on Judgments of Learning and Memory: A Meta-Analytic Review

Yue Yin[1], David R. Shanks[2], Baike Li[3], Tian Fan[1], Xiao Hu[4], Chunliang Yang[3,4*], Liang Luo[1,3,5*]

[1] Collaborative Innovation Center of Assessment for Basic Education Quality, Beijing Normal University, Beijing, China.

[2] Division of Psychology and Language Sciences, University College London, London, UK.

[3] Institute of Developmental Psychology, Faculty of Psychology, Beijing Normal University, Beijing, China.

[4] Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education, Beijing Normal University, Beijing, China.

[5] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China.

Author Note

The data contained in this project are publicly available at Open Science Framework (https://osf.io/3wkhm/). Correspondence concerning this article should be addressed to Liang Luo (luoliang@bnu.edu.cn) or Chunliang Yang (chunliang.yang@bnu.edu.cn), 19 Xinjiekou Wai Street, Beijing 100875, China.

Conflicts of Interest

We have no known conflict of interest to disclose.

Acknowledgments

This research was supported by the Natural Science Foundation of China (32171045; 32000742; 32200841), and the UK Economic and Social Research Council (ES/S014616/1).

Abstract

Emotional information pervades experiences in daily life. Numerous studies have established that emotional materials and information are easier to remember than neutral ones, a phenomenon known as the emotional salience effect on memory. In recent years, an emerging body of research has begun to explore the effect of emotion on metamemory. Preliminary findings show that participants offer higher judgments of learning (JOLs) to emotional than to neutral stimuli, a phenomenon termed the emotional salience effect on JOLs. The present meta-analysis integrated data from 1,887 participants, extracted from 17 qualifying studies, to examine the effects of emotion on JOLs and memory and to explore potential moderators of these effects. The results showed a medium-sized (g = 0.53 [0.41, 0.64]) emotional salience effect on JOLs, which was moderated by age and material type, as well as a small to medium (g = 0.38 [0.25, 0.51) emotional salience effect on memory, which was moderated by test format. These findings establish that emotionality is a salient cue in the theoretical framework of metamemory, and also provide some practical implications (e.g., in eyewitness testimony). However, more research is needed, especially employing high-powered pre-registered experiments, to address the signals of publication bias detected in this meta-analysis.

Keywords: Emotion; Judgments of learning; Memory; Metamemory; Meta-analysis

Emotional information pervades experiences in daily life, such as when children read emotionally-charged fairy tales or adults witness frightening crimes. The mnemonic consequences of variations in emotion have attracted substantial research interest over the last century, and numerous studies have established that emotional materials and events tend to be better remembered than neutral ones, a phenomenon referred to as the *emotional salience effect* on memory (Murphy & Isaacowitz, 2008). Besides investigating the effect of emotion on memory, an emerging body of studies has begun to explore the effect of emotion on metamemory – people's insight into their own memory status – but the results are somewhat fragmented and inconsistent (e.g., Efklides, 2016; Witherby et al., 2021). As emotion plays a key role in metamemory (Efklides, 2006, 2016; Fairfield et al., 2015), research on this topic is not only fundamental to assist the development of theoretical frameworks for metamemory, but also has the potential to address important practical implications (e.g., confidence and accuracy in eyewitness testimony). The current review conducts the first meta-analysis to clarify the effect of emotion on metamemory.

Metamemory can be defined as "knowledge about memory" (Flavell & Wellman, 1977), involving a complex set of cognitive processes including monitoring and control components (Nelson & Narens, 1990). Judgments of learning (JOLs; metacognitive estimates about the likelihood that a given item will be successfully recalled or recognized on a later test) are a widely-studied measure of metamemory monitoring (Rhodes, 2016). Accurately monitoring one's ongoing learning is critical for being a successful learner because individuals typically regulate their study activities (e.g., how to allocate limited study time, which learning strategy to select) according to their JOLs (Metcalfe & Finn 2008; Yang et al., 2017). Inaccurate JOLs generally lead to inefficient regulation of learning strategies and poor learning outcomes. For instance, students might stop learning course material prematurely if they are overconfident about their learning progress. In contrast, if a person is underconfident about the durability of their learning, they might unnecessarily expend extra efforts toward re-studying well-mastered material. Given the importance of JOLs in self-regulated learning, the current review seeks to explore whether (and if so, how) emotion, a prevailing factor experienced in daily life, affects JOL formation.

It is well-established that emotional materials differ from neutral ones on two fundamental dimensions, namely valence and arousal (Lang et al., 1990; Bliss-Moreau et al., 2020). Valence is defined as subjective positive-to-negative evaluations of the inherent emotional quality of information (Lang et al., 1993). Arousal refers to low-to-high activation of the sympathetic nervous system, which is associated with emotionally-laden items (Fairfield et al., 2015). Previous studies have suggested that both dimensions (valence and arousal) may contribute to JOL formation, but previous findings about their contributions are inconsistent (Hourihan et al., 2017; Zimmerman & Kelley, 2010). From a theoretical perspective, exploring whether these two dimensions of emotion contribute to JOL formation can sharpen our understanding about the mechanisms underlying the effect of emotionality on metamemory (see below for details). The current meta-analysis attempts to further explore this critical question through integrating the inconsistent results observed in previous studies.

Emotionality and JOLs

Much research confirms that emotion affects various aspects of cognition, such as attention and memory (Hamann, 2001; Murphy & Isaacowitz, 2008; Kensinger & Corkin, 2003; Palombo et al., 2021; Talmi & Moscovitch, 2004). Furthermore, recent studies have found that emotion also tends to affect JOLs by showing that participants provide higher JOLs to emotional than to neutral stimuli (Caplan et al., 2019; Nomi et al., 2013; Schmoeger et al., 2020; Tauber & Dunlosky, 2012; Witherby & Tauber, 2018; Zimmerman & Kelley, 2010), in other words an emotional salience effect on JOLs (Tauber et al., 2017).

According to the cue-utilization framework of metamemory (Koriat, 1997), people construct JOLs based on a range of cues, and emotionality can be well situated within this framework (Hourihan & Bursey, 2017; Zimmerman & Kelley, 2010). Emotion can affect JOLs through two modes: (1) theory-based (analytic) inference, and (2) experience-based (non-analytic) heuristics. Emotionality may influence JOLs in a theory-based way (Tauber et al., 2019; Zimmerman & Kelley, 2010). That is, people may believe that emotional stimuli are easier to remember than neutral ones, and therefore offer higher JOLs to emotional stimuli. Alternatively, emotionality may influence JOLs in an experience-based way (Hourihan & Bursey, 2017; Mitton, 2020). Compared with neutral stimuli, emotional ones capture greater levels of attention (Hamann, 2001), provoke physiological arousal (Mitton, 2020), or induce stronger subjective feelings (Hourihan et al., 2017), hence leading to higher JOLs to emotional than to neutral stimuli.

Some researchers have connected the two dimensions of emotion (valence and arousal) with these two theoretical explanations (beliefs and experiences) to account for the emotional salience effect on JOLs (Hourihan et al., 2017; Tauber et al., 2017; Witherby & Tauber, 2018; Witherby et al., 2021). Specifically, if emotion affects JOLs mainly via beliefs (i.e., a conscious strategy), valence should be the most salient cue contributing to the emotional salience effect on JOLs, because valence is a categorical descriptor (e.g., negative, positive, or neutral) involving subjective interpretations of individual items (for detailed discussion, see Witherby et al., 2021). Participants explicitly notice the categorical difference in valence among the stimuli and then provide higher JOLs for the emotional items based on the belief that emotional items are easier to remember than neutral ones. Conversely, if emotion affects JOLs primarily through processing experience (i.e., an implicit strategy), then arousal is expected be the main contributor to the emotional salience effect on JOLs, because arousal-provoking stimuli usually lead to greater physiological responses (a kind of processing experience; Hourihan et al., 2017; Tauber et al., 2017). Participants implicitly interpret feelings of arousal as being predictive of later memory performance and then assign higher JOLs to emotional items. Testing the contributions of valence and arousal to the emotional salience effect on JOLs can refine our understanding about how emotion affects JOLs.

Although previous studies have emphasized the important role of valence in the emotional salience effect on JOLs (Hourihan et al., 2017; Tauber et al., 2017), it remains unclear whether positive and negative emotions have different effects on JOLs. To investigate this question, Hourihan (2020) instructed undergraduates to study a mixed list of positive, negative, and neutral images, and make item-by-item JOLs. The results showed that participants provided higher JOLs to negative than to positive images. The same pattern was detected by Tauber and Dunlosky (2012), who employed word lists as learning materials. These findings suggest that negative emotion can produce a larger emotional salience effect on JOLs than positive emotion.

However, it must be highlighted that many other studies observed no difference between the effects of positive and negative emotion on JOLs. For instance, Nomi et al. (2013) found no difference in JOLs for faces with positive or negative expressions. Consistent with Nomi et al.'s (2013) findings, other studies found no statistically detectable difference using other types of learning materials, such as word lists (Gallant et al., 2019), word pairs (Zimmerman & Kelley, 2010) and images (Witherby, 2019).

Overall, previous results about whether positive and negative emotion have different effects on JOLs are rather inconsistent. Hence, a meta-analysis, integrating existing results to increase statistical power, is necessary to provide an answer to this question.

Besides valence, arousal is considered as another important cue to inform JOLs. For instance, Hourihan et al. (2017) found that participants provided higher JOLs to high-arousal than to low-arousal words. However, only a few studies have explored the effect of arousal on JOLs (e.g., Hourihan et al., 2017; Tauber et al., 2017). Hence, it is premature to draw any firm conclusion about its role in JOL formation.

In some previous studies which examined the effect of emotion on JOLs, the arousal levels between emotional and neutral stimuli were explicitly matched (e.g., Gallant et al., 2019; Tauber et al., 2017). By contrast, in other studies, arousal levels were unmatched. Specifically, in these unmatched studies, emotional stimuli were more arousal-provoking than neutral ones (e.g., Schmoeger et al., 2020; Tauber & Dunlosky, 2012; Zimmerman & Kelley, 2010). The current meta-analysis seeks to further explore the potential role of arousal in the emotional salience effect on JOLs by investigating whether arousal match between emotional and neutral stimuli moderates the emotional salience effect on JOLs. If arousal contributes to the construction of JOLs, a stronger emotional salience effect on JOLs would be observed in arousal-unmatched than in arousal-matched studies, since in arousal-unmatched studies the emotional stimuli are more arousing than the neutral ones.

Numerous studies have investigated the influence of aging on metacognition (Connor et al., 1997; Serra et al., 2008; Price et al., 2016), but it remains unclear whether metamemory monitoring varies as a function of age across adulthood (Hines et al., 2015; Kelley & Sahakyan, 2003). To further explore age differences in metamemory monitoring, several studies have investigated whether young and older adults exhibit different patterns of emotion effects on JOLs, but the results are again inconsistant, with some studies observing age differences (e.g., Sanders & Berry, 2021; Tauber & Dunlosky, 2012) and others not (e.g., Flurry, 2016; Gallant et al., 2019). In order to shed new light on our understanding of metacognitive aging, the current meta-analysis investigates if there is any age difference in the emotional salience effect on JOLs. Because prior findings on this question are conflicting, we claim that we had no *a priori* prediction regarding age differences.

To our knowledge, the question of whether material type (i.e., images vs. verbal stimuli) moderates the emotional salience effect on JOLs has never been explored. Given that images tend to be more provocative and contain richer emotional details than verbal stimuli

(Palombo et al., 2021; Schlochtermeier et al., 2013; Tauber et al., 2017), it is reasonable to assume that the emotional salience effect on JOLs would be stronger for images than for verbal stimuli. Another way for categorizing stimuli is based on whether the study stimuli were single items (e.g., single words or single images) or paired-associates (e.g., word pairs or image pairs). To avoid confusion with the category name “material type” discussed above (i.e., images vs. verbal stimuli), we define the category of “single items vs. paired associates” as stimulus type. Because JOLs for single items and for paired associates are formed based on different cues (e.g., cue-target relations are typically used as a cue to inform JOLs for paired associates, whereas item distinctiveness is generally used as a cue to inform JOLs for single items), it is important to explore whether emotion exerts differential effects on JOLs for single items and paired associates. To our knowledge, the question of whether stimulus type moderates the emotional salience effect on JOLs has also not been investigated. Hence, we had no *a priori* prediction about the moderating effect of stimulus type.

Additionally, although previous studies found that test format (for instance, recall versus recognition) reliably moderates the emotional salience effect on memory (Murphy & Isaacowitz, 2008), little research has been conducted to explore whether test format moderates the effect of emotion on JOLs (Zimmerman & Kelly, 2010). For exploratory purposes, the current meta-analysis also assesses the moderating effect of test format.

Emotionality and memory

Several meta-analyses observed a reliable emotional salience effect on memory (e.g., Murphy & Isaacowitz, 2008; Murty et al., 2010). However, among previous JOL studies, there is no consensus about whether emotion has an enhancing effect on memory. For instance, some JOL studies observed superior memory for emotional than for neutral stimuli (Tauber & Dunlosky, 2012; Tauber et al., 2017; West, 2021; Zimmerman & Kelley, 2010), whereas others found no difference (Flurry, 2016; Witherby & Tauber, 2018). Furthermore, some studies even observed poorer memory for emotional than for neutral stimuli (Caplan et al., 2019; Hourihan, 2020; Nomi et al., 2013).

Considering that existing results about the effect of emotionality on memory were substantially inconsistent among previous JOL studies, another aim of the current meta-analysis is to integrate existing results to determine whether emotionality produces a detectable effect on memory, as it does in studies not restricted to include measurement of JOLs (Murphy & Isaacowitz, 2008; Murty et al., 2010). It is critical to address this issue because if emotion has no influence on memory, then the emotional salience effect on JOLs would necessarily constitute a metacognitive illusion. Put differently, if emotional materials evoke higher JOLs but are no more memorable than neutral materials, then this must be indicative of misalignment between memory and metamemory. Finally, it is also important to determine whether the effect of emotionality on JOLs is different (i.e., larger or smaller) from the effect on memory, a key question that has little been explored in previous research.

Overview of the Current Meta-analytic Review

The main purpose of the current review is to examine the effect of emotionality on JOLs. To achieve this aim, a meta-analysis was performed to quantify the standardized difference in JOLs between emotional (positive and negative) and neutral stimuli. Meta-regression analyses were implemented to examine whether the effect of emotionality on JOLs is moderated by valence, arousal match, age, material type, stimulus type and test format. The same analyses were also performed to investigate the effect of emotionality on memory.

The accuracy of JOLs is often of critical importance. However, most of the included studies did not report sufficient data to calculate absolute accuracy of JOLs (i.e., signed difference between JOLs and memory performance). Therefore, to examine the potential difference between the effects of emotionality on JOLs and memory, we directly compared the effect size for the emotionality effect on JOLs with that for the equivalent effect on memory. Additionally, because only 10 (out of 17) studies provided sufficient data about relative accuracy of JOLs, the meta-analytic results regarding this aspect of JOLs are reported in the online supplementary materials.

Method

Literature Search

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009), the literature search was conducted by the first author and a research assistant in May 2022 in the following electronic databases: Web of Science, PubMed, PsycINFO, PsycARTICLES, ProQuest Dissertation & Theses Global Database, and Google Scholar. The search terms were [("emotion*" OR "mood*" OR "affective" OR "arousal" OR "valence") AND ("metamemory" OR "judgment* of learning" OR "judgement* of learning" OR "JOL*")]. To ensure that the meta-analysis is as comprehensive as possible, we also manually screened the reference lists and Google Scholar citations of four narrative reviews (Efklides, 2006, 2016; Fairfield et al., 2015; Witherby et al., 2021). In addition, some researchers, who have previously explored the emotional salience effect on JOLs, were contacted to obtain unpublished data (Witherby et al., 2022).

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria were as follows:

a. Only empirical studies written in English were considered.

b. Duplicates were excluded.

c. Only studies involving item-by-item JOLs were included. Studies on other forms of metamemory monitoring (e.g., feelings of knowing, retrospective confidence ratings about answer accuracy) or meta-comprehension were excluded (e.g., Baumeister et al., 2015; Hoy, 2018; Saoud, 2020; Strain et al., 2013). Studies asking participants to make JOLs for another person were excluded (Tauber et al., 2019). All studies included in the current meta-analysis collected immediate JOLs (i.e., JOLs made immediately after participants studied each item), and none of them collected delayed JOLs.

d. Only studies examining the differences in JOLs between emotional (either positive or negative) and neutral stimuli were included. Studies that combined positive and negative stimuli as “emotional materials” and did not separately report the results for positive and negative materials were excluded (e.g., Undorf & Bröder, 2020; Undorf et al., 2018). Studies which did not include a neutral (control) condition were excluded (e.g., Kelly & Metcalfe, 2011).

e. Only studies reporting sufficient data for effect size calculation were included.

Coding of Studies

All studies were independently coded by the first author and a research assistant. They agreed 99% on all coding information. All divergences were checked and settled through discussion. The coded moderators were as follows:

Valence. Based on the valence of the stimuli, the effects were divided into two sub-categories: (a) positive emotion and (b) negative emotion.

Arousal match. The effects were divided into three sub-categories according to whether the arousal levels between emotional and neutral stimuli were matched or not: (a) matched, (b) unmatched (i.e., emotional stimuli were more arousal-provoking than neutral ones), and (c) unknown (i.e., no available information to judge whether the arousal levels were matched or not). As an additional note, effects were coded as unmatched when the original article explicitly reported that the arousal level of emotional stimuli was significantly higher than that of neutral ones.

Age. The effects were coded into two sub-categories according to participants' age: (a) young adults (mean age ranging from 18.63 to 24.73) and (b) older adults (mean age ranging from 67.79 to 73.95).

Material type. The effects were divided into two sub-categories according to study materials: (a) verbal materials (including words and word pairs) and (b) images (including images of facial expressions, individual images and image pairs of animals, inanimate objects, landscapes and life scenes).

Stimulus type. The effects were divided into two sub-categories according to the type of study stimuli: (a) single items and (b) paired associates.

Test format. The effects were divided into three sub-categories according to test format: (a) free recall, (b) cued recall, and (c) recognition (including both old/new and forced choice recognition).

Effect Size Calculation

The effects of emotionality on both JOLs and memory were quantified as standardized differences (Cohen's *d*) between emotional (either positive or negative) and neutral stimuli. When test format in a given study was old/new recognition, Cohen's *d*s for memory were based on hit rates rather than discriminability (i.e., *d*'), because some studies did not report sufficient data to calculate *d*' (e.g., Gallant et al., 2019; Mitton, 2020). For the extraction of both JOLs and memory effects, we directly extracted the reported effect size values if Cohen's *d*s were reported in the original articles. Otherwise, the formulae provided by Borenstein et al. (2009) were employed to calculate them.

In order to reduce potential bias in effects with small sample sizes, we applied the bias correction function provided by Hedges (1982) to transform Cohen's *d*s into Hedges' *g*s. For within-subjects design effects, the correlation coefficients (*r*s) for dependent measures between emotional and neutral conditions were required to transform d_z to d_{rm}. Thus, we directly calculated *r*s for studies for which their raw data were available. In addition, for studies which reported group means, standard deviations (or standard errors), and paired-samples *t* values, we calculated *r*s using the formula provided by Morris and DeShon (2002).

In total, there were 28 JOL effects and 24 memory effects for which the corresponding *r*s could be calculated. We converted those *r*s into Fisher's *Z* scores (Silver & Dunlap, 1987) and then conducted three-level random-effects meta-analyses to obtain the weighted mean of these Fisher's *Z* scores. The results showed a positive correlation between dependent measures for both JOLs, $Z = 1.30$ [1.08, 1.52], $p < .001$, and memory, $Z = 0.36$ [0.26, 0.46], $p < .001$. These Z scores were then transformed back to *r*s, with $r = .86$ [.79, .91] for JOLs, and $r = .35$ [.26, .43] for memory. For the 48 JOL effects whose *r*s were unknown, their *r*s were set to .86, and for the 52 memory effects whose *r*s were unknown, their *r*s were set to .35.

Data Analysis Strategy

We conducted separate meta-analyses to examine the effects of emotionality on JOLs and memory. Since some effects were extracted from the same sample of participants in a single study, which might violate the assumption of independence, the meta-analyses were performed using three-level random-effects models (Cheung, 2014; Assink & Wibbelink, 2016) where the effects from the same sample of participants were coded as dependent. Q statistics were used to measure heterogeneity among effects, and significant heterogeneity was indicated by a Q test with $p \leq .05$ (Cochran, 1954). Additionally, I^2 within clusters of dependent effects (I^2_{within}; the percentage of the total variability of effects attributable to heterogeneity within clusters of dependent effects) and I^2 between effects based on independent samples ($I^2_{between}$; the percentage of the total variability of effects attributable to heterogeneity between effects based on independent samples) were estimated (Cheung, 2014). The typical within-study variance was estimated using the formula provided by Higgins and Thompson (2002).

To explain potential sources of heterogeneity, univariate three-level random-effects meta-regression analyses were performed. Considering that multivariate approaches can examine each moderator's effect while controlling for the effects of other moderators, multivariate three-level random-effects meta-regression analyses were also conducted. All analyses were conducted via the R *metafor* package (Viechtbauer, 2010). Harrer et al. (2021) recommended using *t* tests rather than the default *z* tests in multi-level meta-analysis. Accordingly, the current meta-analysis adopted *t* tests to determine significance of the weighted mean effect sizes and meta-regression coefficients.

To assess the likelihood that publication bias – the preferential publishing of statistically significant studies – leads to an inflated effect size estimate, we first examined the moderating role of publication year and publication status (published versus unpublished). Then we visually inspected the funnel plot of effect sizes against their precision (standard error, *SE*) for asymmetry. Finally, we corrected for publication bias using Robust Bayesian Meta-Analysis (RoBMA; Bartoš et al., 2022; Maier et al., in press). RoBMA – which has been shown to be superior to other bias-correction methods via several simulation studies (Bartoš et al., 2022; Maier et al., in press) – uses Bayesian model-averaging to combine estimates from multiple models — including PET, PEESE (Stanley & Doucouliagos, 2014) and selection models, both with and without publication bias. Each method is fit to the data and then the estimated effect is computed by weighting each of them by its likelihood, given the data. This method computes Bayes factors to quantify the evidence for the presence or absence of an effect as well as of heterogeneity and publication bias. We fit RoBMA assuming equal prior probabilities across model types. Because bias-correction methods are generally not applicable to multilevel meta-analysis, we applied RoBMA to a dataset in which effects from the same study were averaged so as to avoid dependencies between effects.

All data have been made publicly available at Open Science Framework (https://osf.io/3wkhm/).

Results

Our initial search returned 4,432 articles, and an additional 49 studies were identified through correspondence or via manually screening the reference lists of relevant literature reviews. In total, 17 studies were identified as qualifying for the meta-analyses, from which 76 JOL effects and 76 memory effects (including data from 1,887 participants) were extracted from 32 experiments.[1] The screening procedure is reported in a flowchart (see Fig. S1 in the online supplemental materials) and the characteristics of the included studies are presented in Table S1 in the online supplemental materials.

As an aid to readers, Table 1 summaries the main research questions explored in the current meta-analyses and the corresponding findings. Below we first report results relating to the effect of emotionality on JOLs, and then the effect of emotionality on memory. Finally, we compare the effect of emotionality on JOLs versus its effect on memory.

[1] Among the 32 experiments, 10 provided a single effect size, and the other 22 provided more than one effect sizes.

Effect of emotionality on JOLs

The weighted mean effect size, estimated by a three-level random-effects meta-analysis, was g = 0.53 [0.41, 0.64], $p < .001$, indicating a medium-sized emotional salience effect on JOLs. Participants gave significantly higher JOLs to emotional than to neutral stimuli. Heterogeneity among the effects was substantial, $Q(75) = 1212.47$, $p < .001$, I^2_{within}= 7.40%, $I^2_{between}$ = 87.71%, indicating the need to conduct moderator analyses to identify possible sources of heterogeneity.

Random-effects meta-regression analyses were conducted using both univariate and multivariate models to detect potential sources of heterogeneity. The results showed similar patterns in univariate and multivariate analyses. Because the number of included effects is relatively small and many moderators were tested here, below we report the results from the univariate meta-analyses. The main results are shown in Table 2.

Moderator Analyses

Valence. There was no statistically detectable moderating effect of valence, $F(1, 74)$ = 0.18, $p = .673$. Both positive, g = 0.53 [0.41, 0.66], $p < .001$, and negative, g = 0.52 [0.39, 0.64], $p < .001$, stimuli received significantly higher JOLs than neutral ones. Overall, these results suggest little difference in the emotional salience effects on JOLs between positive and negative emotion.

Arousal match. Arousal match did not significantly moderate the emotional salience effect on JOLs, $F(2, 73) = 2.62$, $p = .079$. Regardless of whether arousal levels between emotional and neutral stimuli were matched, g = 0.48 [0.26, 0.70], $p < .001$, or unmatched, g

= 0.58 [0.46, 0.70], $p < .001$, there was a reliable emotional salience effect on JOLs. Excluding cases where the approach to matching was unknown, there was little difference between arousal-matched and arousal-unmatched effects, $F(1, 73) = 0.85$, $p = .359$. In other words, the effect size for the emotional salience effect on JOLs in arousal-matched studies (i.e., when the arousal levels between emotional and neutral stimuli were matched) was roughly equal to that in arousal-unmatched studies (i.e., when the arousal levels were significantly higher for emotional than for neutral stimuli). These findings suggest that arousal tends to contribute minimally to the emotional salience effect on JOLs.

The effect for the unknown category was not statistically significant, $g = 0.19$ [-0.15, 0.53], $p = .267$, possibly due to low statistical power as there were only $k = 6$ effects in this category. Considering that the number of effects in the unknown category was too small to generate a reliable conclusion, we do not discuss this result further.

Age. The moderating effect of age was significant, $F(1, 74) = 13.73$, $p < .001$, with young adults demonstrating a larger emotional salience effect on JOLs, $g = 0.62$ [0.51, 0.73], $p < .001$, than older adults, $g = 0.15$ [-0.08, 0.37], $p = .204$. These results reflect that young adults' JOLs are more sensitive to emotion than those of older adults.

Material type. Material type was a significant moderator, $F(1, 74) = 11.54$, $p = .001$, with a larger emotional salience effect on JOLs for images, $g = 0.73$ [0.57, 0.89], $p < .001$, than for verbal materials, $g = 0.38$ [0.24, 0.51], $p < .001$.

Stimulus type. There was no statistically detectable moderating effect of stimulus type, $F(1, 74) < 0.001$, $p = .977$. Regardless of whether the study stimuli were single items, g =

0.53 [0.40, 0.65], $p < .001$, or paired associates, $g = 0.53$ [0.14, 0.93], $p = .009$, JOLs were always higher for emotional than for neutral items.

Test format. Test format did not significantly moderate the emotional salience effect on JOLs, $F(2, 73) = 2.50$, $p = .089$. All test formats were associated with significant emotional salience effects on JOLs, with $g = 0.58$ [0.14, 1.02], $p = .01$, for cued recall tests, $g = 0.43$ [0.29, 0.57], $p < .001$, for free recall tests, and $g = 0.71$ [0.50, 0.91], $p < .001$, for recognition tests.

Publication Bias

A three-level meta-regression analysis was performed to examine the relationship between effect sizes and year of publication (ranging from 2010 to 2021), which showed a significantly positive relationship between the two variables, $b = 0.04$, $F(1, 74) = 7.35$, $p = .008$ (see Fig. S2 in the online supplemental materials). It should be noted that if an effect is spurious, then the observed effect sizes ought to gradually decrease across years of publication (Borenstein & Cooper, 2009). The current meta-regression analysis found the opposite result that the observed effect sizes increased across years, indicating little risk of sequence-based publication bias. It is possible that some aspect(s) of the research methods changed across time, increasing the obtained effect sizes. Regarding publication status, 62 effects were coded as published and the remaining 14 were unpublished. A three-level sub-group meta-analysis showed no significant moderating effect of publication status, $F(1, 74) = 0.05$, $p = .829$, again suggesting little need to worry about publication bias.

Fig. 1 shows the funnel plot of effect sizes against standard error (Sterne et al., 2011), which is asymmetric to some extent. The RoBMA analysis revealed very strong evidence of residual heterogeneity, *BF* = 7.34e+66, but little evidence of an overall emotional salience effect on JOLs, *BF* = 0.69, together with strong evidence of publication bias, *BF* = 24.55. The Bayes factor for the emotional salience effect indicates that, if anything, the evidence is slightly more consistent with there being no effect than with there being one, and the mean estimated effect is 0.10 [-0.17, 0.52]. Thus, when corrected for bias, the effect becomes very small, though with a wide estimation interval. We further discuss the RoBMA results in the General Discussion section.

Effect of emotionality on memory

We next turn to the effect of emotionality on memory. A three-level random-effects meta-analysis showed that the weighted mean effect size was $g = 0.38$ [0.25, 0.51], $p < .001$, indicating a small to medium emotional salience effect on memory. There was substantial heterogeneity among the effects, $Q(75) = 485.01$, $p < .001$, $I^2_{within} = 20.33\%$, $I^2_{between} = 63.31\%$.

Moderator Analyses

Random-effects meta-regression analyses were conducted using both univariate and multivariate approaches to identify potential sources of heterogeneity. Univariate and multivariate analyses yielded converging results. Hence, below we focus on the univariate analyses (see Table 2).

Valence. The moderating effect of valence was not statistically reliable, $F(1, 74) = 1.02$, $p = .317$, with $g = 0.34$ [0.19, 0.50], $p < .001$, for positive emotion, and $g = 0.43$ [0.27, 0.58], $p < .001$, for negative emotion.

Arousal match. The moderating effect of arousal match was not statistically significant, $Q(2, 73) = 1.83$, $p = .168$. Unmatched-arousal, $g = 0.42$ [0.28, 0.57], $p < .001$, and matched-arousal, $g = 0.44$ [0.12, 0.75], $p = .008$, generated similar emotional salience effects on memory, $F(1, 73) = 0.01$, $p = .935$. The emotional salience effect on memory for the unknown category was not statistically detectable, $g = 0.02$ [-0.39, 0.42], $p = .937$.

Age. The moderating effect of age was not significant, $F(1, 74) = 3.22$, $p = .077$, with $g = 0.33$ [0.19, 0.47], $p < .001$, for young adults, and $g = 0.63$ [0.33, 0.93], $p < .001$, for older adults.

Material type. The moderating effect of material type was not statistically detectable, $F(1, 74) = 2.26$, $p = .137$, with $g = 0.27$ [0.07, 0.47], $p = .008$, for images, and $g = 0.47$ [0.30, 0.64], $p < .001$, for verbal materials.

Stimulus type. The moderating effect of stimulus type was not statistically detectable, $F(1, 74) = 1.92$, $p = .170$, with $g = 0.41$ [0.27, 0.55], $p < .001$, for single items, and $g = 0.08$ [-0.36, 0.53], $p = .709$, for paired associates.

Test format. The moderating effect of test format was significant, $F(2, 73) = 12.76$, $p < .001$. The emotional salience effect on memory was larger in free recall, $g = 0.59$ [0.45, 0.73], $p <.001$, than in recognition tests, $g = 0.01$ [-0.18, 0.20], $p = .910$, $F(1, 73) = 24.83$, $p < .001$. There was no statistically detectable difference between cued recall, $g = 0.21$ [-0.22, 0.64], $p = .340$, and free recall tests, $F(1, 73) = 2.87$, $p = .094$, nor between cued recall and recognition tests, $F(1, 73) = 0.70$, $p = .406$.

Publication Bias

There was no statistically detectable relationship between effect size and publication year, $b = -0.01$, $F(1, 74) = 0.64$, $p = .425$ (see Fig. S3 in the online supplemental materials), and no detectable moderating effect of publication status, $F(1, 74) = 1.31$, $p = .257$. These results jointly suggest little need to worry about publication bias of the included studies.

The funnel plot is displayed in Fig. 2, which is again asymmetric to some extent. RoBMA revealed a pattern similar to that for the emotional salience effect on JOLs: Very strong evidence of residual heterogeneity, $BF = 2.03e+12$, but little evidence of an overall emotional salience effect on memory, $BF = 0.34$, and strong evidence of publication bias, $BF = 305.15$. The Bayes factor for the emotional salience effect on memory indicates that, if anything, the evidence is about 3 times more consistent with there being no effect than with there being one, and the mean estimated effect is -0.08 [-0.78, 0.09]. Thus, when corrected for bias, the emotional salience effect on memory becomes negligible.

Difference between the effects of emotion on JOLs and memory

To determine if emotion has different effects on JOLs and memory, we combined the JOL and memory effects, and conducted a three-level meta-regression analysis, with effect type (JOL versus memory) as a moderator. The results revealed a significant moderating role of effect type, $F(1, 150) = 6.54$, $p = .012$, indicating that the emotional salience effect was significantly larger on JOLs than on memory.[2] Thus, even though people can metacognitively recognize the effect of emotion on memory (that is, JOLs and memory vary in the same direction as a function of emotionality), they tend to overestimate the actual magnitude of the emotional salience effect on memory.

General Discussion

The current review conducted the first meta-analysis to examine the effect of emotionality on JOLs. The results showed a medium-sized (g = 0.53) emotional salience effect on JOLs and a small-to-medium sized (g = 0.38) emotional salience effect on memory. A cutting-edge method for correction publication bias, RoBMA (Bartoš et al., 2022; Maier et al., in press), indicated that when corrected, both the residual emotional salience effect on JOLs and the effect on memory are small, with Bayes factors favoring of the null hypothesis. Below, we first discuss the main meta-analysis results and then comment on publication bias.

Effect of emotionality on JOLs

In line with previous studies, the meta-analysis showed an emotional salience effect on JOLs. According to previous studies, emotion affects JOLs through two distinct pathways (Koriat, 1997; Koriat et al., 2004; Mueller & Dunlosky, 2017; Witherby et al., 2021). One is theory-based inference (i.e., beliefs), which is linked more with valence. That is, emotionally-valenced stimuli have properties that make them distinctive from neutral ones, leading participants to hold the belief that they are more memorable than neutral ones (Tauber et al.,

[2] This result was based on the direct mixing of JOL effects with memory effects and then testing the moderating effect of dependent variable (JOL versus memory), which did not take publication bias into account. Indeed, due to technical limitations, it is difficult to compare the difference between JOL and memory effects in the case where publication bias is accounted for. For more details on the interpretation of publication bias, see the General Discussion section.

2017; Witherby & Tauber, 2018). Another is experience-based heuristics (i.e., processing experience), which are linked more with arousal. Participants experience different physiological reactions (e.g., heart rate, eye movements, brain signaling, skin conductance responses) when encoding high-arousal (emotional) stimuli relative to low-arousal (neutral) ones, and thus provide higher JOLs to emotional stimuli based on their processing experience (Hourihan et al., 2017).

According to the aforementioned theoretical explanations, if arousal contributes to JOL formation, a larger emotional salience effect on JOLs should be observed in arousal-unmatched (i.e., emotional stimuli were more arousal-provoking than neutral ones) than in arousal-matched studies. However, the meta-analysis found that arousal match failed to moderate the effect of emotionality on JOLs, with roughly equal effect sizes in arousal-matched and arousal-unmatched studies. Hence, it is reasonable to speculate that physiological arousal plays a minimal role in the emotional salience effect on JOLs. The emotional salience effect on JOLs may derive from the fact that participants hold the belief that emotionally-valenced stimuli are easier to remember. Previous studies did provide support for this explanation (Witherby & Tauber, 2018; Witherby et al., 2022). For instance, Witherby et al. (2022, Experiment 6) found that 96.9% of participants believed that their memory is better for emotional than for neutral stimuli. Further supporting evidence comes from Undorf and Bröder (2020). Undorf and Bröder instructed participants to make pre-study JOLs (that is, JOLs provided before participants saw and studied each item) and found that pre-study JOLs were higher for emotional than for neutral stimuli.

To sum up, emotion may affect JOLs mainly through beliefs, and processing experience contributes minimally to the emotional salience effect on JOLs. These findings support the analytic processing (AP) theory (Mueller & Dunlosky, 2017), which emphasizes the central role of beliefs in JOL formation. According to the AP theory, when participants are instructed to predict their future memory performance, they try to search for cues (e.g., concreteness, semantic relatedness) which they believe are related to future memory performance, and then they make JOLs based on those identified cues. Importantly, the AP theory proposes that JOLs are largely driven by *a priori* or newly developed beliefs, whereas processing experience plays a less important (or even no) role in JOL formation. Consistent with the AP theory, the current meta-analysis found that emotional valence significantly affected JOLs (that is, both positive and negative emotion enhanced JOLs), whereas arousal tended to contribute minimally to JOL formation.

It has to be acknowledged that meta-analysis only provides a blunt instrument to test theoretical accounts, and it is premature to draw a firm conclusion about the mechanisms underlying the emotional salience effect on JOLs based on the meta-analytic results observed here. It is possible that, besides physiological arousal, emotion affects JOLs through other types of processing experience, such as processing fluency (Witherby et al., 2021). Hence, it is too soon to completely rule out any role of processing experience in the emotional salience effect on JOLs. More experimental research on the underlying mechanisms is called for.

Another noteworthy is that the emotional salience effect on JOLs was numerically larger in arousal-unmatched (g = 0.58) than in arousal-matched studies (g = 0.48), even though the difference was not statistically significant (p = .359). A common methodological issue in detecting moderating effects in a meta-analysis is that second order sampling error stemming from the random sampling of studies affects the precision of the meta-analytic estimates, especially when a small number of studies are included (Hunter & Schmidt, 2004). In other words, it is difficult to quantify how much of the variance across meta-analytic estimates is explained by the moderators due to the presence of second order sampling error, and the statistical power of sub-group meta-analyses is generally low (Cuijpers et al., 2021; Griffin, 2021). Hence, we strongly recommend researchers to conduct updated meta-analyses to re-assess the moderating role of arousal match (or other variables) in the emotional salience effect on JOLs when more data are available.

Effect of emotionality on memory

Previous meta-analyses, which assessed the influences of emotion on memory when making JOLs was not required, found small-to-medium sized emotional salience effects on memory. For instance, Murphy and Isaacowitz (2008) found that both young ($\bar{d}_w$= 0.46, p < .01) and older adults ($\bar{d}_w$= 0.41, p < .01) exhibited superior memory performance for emotional than for neutral materials. Consistently, the current meta-analysis observed a small-to-medium sized emotional salience effect on memory in the studies in which making JOLs was required.

The results obtained here confirm a memory advantage for emotional information. There are several explanations for this phenomenon. The first possibility is that emotional stimuli involuntarily capture greater attention than neutral ones (Murphy & Isaacowitz, 2008; Yiend, 2010), thus leading to better memory encoding. It is also possible that the characteristics of emotional stimuli make them stand out when mixed with neutral ones, and the distinctiveness associated with emotional stimuli contributes to the emotional salience effect on memory (Schmidt & Saari, 2007; Talmi, 2013). Additionally, individuals may realize that emotional stimuli are inherently more interlinked with each other (e.g., emotionally-negative words, such as *gun* and *injury*, are semantically related) than randomly selected neutral stimuli (e.g., *book*, *pond*), and the structural inter-item organization of emotional stimuli may hence produce superior memory (Palombo et al., 2021; Talmi & Moscovitch, 2004).

Difference between the effects of emotionality on JOLs and memory

The emotional salience effects on JOLs and on memory jointly imply that people's JOLs are at least somewhat accurate because JOLs and memory vary in the same direction as a function of emotion (that is, people metacognitively appreciate the enhancing effect of emotion on memory). However, the meta-analysis found that the magnitude of the JOL effect was larger than that of the memory effect, suggesting that people tend to metacognitively overestimate the emotional salience effect on memory. To our knowledge, no previous studies have explored why this happens, and future studies could profitably address this issue. More importantly, interventions should be developed to reduce such a metacognitive illusion, which is of practical importance for eyewitness testimony.

Age

The meta-analysis found that young adults showed a larger emotional salience effect on JOLs than older adults, suggesting that young adults' JOLs are more sensitive to emotionality than those of older adults. A possible explanation is that young and older adults differ either in their beliefs or in experience (e.g., fluency) when processing emotionally-valenced stimuli, leading to age-related differences in the emotional salience effect on JOLs (Tauber & Dunlosky, 2012). Another possibility is that older adults' cognitive resources are limited (Zacks et al., 2000), and concurrently making JOLs and performing the learning task is highly challenging for them (Tauber & Witherby, 2019). The requirement of making item-by-item JOLs diverts older adults' attention from the encoding task, in turn leading to a smaller emotional salience effect on JOLs for older than for young adults. The reasons why older adults' JOLs are less sensitive to emotion are clouded by lack of relevant studies and deserves further investigation.

Material type

The meta-analysis observed that images produced a greater emotional salience effect on JOLs than verbal materials, while at the same time there was no statistically detectable difference in the emotional salience effects on memory between these two types of materials. A persuasive explanation for the larger emotional salience effect on JOLs for images is that such materials (e.g., facial expressions, scenes images) contain more emotionally-relevant details than verbal materials (Bradley et al., 2001; Hinojosa et al., 2009; Tauber et al., 2017). Rich and salient emotional cues, delivered by images, provoke people's beliefs about how emotion affects memory, in turn leading to a stronger effect on JOLs.

Test format

It is hardly surprising that test format did not moderate the emotional salience effect on JOLs because JOLs are provided before the test phase (for related findings, see Chang & Brainerd, 2022). By contrast, previous studies established that test format reliably moderated the emotional salience effect on memory (Charles et al., 2003; Hourihan, 2020). For instance, the meta-analysis conducted by Murphy and Isaacowitz (2008) found a moderating role of test format in the emotional salience effect on memory, with a larger effect in recall than in recognition tests. The same result pattern was observed here: Free recall tests were associated with a larger emotional salience effect on memory by comparison with recognition tests. The larger emotional salience effect on memory in free recall tests may result from the stronger semantic cohesion or relatedness among emotional items (see above). Overall, the emotional salience effect on memory, but not the emotional salience effect on JOLs, is moderated by test format, reflecting a dissociation between JOLs and memory.

Publication bias

The meta-analysis found evidence of publication bias from the visually asymmetric funnel plots displayed in Fig. 1 and 2, and from RoBMA (Bartoš et al., 2022; Maier et al., in press), indicating that this set of studies is probably contaminated by publication bias. In particular, the funnel plots suggest an absence of moderate-precision studies with small effect sizes. For both the residual emotional salience effects on JOLs and on memory, when corrected for publication bias, the Bayes factors if anything support the null hypothesis. It has to be acknowledged that the discussion and interpretation of key findings should be read in light of publication bias.

However, it is also important not to overinterpret these publication bias findings. In other words, we highly recommend not taking them as incontrovertible evidence against the key effects. First, visual inspection of funnel plots is subjective. Secondly, the confidence intervals on the bias-corrected effect sizes are wide (particularly for the emotional salience effect on JOLs) and do not exclude a medium-to-large true effect. Thirdly, it is well-known that bias-correction methods, including funnel plots and RoBMA, are imperfect and their accuracy depends on properties of the dataset that are unknowable, such as the true level of heterogeneity. Although the performance of RoBMA has been shown to be superior to other bias-correction methods by applying PET, PEESE and selection models to the data simultaneously, the meta-analytic estimate might still suffer from inaccurate estimation if none of the models approximate the data generating process well (Bartoš et al., 2022). Fourthly, the meta-analysis demonstrated minimal difference in the effects of emotionality on both JOLs and memory between published and unpublished studies, which mitigates potential concern about the risk of bias due to unpublished results. In addition, the emotional salience effects on both JOLs and memory did not systematically fluctuate as a function of publication year, again suggesting little need to worry about publication bias.

There is substantial heterogeneity among the included JOL and memory effects. Even if all of the included studies were well-designed (e.g., pre-determined their sample sizes before data-collection) and suffered from no publication bias, there would still be a negative correlation between effect sizes and sample sizes (or a positive correlation between effect sizes and standard errors) because large effect sizes require smaller numbers of participants to achieve a specific statistical power (Peters et al., 2010; Terrin et al., 2003). In that event, RoBMA would spuriously detect misleading evidence of "publication bias". Moreover, the evidence obtained here in no way suggests that this particular domain is any more tainted by publication bias than many other domains in behavioral research (see Kvarven et al., 2020). Rather, this evidence of publication bias emphasizes the pressing need for future research employing high-powered, pre-registered, confirmatory experiments.

Conclusion

Emotionality has a medium-sized salience effect on JOLs and a small-to-medium sized salience effect on memory, and the effect on JOLs is larger than the effect on memory. Both positive and negative emotion produce an emotional salience effect on JOLs. Arousal tends to contribute minimally to the emotional salience effect on JOLs. Young adults' JOLs are more sensitive to emotion than those of older adults. Image materials produce a larger emotional salience effect on JOLs than verbal materials. Test format moderates the emotional salience effect on memory, but not the effect on JOLs. All of the above results are tentative, however, in light of potential publication bias detected in this literature.

References

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

Assink, M., & Wibbelink, C. J. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12*(3), 154-174. https://doi.org/10.20982/tqmp.12.3.p154

Bartoš, F., Maier, M., Quintana, D. S., & Wagenmakers, E. J. (2022). Adjusting for publication bias in JASP and R: Selection models, PET-PEESE, and robust Bayesian meta-analysis. *Advances in Methods and Practices in Psychological Science, 5*(3). https://doi.org/10.1177/25152459221109259

Baumeister, R. F., Alquist, J. L., & Vohs, K. D. (2015). Illusions of learning: Irrelevant emotions inflate judgments of learning. *Journal of Behavioral Decision Making, 28*(2), 149-158. https://doi.org/10.1002/bdm.1836

Bliss-Moreau, E., Williams, L. A., & Santistevan, A. C. (2020). The immutability of valence and arousal in the foundation of emotion. *Emotion, 20*(6), 993-1004. https://doi.org/10.1037/emo0000606

Borenstein, M., & Cooper, H. (2009). *The handbook of research synthesis and meta-analysis* (Vol. 2). New York, NY: Russell Sage Foundation.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Converting among effect sizes. In U. Chichester (Ed.), *Introduction to meta-analysis* (pp. 45-49). Hoboken, NJ: Wiley. https://doi.org/10.1002/9780470743386.ch7

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation

I: defensive and appetitive reactions in picture processing. *Emotion, 1*(3), 276-298. http://dx.doi.org/10.1037/1528-3542.1.3.276

*Caplan, J. B., Sommer, T., Madan, C. R., & Fujiwara, E. (2019). Reduced associative memory for negative information: Impact of confidence and interactive imagery during study. *Cognition and Emotion*, *33*(8), 1745-1753. https://doi.org/10.1080/02699931.2019.1602028

Chang, M., & Brainerd, C. J. (2022). Association and dissociation between judgments of learning and memory: A Meta-analysis of the font size effect. *Metacognition and Learning, 17*(2), 443-476. https://doi.org/10.1007/s11409-021-09287-3

Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: The forgettable nature of negative images for older adults. *Journal of Experimental Psychology. General, 132*(2), 310-324. https://doi.org/10.1037/0096-3445.132.2.310

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychological Methods, 19*(2), 211. https://psycnet.apa.org/doi/10.1037/a0032968

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics, 10*(1), 101-129. https://doi.org/10.2307/3001666

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and aging, 12*(1), 50-71. https://psycnet.apa.org/doi/10.1037/0882-7974.12.1.50

Cuijpers, P., Griffin, J. W., & Furukawa, T. A. (2021). The lack of statistical power of subgroup analyses in meta-analyses: A cautionary note. *Epidemiology and Psychiatric Sciences, 30*. https://doi.org/10.1017/S2045796021000664

Denburg, N. L., Buchanan, T. W., Tranel, D., & Adolphs, R. (2003). Evidence for preserved emotional memory in normal older persons. *Emotion, 3*(3), 239-253. https://doi.org/10.1037/1528-3542.3.3.239

Mueller, M. L., & Dunlosky, J. (2017). How beliefs can impact judgments of learning: Evaluating analytic processing theory with beliefs about fluency. *Journal of Memory and Language, 93*, 245-258. https://doi.org/10.1016/j.jml.2016.10.008

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1*(1), 3-14. https://doi.org/10.1016/j.edurev.2005.11.001

Efklides, A. (2016). Metamemory and Affect. In J. Dunlosky & S.K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 245–267). Oxford University Press.

Fairfield, B., Mammarella, N., Palumbo, R., & Di Domenico, A. (2015). Emotional meta-memories: A review. *Brain Sciences, 5*(4), 509-520. https://doi.org/10.3390/brainsci5040509

Flavell, J. H., & Wellman, H. M. (1977). *Metamemory*. Hillsdale, NJ: Erlbaum.

*Flurry, S. E. (2016). *Examining age differences in metamemory for emotional words* (Order No. 10244251). [Master's Thesis, Mississippi State University]. ProQuest Dissertations and Theses Global. https://www.proquest.com/openview/6742607e9958fc8fdac650ba859e4d2e/1?pq-origsite=gscholar&cbl=18750

*Gallant, S. N., Spaniol, J., & Yang, L. (2019). Age differences in cue utilization during prospective and retrospective memory monitoring. *Psychology and Aging, 34*(4), 545-557. http://dx.doi.org/10.1037/pag0000352

Griffin, J. W. (2021). Calculating statistical power for meta-analysis using metapower. *The Quantitative Methods for Psychology. 17*(1), 24–39. https://doi.org/doi:10.20982/tqmp.17.1.p024

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences, 5*(9), 394-400. https://doi.org/10.1016/S1364-6613(00)01707-1

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). Doing meta-analysis with R: A hands-on guide. Chapman and Hall/CRC. https://doi.org/10.1201/9781003107347

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92*(2), 490–499. https://doi.org/http://dx.doi.org/10.1037/0033-2909.92.2.490

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in medicine, 21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Hines, J. C., Hertzog, C., & Touron, D. R. (2015). Younger and older adults weigh multiple cues in a similar manner to generate judgments of learning. *Aging, Neuropsychology, and Cognition, 22*(6), 693-711. https://doi.org/10.1080/13825585.2015.1028884

Hinojosa, J. A., Carretié, L., Valcárcel, M. A., Méndez-Bértolo, C., & Pozo, M. A. (2009). Electrophysiological differences in the processing of affective information in words and pictures. *Cognitive, Affective, & Behavioral Neuroscience, 9*(2), 173-189. https://doi.org/10.3758/CABN.9.2.173

*Hourihan, K. L. (2020). Misleading emotions: Judgments of learning overestimate recognition of negative and positive emotional images. *Cognition and Emotion, 34*(4), 771-782. https://doi.org/https://doi.org/10.1080/02699931.2019.1682972

*Hourihan, K. L., & Bursey, E. (2017). A misleading feeling of happiness: Metamemory for positive emotional and neutral pictures. *Memory, 25*(1), 35-43. https://doi.org/10.1080/09658211.2015.1122809

*Hourihan, K. L., Fraundorf, S. H., & Benjamin, A. S. (2017). The influences of valence and arousal on judgments of learning and on recall. *Memory & Cognition, 45*(1), 121-136. https://doi.org/10.3758/s13421-016-0646-3

Hoy, R. C. (2018). *Emotion and metacognitive monitoring: The role of emotion in the development of learning beliefs* (Order No. 10791288). [Doctoral dissertation, The University of New Mexico]. ProQuest Dissertations and Theses Global. https://www.proquest.com/docview/2088448703?pq-origsite=primo

Hunter, J. E., & Schmidt, F. L. (2004). Methods of meta-analysis: Correcting error and bias in research findings（2nd ed.). Thousand Oaks, CA: Sage.

Kelly, K. J., & Metcalfe, J. (2011). Metacognition of emotional face recognition. *Emotion, 11*(4), 896-906. https://doi.org/10.1037/a0023746

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48*(4), 704-721.

https://doi.org/10.1016/S0749-596X(02)00504-1

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition, 31*(8), 1169-1180. https://doi.org/10.3758/BF03195800

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*(4), 349-370. https://doi.org/10.1037/0096-3445.126.4.349

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based orocesses. *Journal of Experimental Psychology: General, 133*(4), 643-656. https://doi.org/10.1037/0096-3445.133.4.643

Kvarven, A., Strømland, E., & Johannesson, M. (2020). Comparing meta-analyses and preregistered multiple-laboratory replication projects. *Nature Human Behaviour, 4*(4), 423-434. https://doi.org/10.1038/s41562-019-0787-z

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review, 97*(3), 377-395. https://doi.org/10.1037/0033-295X.97.3.377

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3). 261-273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Maier, M., Bartoš, F., & Wagenmakers, E.-J. (in press). Robust Bayesian meta-analysis: Addressing publication bias with model-averaging. *Psychological Methods*. https://doi.org/10.31234/osf.io/u4cns

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review, 15*(1), 174-179. https://doi.org/10.3758/PBR.15.1.174

*Mitton, E. (2020). *Understanding the influence of emotion on judgments of learning (JOLs): Using physiological measures to probe the contribution of experience-based processes*. [Master's Thesis, The University of Guelph]. https://atrium.lib.uoguelph.ca/xmlui/handle/10214/21191

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–269. https://doi.org/http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00135

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105-125. https://doi.org/10.1037//1082-989X.7.1.105

Murphy, N. A., & Isaacowitz, D. M. (2008). Preferences for emotional information in older and younger adults: a meta-analysis of memory and attention tasks. *Psychology and Aging, 23*(2), 263-286. https://doi.org/10.1037/0882-7974.23.2.263

Murty, V. P., Ritchey, M., Adcock, R. A., & LaBar, K. S. (2010). fMRI studies of successful emotional memory encoding: A quantitative meta-analysis. *Neuropsychologia, 49*(4), 695-705. https://doi.org/10.1016/j.neuropsychologia.2010.07.030

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings.

Psychology of Learning and Motivation, 26, 125-173. https://doi.org/10.1016/S0079-7421(08)60053-5

*Nomi, J. S., Rhodes, M. G., & Cleary, A. M. (2013). Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition & Emotion, 27*(1), 141-149. https://doi.org/10.1080/02699931.2012.679917

Palombo, D. J., Te, A. A., Checknita, K. J., & Madan, C. R. (2021). Exploring the facets of emotional episodic memory: Remembering "What," "When," and "Which". *Psychological Science, 32*(7), 1104-1114. https://doi.org/10.1177/0956797621991548

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., Rushton, L., & Moreno, S. G. (2010). Assessing publication bias in meta-analyses in the presence of between-study heterogeneity. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 173*(3), 575-591. https://doi.org/10.1111/j.1467-985X.2009.00629.x

Price, J., McElroy, K., & Martin, N. J. (2016). The role of font size and font style in younger and older adults' predicted and actual recall performance. *Aging, Neuropsychology, and Cognition, 23*(3), 366-388. https://doi.org/10.1080/13825585.2015.1102194

Rhodes, M.G. (2016). Judgments of learning: Methods, data, and theory. In J. Dunlosky and S.K. Tauber (Eds.), *The Oxford handbook of metamemory* (pp. 65-80). Oxford University Press.

*Sanders, E. C., & Berry, J. M. (2021). Evidence for an age-related positivity effect in metacognitive judgments. *The Journals of Gerontology: Series B, 76*(7), 1282-1291. https://doi.org/10.1093/geronb/gbaa177

Saoud, W. (2020). Monitoring accuracy across domains of general knowledge and emotional face recognition. [Doctoral dissertation, York University]. https://yorkspace.library.yorku.ca/xmlui/handle/10315/37709

Schlochtermeier, L. H., Kuchinke, L., Pehrs, C., Urton, K., Kappelhoff, H., & Jacobs, A. M. (2013). Emotional picture and word processing: An fMRI study on effects of stimulus complexity. *PLoS One, 8*(2), e55619-e55619. https://doi.org/10.1371/journal.pone.0055619

Schmidt, S. R., & Saari, B. (2007). The emotional memory effect: Differential processing or item distinctiveness? *Memory & Cognition, 35*(8), 1905-1916. https://doi.org/10.3758/BF03192924

*Schmoeger, M., Deckert, M., Loos, E., & Willinger, U. (2020). How influenceable is our metamemory for pictorial material? The impact of framing and emotionality on metamemory judgments. *Cognition, 195*, 104112. https://doi.org/10.1016/j.cognition.2019.104112

Serra, M. J., Dunlosky, J., & Hertzog, C. (2008). Do older adults show less confidence in their monitoring of learning? *Experimental aging research, 34*(4), 379-391. https://doi.org/10.1080/03610730802271898

Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's z transformation be used? *Journal of Applied Psychology, 72*(1), 146-148. https://doi.org/10.1037/0021-9010.72.1.146

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*(1), 60-78.

https://doi.org/10.1002/jrsm.1095

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., ... & Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ, 343*(7818), 302-307. https://www.jstor.org/stable/23051949

Strain, A. C., Azevedo, R., & D'Mello, S. K. (2013). Using a false biofeedback methodology to explore relationships between learners' affect, metacognition, and performance. *Contemporary Educational Psychology, 38*(1), 22-39. https://doi.org/10.1016/j.cedpsych.2012.08.001

Talmi, D. (2013). Enhanced Emotional Memory: Cognitive and Neural Mechanisms. *Current Directions in Psychological Science, 22*(6), 430-436. https://doi.org/10.1177/0963721413498893

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition, 32*(5), 742-751. https://doi.org/10.3758/BF03195864

*Tauber, S. K., & Dunlosky, J. (2012). Can older adults accurately judge their learning of emotional information? *Psychology and Aging, 27*(4), 924-933 https://doi.org/10.1037/a0028447

*Tauber, S. K., Dunlosky, J., Urry, H. L., & Opitz, P. C. (2017). The effects of emotion on younger and older adults' monitoring of learning. *Aging, Neuropsychology, and Cognition, 24*(5), 555-574. https://doi.org/10.1080/13825585.2016.1227423

Tauber, S. K., & Witherby, A. E. (2019). Do judgments of learning modify older adults' actual learning? *Psychology and Aging, 34*(6), 836-847. https://doi.org/10.1037/pag0000376

Tauber, S. K., Witherby, A. E., & Dunlosky, J. (2019). Beliefs about memory decline in aging do not impact judgments of learning (JOLs): A challenge for belief-based explanations of JOLs. *Memory & Cognition, 47*(6), 1102-1119. https://doi.org/10.3758/s13421-019-00919-3

Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in medicine, 22*(13), 2113-2126. https://doi.org/10.1002/sim.1461

Undorf, M., & Bröder, A. (2020). Cue integration in metamemory judgements is strategic. *Quarterly Journal of Experimental Psychology, 73*(4), 629-642. https://doi.org/10.1177/1747021819882308

Undorf, M., Söllner, A., & Bröder, A. (2018). Simultaneous utilization of multiple cues in judgments of learning. *Memory & Cognition, 46*(4), 507-519. https://doi.org/10.3758/s13421-017-0780-6

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. https://doi.org/http://dx.doi.org/10.18637/jss.v036.i03

*West, J. T., & Mulligan, N. W. (2021). Investigating the replicability and boundary conditions of the mnemonic advantage for disgust. *Cognition and Emotion, 35*(4), 753-773. https://doi.org/10.1080/02699931.2020.1863187

*Witherby, A. E. (2019). *How do students regulate their learning of emotional and neutral* *information?* (Order No. 13898892). [Doctoral dissertation, Texas Christian University]. ProQuest Dissertations and Theses Global. https://www.proquest.com/docview/2299127839?pq-origsite=primo

*Witherby, A. E., & Tauber, S. K. (2018). Monitoring of learning for emotional faces: How do fine-grained categories of emotion influence participants' judgments of learning and beliefs about memory? *Cognition and Emotion, 32*(4), 860-866. https://doi.org/10.1080/02699931.2017.1360252

Witherby, A. E., Tauber, S. U. K., & Dunlosky, J. (2021). Why do emotional stimuli influence judgments of learning? Theory, evidence, and future directions. In *D. Moraitou & P. Metallidou (Eds.), Trends and Prospects in Metacognition Research Across the Life Span* (pp. 3-24). Springer. https://doi.org/10.1007/978-3-030-51673-4_1

* Witherby, A. E., Tauber, S. K., & Goodrich, M. (2022). People hold mood-congruent beliefs about memory but do not use these beliefs when monitoring their learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(4), 499. https://doi.org/10.1037/xlm0001096

Yang, C., Potts, R., & Shanks, D. R. (2017). Metacognitive unawareness of the errorful generation benefit and its effects on self-regulated learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 43*(7), 1073-1092. https://doi.org/10.1037/xlm0000363

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion, 24*(1), 3-47.

https://doi.org/10.1080/02699930903205698

Zacks, R. T., Hasher, L., & Li, K. Z. (2000). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 293–357). Mahwah, NJ: Erlbaum.

*Zimmerman, C. A., & Kelley, C. M. (2010). "I'll remember this!" Effects of emotionality on memory predictions versus memory performance. *Journal of Memory and Language, 62*(3), 240-253. https://doi.org/10.1016/j.jml.2009.11.004

Table 1[3]

Questions explored in the meta-analysis and the corresponding findings

Questions	Answers (research findings)
Q1. Does (and if so, to what extent) emotion affect JOLs?	Emotion produces a medium-sized (g = 0.53) effect on JOLs, but publication bias is evident.
Q2. Does (and if so, to what extent) emotion enhance memory?	Emotion has a small-to-medium sized (g = 0.38) effect on memory, but publication bias is evident.
Q3. Does emotion have different effects on JOLs and memory?	The emotional salience effect on JOLs is larger than the effect on memory, indicating that, although people can metacognitively appreciate the effect of emotionality on memory, they tend to overestimate the magnitude of this effect.
Q4. Do positive and negative emotion have different effects on JOLs?	Both positive (g = 0.53) and negative (g = 0.52) stimuli receive higher JOLs than neutral ones, and there is minimal difference between their effects on JOLs.

[3] Given the common problem that second order sampling error affects the precision of the meta-analytic estimates, especially when a small number of studies were included (Hunter & Schmidt, 2004), the results, especially for the moderating effects shown here, should be interpreted as tentative.

Q5. Do positive and negative emotion have different effects on memory?	Both positive (*g* = 0.34) and negative (*g* = 0.43) stimuli are remembered better than neutral ones, and there is no statistically detectable difference between their effects on memory.
Q6. Does arousal contribute to the emotional salience effect on JOLs?	There is minimal difference in the emotional salience effects on JOLs between arousal-matched (*g* = 0.48) and arousal-unmatched (*g* = 0.58) studies, suggesting that arousal may contribute minimally to the emotional salience effect on JOLs.
Q7. Does arousal affect memory?	The moderating role of arousal match in the emotional salience effect on memory is not statistically detectable, with unmatched-arousal (*g* = 0.42) and matched-arousal (*g* = 0.44) stimuli producing similar enhancing effects on memory.
Q8. Does emotionality affect JOLs to different extents for young and older adults?	Young adults exhibit a larger emotional salience effect on JOLs (*g* = 0.62) than older adults (*g* = 0.15), indicating that young adults' JOLs are more sensitive to emotion than those of older adults.
Q9. Does emotionality affect memory to different extents for young and older adults?	There is no statistically detectable difference in the emotional salience effects on memory between young (*g* = 0.33) and older adults (*g* = 0.63).

Q10. Does material type moderate the emotional salience effect on JOLs?	The emotional salience effect on JOLs is larger for images (g = 0.73) than for verbal materials (g = 0.38).
Q11. Does material type moderate the emotional salience effect on memory?	There is no statistically detectable difference in the emotional salience effects on memory between images (g = 0.27) and verbal materials (g = 0.47).
Q12. Does stimulus type moderate the emotional salience effect on JOLs?	Regardless of whether the study stimuli are single items (g = 0.53) or paired associates (g = 0.53), JOLs are always higher for emotional than for neutral items, and there is no statistically detectable difference in the emotional salience effect on JOLs between these two categories.
Q13. Does stimulus type moderate the emotional salience effect on memory?	There is no statistically detectable difference in the emotional salience effects on memory between single items (g = 0.41) and paired associates (g = 0.08).
Q14. Does test format moderate the emotional salience effect on JOLs?	Test format (g = 0.58 for cued recall; g = 0.43 for free recall; g = 0.71 for recognition) does not moderate the emotional salience effect on JOLs.
Q15. Does test format moderate the emotional salience effect on memory?	The emotional salience effect on memory is moderated by test format, with a larger effect in free recall (g = 0.59) than in recognition (g = 0.01) tests. There was no statistically detectable

	difference between cued recall (g = 0.21) and free recall tests, nor between cued recall and recognition tests.

Table 2

Moderator analysis results

	JOL effects (g = 0.53, p < .001)					Memory effects (g = 0.38, p < .001)				
Categorical moderators	k	g	95% CI	F	p	k	g	95% CI	F	p
Valence				0.18	.673				1.02	. 317
Positive emotion	37	0.53	[0.41, 0.66]		<.001	37	0.34	[0.19, 0.50]		<.001
Negative emotion	39	0.52	[0.39, 0.64]		<.001	39	0.43	[0.27, 0.58]		<.001
Arousal match				2.62	.079				1.83	.168
Matched	9	0.48	[0.26, 0.70]		<.001	9	0.44	[0.12, 0.75]		.008
Unmatched	70	0.58	[0.46, 0.70]		<.001	70	0.42	[0.28, 0.57]		<.001
Unknown	6	0.19	[-0.15, 0.53]		.267	6	0.02	[-0.39, 0.42]		.937
Age				13.73	<.001				3.22	.077
Young adults	62	0.62	[0.51, 0.73]		<.001	62	0.33	[0.19, 0.47]		<.001

Older adults	14	0.15	[-0.08, 0.37]		.204	14	0.63	[0.33, 0.93]		<.001
Material type				11.54	.001				2.26	.137
Verbal materials	43	0.38	[0.24, 0.51]		<.001	43	0.47	[0.30, 0.64]		<.001
Images	33	0.73	[0.57, 0.89]		<.001	33	0.27	[0.07, 0.47]		.008
Stimulus type				<0.001	.977				1.92	.170
Single items	70	0.53	[0.40, 0.65]		<.001	70	0.41	[0.27, 0.55]		<.001
Paired associates	6	0.53	[0.14, 0.93]		.009	6	0.08	[-0.36, 0.53]		.709
Test format				2.50	.089				12.76	<.001
Free recall	46	0.43	[0.29, 0.57]		<.001	46	0.59	[0.45, 0.73]		<.001
Cued recall	5	0.58	[0.14, 1.02]		.01	5	0.21	[-0.22, 0.64]		.340
Recognition	25	0.71	[0.50, 0.91]		<.001	25	0.01	[-0.18, 0.20]		.910

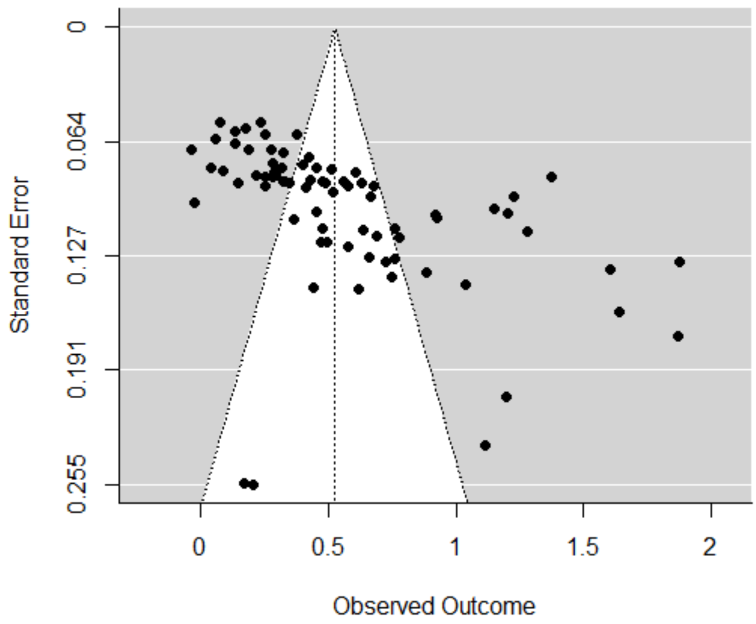

Fig.1 Funnel plot for effect of emotionality on JOLs. Each point represents a sample's composite effect size. The vertical line represents the summary effect size estimate.

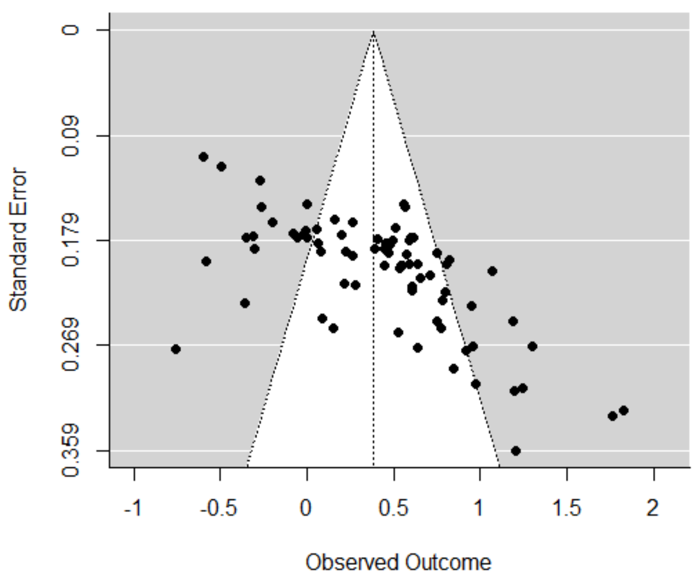

Fig. 2 Funnel plot for effect of emotionality on memory. Each point represents a sample's composite effect size. The vertical line represents the summary effect size estimate.