AMERICAN PSYCHOLOGICAL ASSOCIATION

Emotion

ISSN: 1528-3542

2023, Vol. 23, No. 6, 1740–1763
https://doi.org/10.1037/emo0001188

Examining the Effects of Emotional Valence and Arousal on Source Memory: A Meta-Analysis of Behavioral Evidence

Diana R. Pereira[1], Ana C. Teixeira-Santos[1, 2], Adriana Sampaio[1], and Ana P. Pinheiro[3]
[1] Psychological Neuroscience Laboratory, CIPsi, School of Psychology, University of Minho, Campus de Gualtar
[2] Department of Social Sciences, Institute for Research on Socio-Economic Inequality, University of Luxembourg, Belval Campus
[3] Faculdade de Psicologia, Universidade de Lisboa

The current meta-analysis examined the effects of valence and arousal on source memory accuracy, including the identification of variables that moderate the magnitude and direction of those effects. Fifty-three studies, comprising 85 individual experiments (N = 3,040 participants), were selected. Three separate analyses focusing on valence effects (valence-based: negative-neutral; positive-neutral; negative-positive) and other three focusing exclusively on arousal (arousal-based: high-low; medium-low; high-medium) were considered. Effect sizes varied from very small to medium. For the valence-based analyses, source memory accuracy was impaired for emotional compared with neutral stimuli ($d_{unb} = -.14$ for negative-neutral; $d_{unb} = -.11$ for positive-neutral), with a similar performance found for the negative-positive comparison ($d_{unb} = -.04$). In the case of arousal-based analyses, source memory was improved for stimuli with high and medium arousal versus low arousal ($d_{unb} = .27$, $d_{unb} = .49$, respectively), with no statistically significant difference between high and medium arousal stimuli ($d_{unb} = -.12$). Emotion effects on source memory were modulated by methodological factors. These factors may account for the variety findings typically found in emotion-related source memory research and could be systematically addressed in future studies.

Keywords: source memory, emotion, arousal, valence, meta-analysis

Supplemental materials: https://doi.org/10.1037/emo0001188.supp

A robust body of evidence demonstrates that emotions affect cognitive functions, including how we encode, store, and retrieve information in our daily life (e.g., Dolcos et al., 2017; Levine & Pizarro, 2004; Tyng et al., 2017). Specifically, memory for emotional stimuli appears to be improved compared with neutral ones, an effect that has been reported for both immediate and delayed retrieval conditions, spanning from a few minutes to years (D'Argembeau & Van der Linden, 2004; Davidson et al., 2006; Kensinger & Schacter, 2006b; Mickley Steinmetz & Kensinger, 2009; Waring & Kensinger, 2009; Weymar et al., 2011; Wirkner et al., 2018; Yick et al., 2015). This phenomenon, also known as the emotion-enhanced memory (EEM) effect, is not only characterized by enhanced overall accuracy, but also by an enhanced subjective experience of remembering details from the study episode (Ford & Kensinger, 2019; Ford et al., 2014; Kensinger & Corkin, 2003; Ochsner, 2000; Ritchey et al., 2019; Weymar et al., 2010). Nonetheless, it remains to be clarified whether this sense of subjective vividness for elements of the

This article was published Online First December 8, 2022.
Diana R. Pereira https://orcid.org/0000-0002-3190-9298
Ana C. Teixeira-Santos https://orcid.org/0000-0003-0262-9841
Adriana Sampaio https://orcid.org/0000-0001-7347-1282
Ana P. Pinheiro https://orcid.org/0000-0002-7981-3682
This work was supported by a PhD Fellowship (PD/BD/105964/2014), awarded to Diana R. Pereira, funded by the Portuguese Foundation for Science and Technology (FCT) through national funds and co-funded by the European Social Fund (ESF) through the Operational Program for Human Capital (POCH). It was also supported by a doctoral grant (SFRH/BD/80965/2011) awarded to Ana C. Teixeira-Santos, as well as a research grant (PTDC/MHC-PCN/0101/2014) awarded to Ana P. Pinheiro, both funded by FCT. The study was conducted at the Psychology Research Center (PSI/01662), University of Minho, supported by FCT and the Portuguese Ministry of Science, Technology and Higher Education (UID/PSI/01662/2019) through national funds (PIDDAC) and co-funded by FEDER through COMPETE2020 under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007653). We would like to acknowledge all authors that promptly replied when inquired about data. The authors declare no conflicting interests. The datasets and script generated and analyzed during the current study are available in OSF: https://doi.org/10.17605/OSF.IO/F8S5T.

Diana R. Pereira served as lead for conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing–original draft, and writing–review & editing. Ana C. Teixeira-Santos served in a supporting role for data curation, formal analysis, investigation, methodology, visualization, and writing–review & editing. Adriana Sampaio served as lead for resources and supervision and served in a supporting role for writing–review & editing. Ana P. Pinheiro served as lead for resources and supervision and served in a supporting role for writing–review & editing.

Correspondence concerning this article should be addressed to Diana R. Pereira, Psychological Neuroscience Laboratory, CIPsi, School of Psychology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal. Email: diana.r.pereira@gmail.com

study episode is also reflected in different facets of objective memory performance.

The study of emotion effects on source memory, that is the memory for the origins of an event (Johnson et al., 1993), may grant valuable contributions in this regard. Indeed, emotion may operate differently on source memory compared with the memory for the stimulus itself (i.e., item memory; Johnson et al., 1993; Jurica & Shimamura, 1999). The source monitoring framework, the main theoretical proposal in the field, asserts that the recollection of source-specifying information relies on decision-making processes that may include previous knowledge, stereotypes, and beliefs (e.g., "I think it was Mary who told me this, she always shows strong convictions about this matter."), as well as on the evaluation of different qualitative features of the event (Johnson et al., 1993; Lindsay & Johnson, 2000; Mather et al., 1999). Examples of these qualitative features are perceptual (e.g., stimulus color, identity of the speaker), spatial, and temporal characteristics, as well as thoughts, mental images, strategies, and feelings that occurred during encoding. The discrimination between these distinct qualitative features has been the focus of research concerning the influence of emotional information on source memory, which yielded mixed findings (see Pereira et al., 2019, for a selective overview). Even though previous meta-analyses have examined source memory, particularly the impact of age-related processes (Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995), the influence of emotion remains to be specified. Accordingly, the current study is a contribution to fill this gap by presenting an up-to-date systematic review of the effects of emotional stimuli on source memory accuracy.

In addition, theoretical, methodological, and applied implications may be associated with this review. From a theoretical standpoint and as it will be developed in subsequent sections, this review allows testing some of the predictions that can be derived from existing theoretical proposals and to clarify their potentialities and limitations given the inconsistent findings in this field. From a methodological perspective, it enables the identification of factors that might play a role in the source memory-emotion relationship and that can be considered in experimental designs, as well as in data analysis and interpretation. Lastly, although this review mainly includes laboratory studies with healthy adult participants, it can be used as a reference to devise ways of studying source memory and emotion in real word scenarios. Also, the findings may be tested and potentially extended to more applied research topics. One example is the eyewitness memory research as eyewitnesses are commonly solicited to provide detailed accounts of emotional events. It has been shown that these accounts are prone to distortions, with some details more likely to be remembered than others (e.g., Christianson & Loftus, 1991; Lindsay & Johnson, 2000; Wulff & Thomas, 2021). Another example relates to schizophrenia considering that a tendency to misattribute self-generated information to external sources has been consistently identified (e.g., Brookwell et al., 2013; Subramaniam et al., 2012), particularly when the information is negative (Pinheiro et al., 2016). Source memory impairments have also been reported in individuals with schizophrenia and associated with the presence of auditory hallucinations (Brébion et al., 2007). Training the identification of emotions and visual/auditory exercises have been shown to improve the discrimination between self- and external-generated sources in schizophrenia (Subramaniam et al., 2012). These examples illustrate that source memory and emotion research is also relevant to guide psychological interventions.

Emotion Effects on Source Memory: An Overview of Laboratory Experiments

In a typical source memory paradigm, participants are usually invited to learn a list of items while qualitative features of the stimulus are manipulated, such as stimulus color, color or shape surrounding the stimulus, location on the screen, list or block of appearance, as well as modality of presentation (e.g., visual vs. auditory), who generated the stimulus (e.g., male vs. female; participant-generated vs. experimenter-generated), or encoding task (see Appendix of Pereira et al., 2019, for more examples). After a learning period and a delay interval, memory for the qualitative features manipulated during encoding is tested. For example, imagine that during the study phase some items are presented in color A and others in color B (e.g., Dougal et al., 2007; Kensinger & Corkin, 2003; Wang, 2012; Wang & Fu, 2011; Weigl et al., 2020; Zhou et al., 2020). During the test phase, the experimenter can ask the participants if a given item was previously presented in color A or B, or if it is a novel item. Source memory can be tested directly as in the former example or after testing the memory for the item itself, such as in the case of free recall, cued-recall, old/new recognition, or remember/know judgments.

Pertaining to the manipulation of the emotional quality of a stimulus, the dimensional conceptualization of emotion has been widely adopted in experimental studies. It postulates that dimensions such as arousal (i.e., the degree of activation prompted by the stimulus, ranging from low to high) and valence (i.e., the degree of pleasantness of the stimulus, ranging unpleasant to pleasant) characterize our affective experience (Bradley & Lang, 1994; Lang et al., 1993; Russell, 1980). Based on valence and arousal properties, a list of negative, positive, and neutral stimuli with distinct levels of arousal might be selected to examine their influence on source memory performance.

So far, studies that examined the effects of emotion on source memory have focused on a variety of source features and provided mixed effects, showing that emotion may enhance, exert no effect, or even hinder source memory performance (see Appendix of Pereira et al., 2019). Specifically, considering memory for the perceptual qualities of a stimulus such as font color, some studies revealed that participants are more likely to remember color when the stimulus is emotional (D'Argembeau & Van der Linden, 2004, Experiment 3; Doerksen & Shimamura, 2001, Experiment 1; Kensinger & Corkin, 2003), whereas others failed to identify emotion effects on source memory (D'Argembeau & Van der Linden, 2004, Experiments 1 and 2; Davidson et al., 2006, Experiment 2; Wang & Fu, 2011; but see also MacKenzie et al., 2015, Experiment 2 for an impairment effect for negative stimuli). In contrast, when color is manipulated as the frame or background surrounding the stimulus, deleterious effects of emotion on source memory are more consistently observed (Boywitt, 2015, Experiment 1; MacKenzie et al., 2015, Experiment 1; Rimmele et al., 2011, Experiment 1).

Regarding memory for spatiotemporal features, such as stimulus location on the screen and the list/block of appearance during encoding, most studies showed an advantage of emotional over neutral information (e.g., D'Argembeau & Van der Linden, 2004, Experiment 4; D'Argembeau & Van der Linden, 2005; Rimmele et al., 2011, Experiments 2 and 3; Schmidt et al., 2011; Yick et al., 2015). Nonetheless, source memory impairments were observed in the context of negative stimuli (Maddock & Frein, 2009; Mitchell et al., 2006). Moreover, using pictures with distinct levels of arousal, Boywitt (2015) showed that stimuli with medium arousal ratings were associated with improved source memory for location, whereas stimuli with low or high arousal were associated with the worst source memory performance. These findings indicate that differences in stimulus arousal may account for the discrepant findings reported before.

The color and spatiotemporal features discussed so far require the discrimination of sources that can be deemed as external to the subject according to the source monitoring framework (Johnson et al., 1993). Nonetheless, some situations require discriminating between internal and external sources, such as when one has to distinguish if something was imagined or seen (i.e., reality monitoring; Johnson, 1988). Research probing emotion effects on reality monitoring is also characterized by mixed findings. For example, when discriminating between seen and imagined events, negative and high arousing stimuli were associated with improved performance (e.g., Kensinger & Schacter, 2006b), but impairment effects were also shown for negative and positive stimuli (e.g., Cook et al., 2007). More consistently, an emotion-related impairment effect was observed when participants had to distinguish whether a given stimulus was generated by themselves or by others (e.g., experimenter, another participant; Le Bigot et al., 2018; McKague et al., 2012).

Furthermore, other situations require discriminating between internal sources. In this context, memory for the encoding tasks used in the study phase has been commonly tested (e.g., self-referential vs. nonself-referential task), showing no effects of emotion (Ferré et al., 2019, Experiment 1; Kensinger & Schacter, 2006a; Sharot & Yonelinas, 2008) or deleterious effects (Cook et al., 2007; Ferré et al., 2019, Experiments 2 and 3; Mao et al., 2015), especially in the case of negative items (Newsome et al., 2012; Otani, Jaffa, et al., 2012; Otani, Libkuman, et al., 2012). Taken together, the existing studies show that emotion effects on source memory are modulated by the type of source monitoring task (external, internal, or reality monitoring; Johnson et al., 1993), as well as by the type of source features assessed in the memory test.

Theoretical Accounts of the Effects of Emotion on Source Memory

Several theoretical frameworks have offered possible explanations for the mixed results reported above. Examples are the Yerkes-Dodson law (Yerkes & Dodson, 1908), the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959), the priority-binding mechanism (Hadley & MacKay, 2006), the object-based binding theory (Mather, 2007), and the arousal-biased competition (ABC) theory (Mather & Sutherland, 2011).

The proposal of Yerkes and Dodson (1908) posits that increased arousal facilitates attention, memory, and other cognitive processes to a certain point. When such point is surpassed, the effects of arousal are no longer beneficial but rather detrimental. Accordingly, these authors postulated an inverted-U relationship between arousal and performance, the Yerkes-Dodson law: moderate levels of arousal are expected to lead to optimal performance, whereas lower or higher levels of arousal are likely associated with poorer performance. By testing source memory for the location of pictures with different levels of arousal, Boywitt (2015) provided a sound example of this inverted-U relationship.

The Easterbrook's cue-utilization hypothesis (Easterbrook, 1959) has been one of the most discussed approaches. It states that increased arousal results in the narrowing of attentional resources, which leads to prioritization of salient information whereas peripheral information tends to be disregarded. Consequently, there is a central or peripheral trade-off in memory, which is enhanced for the central details of an episode, but impaired for more secondary details. Several studies have supported this view by showing that central features are more accurately remembered in the context of emotional (vs. neutral) events, whereas peripheral features are more accurately remembered in the context of neutral (vs. emotional) events (e.g., Burke et al., 1992; Christianson, 1984, 1992; Christianson & Loftus, 1987, 1991; Kensinger, Garoff-Eaton, & Schacter, 2007). In this context, emotion effects on source memory seem to depend on whether a detail is central or peripheral (i.e., if it is under the focus of attention or not).

In a more recent proposal, the object-based binding theory (Mather, 2007), the distinction between central/peripheral details theorized by the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959) has a parallel with the distinction between intrinsic/extrinsic features. In this case, arousal grants a memory advantage for source features that are an integral/intrinsic part of the arousing item, that is, when they are perceived as a coherent and unified entity (e.g., stimulus color, location on the screen). On the contrary, when the source features are extrinsic to the arousing item, that is, when they can be perceived as distinct elements thereby requiring associations between them (e.g., background color or shape surrounding the stimulus; encoding task), arousal is more likely to exert a deleterious or no effect on source memory. Thus, arousal may enhance binding between an item and some but not all contextual features: the within-item binding is likely promoted whereas between-item binding is likely hindered or not affected in a particular way.

Regarding the intrinsic/extrinsic distinction, the ABC theory (Mather & Sutherland, 2011) adds the possibility that extrinsic features may also be enhanced if they are deemed as goal-relevant. According to this view, intrinsic and extrinsic features might be similarly relevant, and the stimulus priority is the one determining which source details are more likely to be remembered or forgotten. Specifically, as information competes for limited attentional resources, stimuli that are perceived as relevant to meet current goals and situational demands will likely be prioritized. Goals may range from idiosyncratic and situational goals to universal goals that are shared by many persons (see Levine & Edelstein, 2009). In this context, arousal might facilitate memory for source features that are tagged as high priority, while hampering the memory for features deemed as low priority. Remarkably, this account accommodates empirical evidence showing that emotion may benefit memory for extrinsic features (e.g., Doerksen & Shimamura, 2001, Experiment 2; Kensinger, O'Brien, et al., 2007; Kensinger & Schacter, 2006b) or even impair or exert no effect in the case of intrinsic features (e.g., Cook et al., 2007; Dougal et al., 2007; Ferré et al., 2019; Koenig & Mecklinger, 2008; Le Bigot et al., 2018; Minor & Herzmann, 2019; Wang & Fu, 2011).

Even though the ABC theory (Mather & Sutherland, 2011) asserts that source features associated with neutral or emotional stimuli with low arousal might also be prioritized during encoding and remembered better given one's current goals (Levine & Edelstein, 2009; Ochsner, 2000), most of the theoretical approaches emphasize arousal as the main driving factor of the emotion-related effects on memory. Notwithstanding, there are several instances showing processing differences between negative and positive information and memory is no exception (see Unkelbach et al., 2020 for an overview). Even though valence is commonly studied together with arousal in memory studies, its role on memory performance has also been acknowledged in the literature (e.g., Kensinger, 2004; Kensinger & Corkin, 2004; Levine & Edelstein, 2009; Mickley Steinmetz & Kensinger, 2009).

Furthermore, the neurocognitive mechanisms by which valence and arousal modulate memory are hypothesized to be different (Kensinger, 2004; Kensinger & Corkin, 2004): whereas arousal increases the likelihood that a stimulus is attended during encoding and stored in memory through consolidation processes, implicating the activation of the amygdalar-hippocampal network, valence is more likely to engage controlled encoding and elaborative processes associated with the activation of the hippocampus and prefrontal brain regions. Some authors have also pinpointed differences in the encoding of distinct valence categories. Accordingly, negative stimuli are thought to recruit more strongly occipito-temporal regions associated with sensory processes, whereas frontal regions, typically associated with elaborative processes, are more strongly activated by positive stimuli (Mickley Steinmetz & Kensinger, 2009). Thus, both valence- and arousal-related effects need to be considered when analyzing stimulus emotion effects on source memory.

Candidate Factors to Influence the Effects of Emotion on Source Memory

The complex relationship between emotion and source memory may be moderated by other factors accounting for the mixed results reported in the field. Prior meta-analyses on item (Galli et al., 2019; Shields et al., 2017) and source memory (Old & Naveh-Benjamin, 2008; Spencer & Raz, 1995), and on the effects of emotion on other memory processes (Hostler et al., 2018; Lipinska et al., 2019; Murphy & Isaacowitz, 2008) have suggested that both stimulus-related factors (e.g., type of study material: words vs. pictures) and task-related factors (e.g., encoding intentionality; delay interval between study and test; type of source memory task according to the source monitoring framework) might contribute to a better understanding of the interactions between memory and emotion. In this context, the current study intends to examine the influence of the following factors:

1. Type of source memory feature. This intends to appraise the distinction between intrinsic/central and extrinsic/peripheral features postulated by theoretical frameworks such as the object-based binding theory (Mather, 2007) and the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959). This distinction also agrees with the previous empirical evidence showing that the relation between emotion and source memory differs as a function of the source features manipulated during encoding (e.g., Boywitt, 2015; D'Argembeau & Van der Linden, 2004).

2. Type of source memory task. The other theoretical notion that can be evaluated is whether emotion effects may also vary as a function of the source monitoring task. Indeed, according to the source monitoring framework (Johnson et al., 1993), the type of source monitoring task (external, internal, or reality-monitoring) may affect memory performance. While external source memory decisions are expected to rely on perceptual, spatio-temporal, affective, and semantic features, internal source memory decisions are expected to rely more on cognitive operations (e.g., thoughts; mental images; strategies) that took place during the encoding episode (Ferguson et al., 1992; Johnson, 1988; Johnson & Raye, 1981; Raye & Johnson, 1980). Reality-monitoring tasks that require internal-external source features discrimination are hypothesized to be easier compared with internal and external source memory tasks, whose source-specifying features are more likely to share similar characteristics (Johnson & Foley, 1984; Raye & Johnson, 1980).

3. Stimulus type (words vs. pictures). Previous literature documented that pictures tend to communicate more effectively the emotional significance of an event compared with words, being more likely to be remembered (e.g., Paivio et al., 1968) and to elicit arousal-related changes in the beholder (Citron, 2012; Herbert et al., 2006; Hinojosa et al., 2009; Mathews et al., 2013).

4. Encoding intentionality (i.e., if the participants are instructed to memorize or not the information for a subsequent memory test). In this regard, D'Argembeau and Van der Linden (2004) ran several experiments to test whether the intentional or incidental encoding of source features affected the memory for color and location of emotional and neutral words. They found that memory for stimulus color was improved for emotional compared with neutral words in the incidental learning condition only, whereas memory for stimulus location was also improved for emotional relative to neutral words irrespective of encoding instruction. This finding suggests that encoding intentionality modulates emotion effects on source memory as a function of the source features being tested. Furthermore, this same study revealed that the difference between emotional and neutral words is weakened by intentional learning conditions as participants are more likely to engage in effortful encoding. This could be reflected in similar allocation of attentional resources to both emotional and neutral stimuli or in the implementation of specific encoding strategies that strengthen the memory trace for both types of stimuli.

5. Delay time between study and test. It has been observed that the EEM effect is more noticeable and enhanced with the passage of time, especially when testing item memory (Mitchell et al., 2006; Schaefer et al., 2011; Sharot & Phelps, 2004; Sharot & Yonelinas, 2008; Yick et al., 2015; Yonelinas & Ritchey, 2015). These findings support the notion that emotional information is more likely to be consolidated into a more stable memory trace (Hamann, 2001; Kensinger, 2004). Specifically, Wang and Fu (2011) tested the effects of emotion on source

> memory for font color using eight different delay intervals (from immediate to 2 weeks of delay). However, emotion failed to modulate source memory and, more importantly, this lack of effect was stable across delay intervals (see also Sharot & Yonelinas, 2008). This finding suggests that the role of time may not be the same for item and source memory.

Finally, considering that differences in arousal might be associated with different outcomes in source memory performance (e.g., Boywitt, 2015; Easterbrook, 1959; Yerkes & Dodson, 1908) and that stimulus arousal and valence may influence source memory through different cognitive and neural mechanisms (Kensinger, 2004; Kensinger & Corkin, 2004), valence and arousal characteristics of the study items were addressed in this study in two specific ways. First, we selected studies that manipulated stimulus valence by including negative, positive and/or neutral items, as well as studies focusing on stimulus arousal by including items with low, medium, or high arousal levels. Second, as previously suggested by Hostler and colleagues (2018) and depending on the emotion manipulation adopted in each experiment, we coded the experiments based on whether they controlled stimulus arousal across valence categories or whether they controlled stimulus valence across distinct levels of arousal. These codifications are useful as valence and arousal are commonly studied together due to the observation that negative and positive items are often more arousing than neutral ones (Lang et al., 1998), which poses some barriers to the investigation of their differential effects on memory.

The Current Meta-Analysis

From the evidence presented so far, we observed that the relationship between emotion and source memory encloses mixed findings and remains to be specified. Methodological heterogeneity may have contributed to these different results. In this regard, conducting a systematic review may help to identify factors of interest. Thus, the main goal of the current study was to provide an up-to-date quantitative systematic analysis of behavioral studies probing the influence of emotion on source memory accuracy. Furthermore, we aimed to specify the contribution of different stimulus- and task-related factors, as described above, which could be related to variations in the relationship between emotion and source memory. This approach also granted the opportunity to examine some of the predictions postulated by the theoretical approaches commonly considered in this field of research.

Method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al., 2009) guidelines were adopted in the current meta-analytic review.

Literature Search Strategy

An initial database search was conducted on Web of Science, Scopus, Pubmed, Psychology and Behavioral Sciences Collection (through EBSCOhost), and PsycINFO from inception to July 2021. The keywords are provided in Supplemental Materials Table S1. We also searched for relevant citations in previous reviews discussing the role of emotion in source memory (Bowen et al., 2018; Kensinger, 2007, 2009; Kensinger & Kark, 2018; Kensinger & Schacter, 2016; Mather, 2007; Mather & Sutherland, 2011) and in the list of references of articles subjected to full-text reading.

Eligibility Criteria

During the abstract and full-text screening phase, several eligibility criteria guided the selection of studies. Specifically, full-length articles reporting empirical data and published in peer-reviewed journals were considered, excluding other works such as conference abstracts/proceedings, editorials, book chapters, dissertations, narrative/qualitative/quantitative reviews, and patents. Even though we considered articles written in English, French, Portuguese, or Spanish, only articles written in English were included. The studies had at least one sample of healthy adults. Nonhuman animal studies and studies reporting data exclusively of patients, children, adolescents, and older adults were not considered. This decision was supported by evidence demonstrating that emotional processing and episodic memory are significantly modulated by age (e.g., childhood and adolescence: Ghetti & Bunge, 2012; Ladouceur, 2012; older adults: Davidson et al., 2006; Nashiro & Mather, 2010, 2011; Newsome et al., 2012; Spencer & Raz, 1995; Tyng et al., 2017) and by clinical conditions (e.g., depression: Urban et al., 2018; fibromyalgia: Robin et al., 2018; schizophrenia: Fairfield et al., 2016). By restraining the sample to healthy young and middle-aged adults, we expected to cover most of the studies focused on emotion and source memory, while reducing the sources of variability in data analysis. In the context of source memory, few studies with affective manipulations have been conducted with children, adolescents, older adults, and clinical groups. Furthermore, studies examining pharmacological (including the intake of substances such as alcohol) and nonpharmacological interventions (e.g., noninvasive brain stimulation; cognitive rehabilitation and cognitive training strategies; sleep-related manipulations) were excluded when we were unable to extract data from a control or baseline condition with no intervention.

Articles were included if the experimental procedure involved a source memory task together with a manipulation of the emotional properties of the study stimuli, namely valence and/or arousal properties. An effort was made to select studies that used similar experimental procedures in terms of source memory task and types of stimuli. The intention was to mitigate some of the sources of methodological heterogeneity. For this reason, studies including short-term memory (STM) or working memory paradigms in which participants are required to retain short lists of items for just a few seconds with immediate recall or recognition were not considered (e.g., Borg et al., 2011; Mitchell et al., 2006), which is also in agreement with the eligibility criteria adopted in prior meta-analytic reviews on episodic memory (e.g., Galli et al., 2019). In a similar fashion, studies testing other type of emotional manipulations such as mood induction and emotional encoding contexts (e.g., presentation of task-irrelevant emotional stimuli during the encoding of target stimuli) were not considered. Also, studies using specific stimulus categories such as taboo and erotic materials (e.g., Hadley & MacKay, 2006), spiders or snakes, were not included (e.g., Bell et al., 2017). The reasoning behind is that most of the source memory studies tend to mix various stimulus categories, which attenuates category-specific effects as it may be the case of taboo (Madan et al., 2017) and erotic materials (Laier et al., 2013). Furthermore, only studies that provided the necessary information (e.g., *M*, *SD*, *SE*, and *N*) to compute effect sizes in the published version or by request to the authors were eligible.

Study Selection

Figure 1 presents a schematic illustration of the search strategy adopted in the current study. After removal of duplicates, two independent authors, Diana R. Pereira and Ana C. Teixeira-Santos, screened each record by title and abstract following the eligibility criteria. The disagreements were discussed between the two authors until a consensus was reached. If any doubts persisted, records were included in the full-text reading phase. Then the full-text screening phase was ensued. The interrater reliability obtained at this stage as measured by the Cohen's κ was .61, which was indicative of a moderate level of agreement (McHugh, 2012). The reasons for exclusion can be consulted in Figure 1.

Coding Procedure (Main and Moderator Variables)

The main outcome analyzed in the current meta-analysis was the performance in the source memory task reflected in response accuracy. When studies included data for both correct source responses and errors, priority was given to correct responses. Of note, when errors were the only outcome available, means were multiplied by "−1", following previous studies (e.g., Sala et al., 2019). Means, standard deviations/standard errors of the mean, and sample sizes were extracted for each experiment and emotion-related condition. In the case of figures, we used the web-based tool WebPlotDigitizer (https://apps.automeris.io/wpd/; Rohatgi, 2021) to extract the numerical data as this is proven to be a reliable approach (Lipinska et al., 2019). If the sources available in the articles were insufficient to obtain the data, authors were contacted and asked to provide the missing information (means and standard deviations).

Information about the selected moderator variables was also considered (see Table 1 for an overview). Additionally, other relevant methodological details were extracted, namely: sensory modality of presentation (visual, auditory, or both); total number of study items; if the emotional manipulation was achieved between- or within-participants; average years of formal education of the participants; source of the study materials; total number of stimuli used in the experiment; number of stimuli per emotional condition; mean valence and/or arousal ratings of the stimuli considering each emotional condition; list composition (mixed or pure); stimulus encoding time (in seconds); instruction(s) provided during encoding; number of study-test cycles; retention interval between the study and the test phase (in minutes); time given to the participants during the test phase; total number of new items included in the test phase; type of source memory test (free recall, cued recall, or recognition); the memory test focused on source memory features or evaluated first item memory (e.g., old/new recognition, remember/know judgments) and then source memory (direct or indirect test, respectively); number of source memory features tested; how the source memory accuracy was computed. All the information was coded by the first author (see Supplemental Materials Tables S5–S8).

Figure 1

Flow Diagram of the Search Strategy According to the PRISMA Guidelines (Moher et al., 2009)

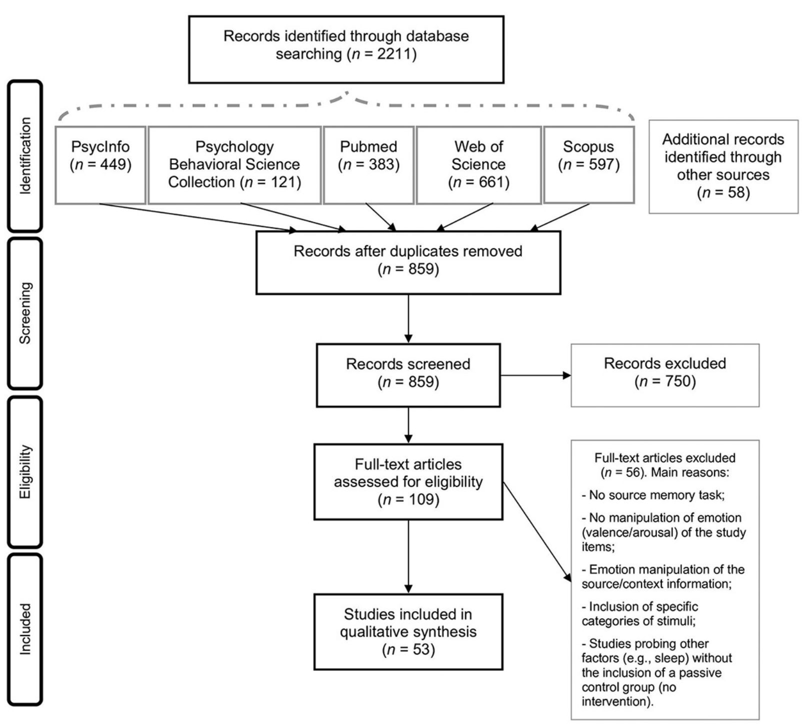


Table 1
Summary of the Moderator Variables Considered in This Study

Moderators	Description
Sociodemographic factors	Total number of women in the final sample.
	M_{age} of the participants.
Stimulus-related factors	Type of study material: Words, pictures, or both.
	The stimuli were controlled for arousal or valence categories? Yes or no.
Task-related factors	Type of SM task (Johnson et al., 1993): External, internal, and reality monitoring.
	Type of SM feature manipulated during the encoding phase and evaluated during the test phase: Temporal, spatial (location), stimulus color, background (color, shape, and scene), identity of whom generated the stimuli (e.g., self-generated, experimenter-generated, partner-generated, and sex of the speaker), modality of stimulus presentation (e.g., seen, heard, and imagined), encoding task (e.g., remember or forgot instruction, common or indoor question, and scene or people judgment).
	Type of SM feature according to the object-based binding theory (Mather, 2007): Extrinsic, intrinsic.
	Participants were instructed to overtly memorize the stimuli and/or they knew their memory would be tested at some point in the future: intentional, incidental.
	Retention interval between the study and the test phase (in minutes).

Note. SM = source memory.

Quality Assessment

To assess the methodological quality of the studies included in the review, a list of 12 items based on the Cochrane Collaboration's tool for assessing risk of bias (Higgins & Altman, 2008) and on the study of Thompson et al. (2017) was adapted. The items covered the following topics: clear identification of goals and outcomes; description of sociodemographic and health-related information regarding the sample; reference to the eligibility criteria used in the study; information about the participants that did complete the study; issues related to the randomization of participants and/or the (pseudo) randomization and counterbalance of the stimuli; information regarding how the emotion and the source memory features were manipulated in the study, and if a control/comparison condition was included; reference if the participants were overtly informed about the emotional manipulation of the stimuli; nonselective reporting of results for the identified outcome(s); other sources of bias (e.g., incomplete information regarding the stimulus selection and properties). The complete list of items is presented in more detail in Supplemental Materials Table S9. Each item was rated from zero to two ("0" = not verified; "1" = verified; "2" = unclear, i.e., not enough information to evaluate the item; thus, "1" was the desired score for each item). To obtain a qualitative index for each experiment, the total number of items scored as "1" was divided by the total number of items, and then categorized following Taneja et al. (2021): low (≤ 33%); moderate (≥ 34% and ≤ 66%); high (≥ 67%).

Meta-Analytic Procedure

Because some studies probed valence and arousal effects whereas others focused solely on arousal effects, six separate comparisons were conducted: three comparing valence categories (negative vs. neutral; positive vs. neutral; negative vs. positive) and three comparing different levels of arousal (high vs. low; medium vs. low; high vs. medium). The manipulation of emotion in all studies included in the current meta-analysis followed a within-subjects design. Accordingly, the standardized mean difference was computed based on the d unbiased (d_{unb}) estimate (Cumming, 2012) and used as a measure of effect size. This measure is a correction to the Cohen's d proposed by Hedges (1981) that attenuates the overestimation bias in studies with small sample sizes. Also, its interpretation is similar to the Cohen's d. Considering recent discussions about the interpretation of effect size in psychological research (see Funder & Ozer, 2019), we adopted the following guidelines: .1 was regarded as very small, .2 as small, .4 as medium, and .6 as large effect. The sample variance of the effect size was computed according to Borenstein et al. (2009; see Supplemental Materials Table S2). However, as the formula requires a correlation coefficient between experimental conditions, which is seldom reported in studies (Dunlap et al., 1996), a correlation of .5 was assumed by default following prior work (e.g., Galli et al., 2019). Nonetheless, a sensitivity analysis was conducted with the values .3, .5, and .7, following the recommendation of Borenstein and colleagues (2009). All formulas and results of the sensitivity analysis are presented in Supplemental Materials Tables S2 and S4, respectively.

Furthermore, because response accuracy data could be extracted for the same sample of participants from distinct experimental conditions, this could represent a problem to conventional meta-analytic procedures by assuming that effect sizes are independent from each other. One way to mitigate this problem is to average effect sizes and sample variances across conditions to obtain one estimate per sample (e.g., Murphy & Isaacowitz, 2008). Alternatively, we can select only one estimate per study (see Assink & Wibbelink, 2016, for other examples). However, such approaches may also lead to loss of information and may bias the estimation of true effect sizes (Borenstein et al., 2009; Del Re, 2015). As an alternative, multilevel models account for dependency among multiple observations extracted from a common sample of participants. Therefore, this approach was adopted, in accordance with previous meta-analyses (e.g., Galli et al., 2019; Schweizer et al., 2019; Teixeira-Santos et al., 2019). It is worth noting that a positive effect of negative or positive valence relative to neutral valence was taken as a boosting effect of emotion on source memory, while a negative effect was interpreted as a deleterious effect of emotion on source memory. A similar train of thought was applied to the comparison between negative and positive valence, as well as to the comparisons across levels of arousal.

The analyses reported here were conducted with the *metafor* package (Version 2.1-0; Viechtbauer, 2010) in RStudio (Version 1.3.1073 "Giant Goldenrod" for macOS; RStudio Team, 2020). Databases and script are available at Open Science Framework (OSF; Pereira et al., 2022). First, an influential outcomes analysis was conducted with the influence function of *metafor* to identify outliers that could exert a strong impact on the results (Viechtbauer & Cheung, 2010). Then, outliers were removed from the initial database following the procedure described in previous studies (e.g., Teixeira-Santos et al., 2019). After this analysis, eight influential outcomes were excluded as they showed effects sizes that diverged from the pool of effect sizes obtained from other studies within the same meta-analysis (see Supplemental Materials Table S3). In other words, these outcomes were identified as influential as they might change the overall effects and, consequently, influence the validity and robustness of the meta-analytic findings (Viechtbauer & Cheung, 2010).

Second, the influence of publication bias was assessed using the Egger's test (Egger et al., 1997) and the rank correlation test (Begg & Mazumdar, 1994) for funnel plot asymmetry, the trim-and-fill method (Duval & Tweedie, 2000), and the Henmi and Copas method (Henmi & Copas, 2010). Nonetheless, these analyses were not performed for two of the comparisons reported here (i.e., medium vs. low; high vs. medium), as tests based on funnel plot asymmetry are not recommended when the number of studies is small (less than 10) and the heterogeneity is large (Ioannidis & Trikalinos, 2007; Sterne et al., 2011). Of note, both the influential outcomes and the publication bias analyses were based on a standard meta-analytic method. Specifically, the *rma* function was used.

Finally, third-level meta-analyses (study sample as the cluster level) were computed using the function *rma.mv* from *metafor*. Of note, a random effect method was selected based on the assumption that the parameters to be estimated are likely to vary from study to study in the population, and that the consequences of applying this method to fixed effect data are more lenient than the other way around (Field & Gillett, 2010; Pigott & Polanin, 2020; Quintana, 2015). Furthermore, moderation analyses were conducted to examine the role of various methodological variables (see Table 1) while keeping the same multilevel approach, in which moderators were included one at a time. The moderation analyses can be particularly informative when the statistical tests of heterogeneity (e.g., Q test) reveal that the included studies display divergent findings, indicating that the effect sizes may be inconsistent among studies (Borenstein et al., 2009; Viechtbauer, 2010). Relatedly, when reporting the moderation analyses, two Q-statistics (Q_E and Q_M) were included in the summary of the results. The first results from the test for residual heterogeneity when moderators are included in the analysis. For example, if a nonsignificant p-value is found, it implies that the moderator accounts for most of the heterogeneity between the studies, even though this is unlikely to occur when testing only one moderator (Del Re, 2015). The second is the test statistic regarding the omnibus test of moderator. A nonsignificant p-value in this case suggests that the overall effect is not moderated by the variable tested in the analysis (Assink & Wibbelink, 2016).

When mixed-effects models were statistically significant and in the specific case of categorical variables, subgroup analyses were also computed to supplement the results. In addition, the subgroup analyses were used in some instances to specify the role of factors with theoretical relevance, such as the type of source memory features according to the object-based binding theory (Mather, 2007).

It is worth noting that conducting multiple moderators and subgroup analyses raise the issue of multiple comparisons and augmented occurrence of Type I error. The Bonferroni correction has been widely used to circumvent this issue. However, with respect to systematic reviews, there are no straightforward solutions to this problem so far. In this regard, one way to address this issue is to distinguish between exploratory and planned subgroup analyses (e.g., Pigott & Polanin, 2020; Whitfield et al., 2022). Accordingly, the exploratory analyses in this study included the sociodemographic factors (see Table 1) as well as the subgroup analyses that were conducted despite no statistically significant mixed-effects models. Even so, the results were only reported if these subgroup analyses reached the threshold of statistical significance.

In addition, due to the small number of experiments (less than 10) included in two of the three arousal-based analyses (i.e., medium vs. low; high vs. medium), moderation analyses were not reported for those cases (e.g., Galli et al., 2019; Mancuso et al., 2016). Also, moderators that presented ranges of values were replaced by the average of the minimum and maximum value. This strategy was adopted in some instances such as the delay interval between study and test or the age of the participants to retain as much information as possible and to attenuate naturally occurring imbalances across moderator levels, which can be problematic (see Field & Gillett, 2010).

Results

The main results of the search strategy are shown in Figure 1. From a total of 2,269 records identified through database search and additional sources, 53 studies published between 1997 and 2021 were included in the meta-analysis. These 53 studies comprised a total of 85 individual experiments: 71 probing valence and arousal effects, and 14 examining arousal effects only. In the following subsections, we start by reporting the main characteristics and results of the quality assessment considering the 53 studies. After this overview, results regarding the publication and reporting bias as well as the main findings are described separately for each comparison.

Characteristics of the Studies Included

A total of 3,040 participants (14–186 participants per sample; $M = 28.55$, $SD = 13.50$ for valence-based studies; $M = 49.93$, $SD = 46.73$ for arousal-based studies; 2,025 women) with M_{age} of 22.45 years and 14.08 years of formal education were included in the pool of studies. Of the 85 experiments, 47 used an external source memory task, 23 used an internal source memory task, whereas 15 used a reality monitoring task. These experiments presented on average 117 items during the encoding phase and gave participants 3.83 s to study each item, even though some of the experiments ($n = 11$) implemented a self-paced approach. Most of the experiments ($n = 76$) presented visual stimuli, four used the auditory modality, and five used both visual and auditory modalities. Also, 39 experiments used pictures as stimuli, 38 used words, seven combined both types of stimuli, and one experiment used faces. Only 16 experiments controlled the arousal/valence ratings of the stimuli across all valence/arousal categories under study, respectively.

Emotional and neutral stimuli were intermixed during study and test phases (i.e., mixed list composition) in most of the experiments, and only two experiments presented stimuli in separate lists per valence (i.e., blocked design). In 48 experiments, participants were not aware that their memory would be tested (i.e., incidental encoding), whereas in 35 experiments participants were overtly instructed to memorize the stimuli (i.e., intentional encoding). Nonetheless, in two of the experiments, it was not possible to determine the encoding intentionality. The source memory feature most often examined was the encoding task (n = 23), followed by spatial location (n = 20), modality of presentation (n = 11), temporal features (n = 9), stimulus color (n = 11), stimulus background (n = 8), and identity of who presented or generated the encoding stimulus (n = 6). The number of source memory features manipulated during encoding and testing ranged between 2 and 56. Nonetheless, testing two features was the most frequent approach (n = 55). Moreover, whereas 41 experiments manipulated extrinsic features, 44 manipulated intrinsic features. Most of the experiments included only one study-test cycle, with only 10 including more than one cycle. The delay interval between study and test ranged between immediate to two weeks; participants' memory was commonly assessed by using recognition tests (n = 77; 7 free recall; 1 cued recall). The memory test had 71 new items on average. Finally, source memory was directly tested in 44 experiments, yet indirectly tested in 41 experiments.

Quality Assessment

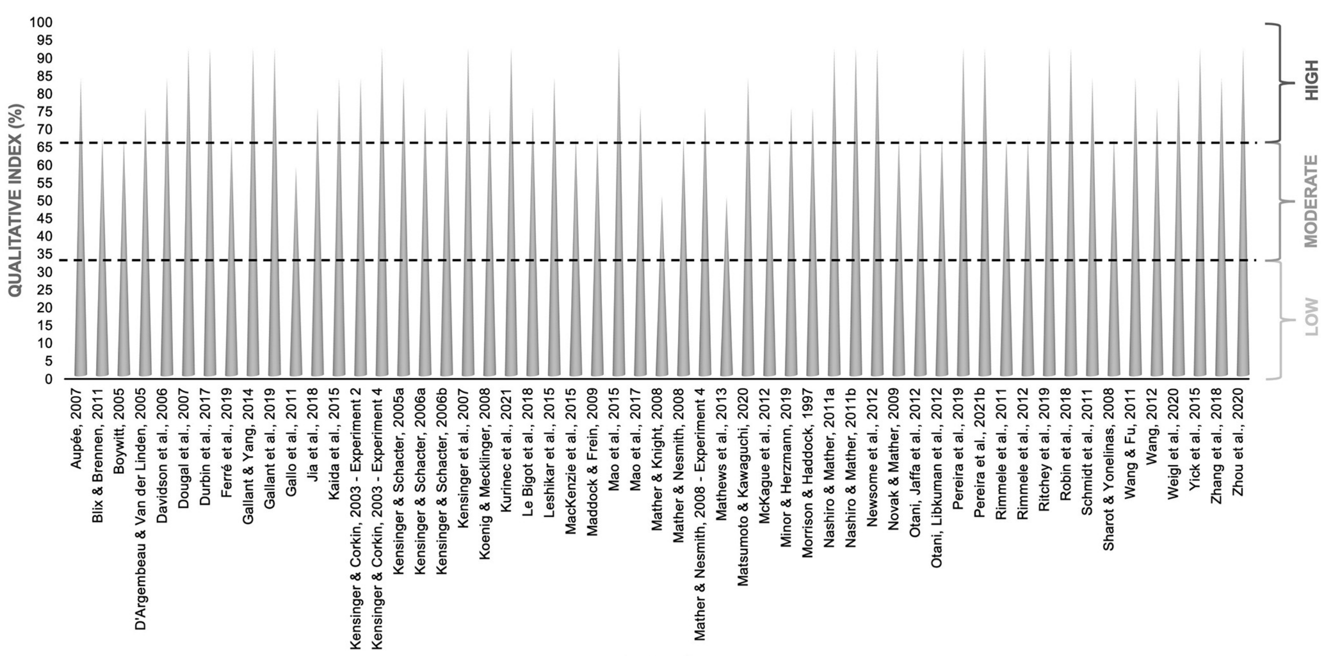
The scores attributed to each item and the final scores are presented for each experiment in Supplemental Materials Table S9. The total scores and the qualitative index are also depicted in Figure 2. Following the qualitative assessment described in the Method section (Taneja et al., 2021), the methodological quality ranged between 50% and 91.67% (M = 77.22%, SD = 11.91): most experiments (n = 51) were characterized by high methodological quality, whereas 28 were characterized by moderate methodological quality.

Publication Bias

The funnel plots for each meta-analysis are displayed in Figure 3. For two of the four comparisons (positive vs. neutral; negative vs. positive), the assessment of publication bias indicated no major bias (positive vs. neutral: z = .46, p = .646; Kendall's τ = .07, p = .405; negative vs. positive: z = −1.49, p = .137; Kendall's τ = −.07, p = .380). No missing studies were imputed by the trim-and-fill method in the case of negative versus positive, yet 10 missing studies were imputed to the left side in the case of positive versus neutral (see Supplemental Materials Figure S1). In these two cases, the estimations derived from the Henmi and Copas method were similar to the estimations provided by the standard random effects meta-analyses. However, regarding the negative versus neutral comparison, the tests yielded evidence of publication bias (z = 2.02, p = .044; Kendall's τ = .16, p = .021). In fact, when using the trim-and-fill method as a correction for publication bias, a total of 10 missing studies were imputed to the left side of the funnel plot (see Supplemental Materials Figure S1).

Pertaining to the high versus low arousal comparison, evidence of publication bias was also found (z = 2.43, p = .015; Kendall's τ = .28, p = .091). However, the trim-and-fill method did not impute any missing studies to either side of the funnel plot (see Supplemental

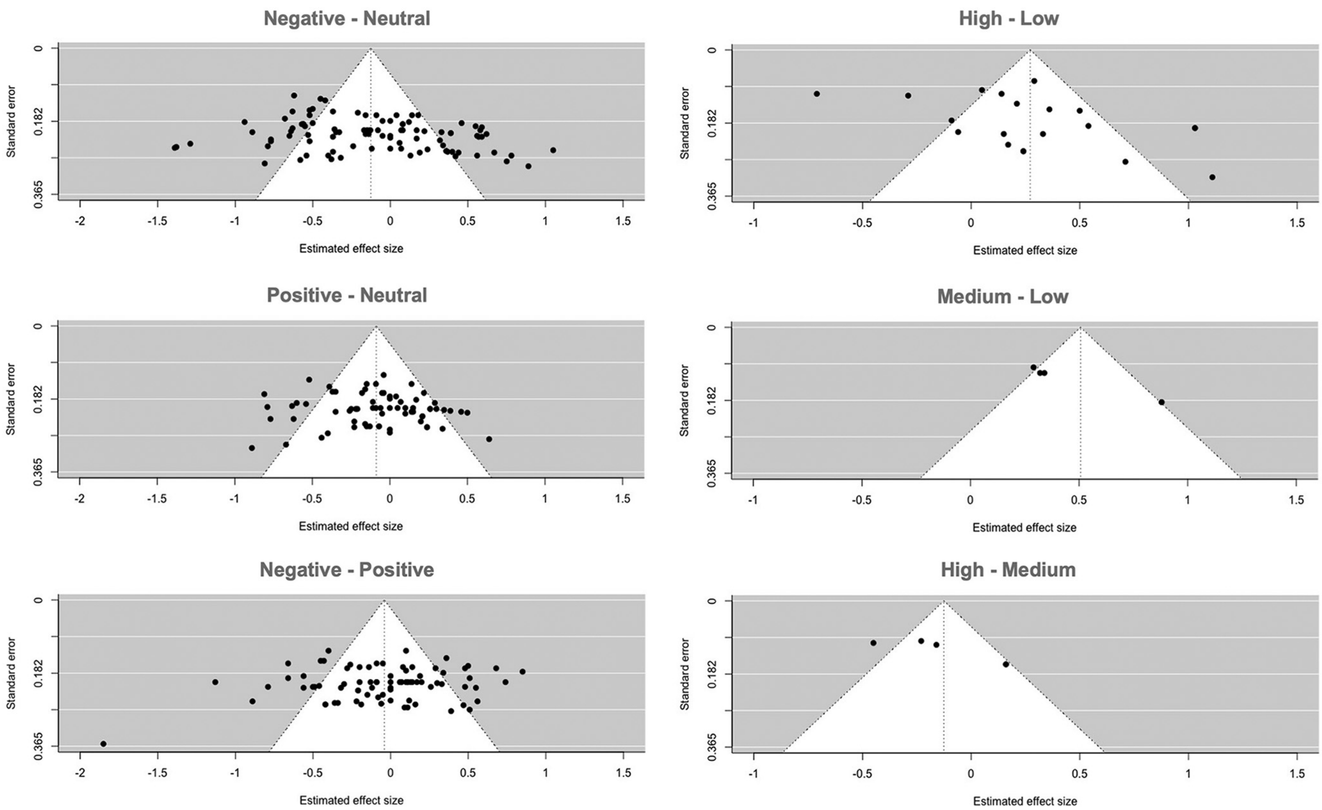
Figure 2
Total Score Obtained by Each Study in the Quality Assessment



Note. For the experiments from the same study whose total scores differed, the qualitative index is plotted separately.

Figure 3

Funnel Plot Graphs for Each Analysis (Valence and Arousal Analyses on the Left Side, and Arousal Analyses on the Right Side)



Materials Figure S1), indicating that the meta-analytic results would remain unaltered.

Because heterogeneity may influence the results of publication bias (Ioannidis & Trikalinos, 2007), caution is warranted regarding the meta-analytical results and their interpretations, especially in the case of the negative versus neutral, high versus medium, and medium versus low comparisons (these last two for containing few studies).

Main Findings on Source Memory Accuracy

An overview of the main findings for each meta-analysis is provided in Figure 4. The forest plots generated during the meta-analytic analyses are shown in Supplemental Materials Figure S2. For each valence- and arousal-based comparison, we first describe the results of the main analyses followed by the moderator and subgroup analyses. The moderator and subgroup analyses are further divided into sociodemographic, stimulus-related, and task-related factors (see Table 1) where applicable.

Valence-Based Effects

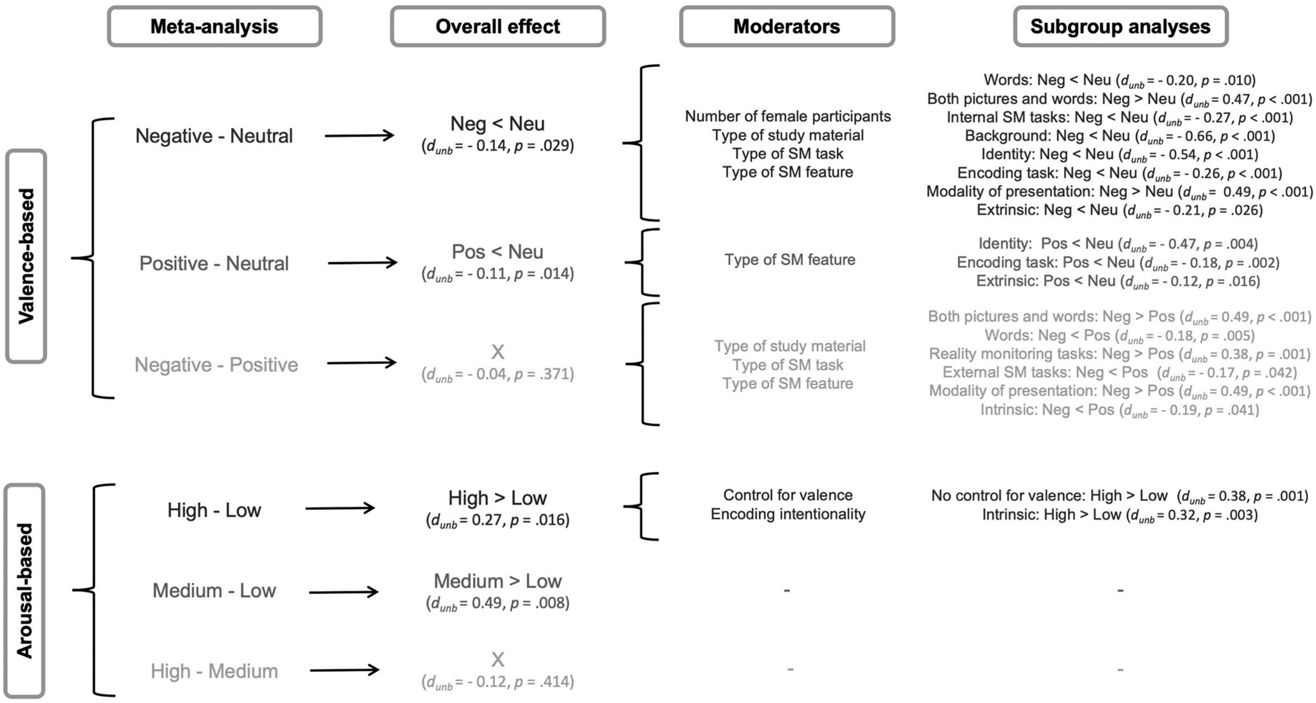
Negative Versus Neutral Valence. There was a statistically significant difference between negative and neutral stimuli considering their influence on source memory performance, $d_{unb} = -.14$, $SE = .07$, $z = -2.18$, $p = .029$, 95% confidence interval, CI [−.27, −.01] (see Supplemental Materials Figure S2A and Table S4). In addition, the heterogeneity Q test was statistically significant, $Q(96) = 521.32$, $p < .001$, indicating that the true effect sizes vary from study to study (Borenstein et al., 2009; Viechtbauer, 2010). Such heterogeneity might be partially explained by study-specific factors. Thereby, moderator and subgroup analyses may be useful in this regard (see Table 2) and are presented in the next subsections.

Moderator and Subgroup Analyses.

Stimulus-Related Factors. The type of material used in the studies ($Q_M(2) = 8.90$, $p = .012$) emerged as a statistically significant moderator. Particularly, the mean effect for pictures (−.21) and for words (−.20) was lower than the mean effect of including both types of stimuli (.46; $z = -2.90$, $p = .004$; $z = -2.85$, $p = .004$, respectively). Subgroup analyses showed that whereas the effect of negative relative to neutral information on source memory was nonsignificant for pictures ($z = -1.90$, $p = .057$), the same was not verified for words ($z = -2.58$, $p = .010$) and studies using both types of stimuli ($z = 5.25$, $p < .001$). In the case of words, the source memory performance was worst for negative in comparison with neutral words ($d_{unb} = -.20$, $SE = .08$). In contrast, the source memory performance was improved for negative relative to neutral information considering studies that included both types of stimuli ($d_{unb} = .47$, $SE = .09$).

Even though the control of stimulus arousal did not emerge as a statistically significant moderator ($Q_M(1) = 2.02$, $p = .155$), we decided to pursue exploratory subgroup analyses to tentatively isolate the effects of valence. In this regard, it can be assumed that

Figure 4
Summary of the Main Findings Obtained for Each Analysis



Note. Neg = negative; Neu = neutral; Pos = positive; SM = source memory. For the moderator and subgroup analyses, only the statistically significant results were presented.

for the studies controlling stimulus arousal across valence categories, the effects observed on source memory are likely to be driven by valence more than arousal. Nonetheless, the subgroup analyses were in agreement with the moderation analysis, showing that the effects of negative relative to neutral stimuli on source memory were not statistically significant regardless of the presence ($z = -1.20$, $p = .232$) or absence ($z = -1.37$, $p = .169$) of control for the arousal properties.

Task-Related Factors. When probing task-related conditions, the type of source memory task ($Q_M(2) = 7.44$, $p = .024$) and the type of source memory feature ($Q_M(6) = 40.39$, $p < .001$) also appeared as moderators of the overall effect of negative emotion on source memory. With respect to the type of task, although the mean effect of internal source memory tasks (−.28) was not significantly lower than the mean effect of external source memory tasks (−.18; $z = -.65$, $p = .516$), the mean effect of reality monitoring tasks (.24) was significantly higher than that of external source memory tasks ($z = 2.36$, $p = .018$). Furthermore, subgroup analyses revealed no statistically significant changes in the overall effect for external source memory ($z = -1.85$, $p = .065$) and reality monitoring tasks ($z = 1.44$, $p = .150$). Notwithstanding, in the case of internal source memory, the performance was improved for neutral compared with negative stimuli ($z = -3.93$, $p < .001$; $d_{unb} = -.27$, $SE = .07$).

Regarding the type of source memory feature, the mean effects of color (−.17), encoding task (−.27), location (−.01), modality of presentation (.49), and temporal features (.12) were larger than the mean effect of background (−.65; see Table 2). No difference was observed between the latter and the mean effect of identity (−.52; $z = .50$, $p = .614$). Subgroup analyses demonstrated worse source memory for negative relative to neutral stimuli in the subset of studies that manipulated background features ($d_{unb} = -.66$, $SE = .19$, $z = -3.56$, $p < .001$). An effect of similar magnitude was found for identity ($d_{unb} = -.54$, $SE = .11$, $z = -4.83$, $p < .001$). This pattern was also observed for encoding task ($d_{unb} = -.26$, $SE = .07$, $z = -4.07$, $p < .001$). In contrast, a beneficial effect of negative information on source memory was obtained in the subset of studies that manipulated modality of presentation ($d_{unb} = .49$, $SE = .07$, $z = -6.60$, $p < .001$). Moreover, no effect of negative (vs. neutral) stimuli on source memory was found for color ($z = -.85$, $p = .393$), location ($z = -.11$, $p = .910$), and temporal features ($z = 1.27$, $p = .204$).

Even though no other task-related factor emerged as statistically significant moderator (see Table 2), additional exploratory subgroup analyses were run on the type of source memory feature to test some theoretical predictions. The results revealed that no statistically significant effect emerged in the subgroup of studies including intrinsic features ($z = -.77$, $p = .439$). In contrast, a small effect was observed in the subgroup of studies testing extrinsic features ($d_{unb} = -.21$, $SE = .09$, $z = -2.23$, $p = .026$), suggesting an impairment effect of negative (vs. neutral) stimuli on source memory.

Sociodemographic Factors. Considering the impact of sociodemographic-related moderators, namely the number of female participants in the sample and the M_{age} of participants, only the first emerged as a statistically significant moderator ($Q_M(1) = 9.94$, $p = .002$). Specifically, as the number of female participants

Table 2
Summary of the Results of the Moderator Analyses for Negative Versus Neutral Valence

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UB]	Q_E, *p*	Q_M, *p*
Neg-Neu									
N of women	91	53	−0.02	0.01	−3.15	.002	[−0.03, −0.01]	388.67, <.001	9.94, .002
M_{age}	94	54	0.001	0.01	0.11	.916	[−0.02, 0.02]	507.46, <.001	0.01, .916
Type of study material	97	56						466.72, <.001	8.90, .012
Pictures			−0.67	0.23	−2.90	.004	[−1.12, −0.22]		
Words			−0.66	0.23	−2.85	.004	[−1.11, −0.21]		
Both (reference)			0.46	0.22	2.17	.030	[0.05, 0.88]		
Control of arousal	90	53						496.23, <.001	2.02, .155
Yes			−0.17	0.12	−1.42	.155	[−0.41, 0.07]		
No (reference)			−0.10	0.07	−1.41	.158	[−0.25, 0.04]		
Type of SM task	97	56						493.88, <.001	7.44, .024
External (reference)			−0.18	0.09	−2.15	.032	[−0.35, −0.02]		
Internal			−0.09	0.15	−0.65	.516	[−0.38, 0.19]		
Reality			0.43	0.18	2.36	.018	[0.07, 0.78]		
Type of SM feature	97	56						316.50, <.001	40.39, <.001
Color			0.48	0.20	2.37	.018	[0.10, 0.88]		
Encoding task			0.38	0.17	2.21	.027	[0.04, 0.72]		
Identity			0.13	0.26	0.50	.614	[−0.38, 0.64]		
Location			0.64	0.19	3.38	<.001	[0.27, 1.01]		
Modality of presentation			1.14	0.21	5.47	<.001	[0.73, 1.55]		
Temporal			0.78	0.19	4.16	<.001	[0.41, 1.14]		
Background (reference)			−0.65	0.15	−4.49	<.001	[−0.94, −0.37]		
Extrinsic versus intrinsic feature	97	56						511.39, <.001	0.99, .320
Extrinsic (reference)			−0.21	0.09	−2.27	.023	[−0.38, 0.03]		
Intrinsic			0.13	0.13	1.00	.320	[−0.13, 0.39]		
Encoding intentionality	95	55						508.53, <.001	0.20, .652
Intentional			0.06	0.13	0.45	.652	[−0.19, 0.31]		
Incidental (reference)			−0.17	0.10	−1.77	.077	[−0.36, 0.02]		
Retention interval	89	49	1.80×10^{-5}	1.16×10^{-5}	1.54	.123	[-4.86×10^{-6}, 4.08×10^{-5}]	465.12, <.001	2.38, .123

Note. CI = confidence interval; LB = lower bound; UB = upper bound; Neg = negative; Neu = neutral; SM = source memory.

increased, the source memory performance for negative stimuli decreased compared with neutral stimuli ($z = -3.15$, $p = .002$), suggesting a greater impairment effect of negative stimuli on source memory.

Positive Versus Neutral Valence. The overall effect of positive information on source memory accuracy was statistically significant, $d_{unb} = -.11$, $SE = .05$, $z = -2.47$, $p = .014$, 95% CI [−.20, −.02] (see Supplemental Materials Figure S2B and Table S4). Once again, moderator and subgroup analyses may clarify further the effects of positive information on source memory.

Moderator and Subgroup Analyses. Only the type of source memory feature ($Q_M(6) = 21.91$, $p = .001$) emerged as a candidate moderator of the overall effect of positive stimuli on source memory. No other moderator analysis was statistically significant (see Table 3).

Task-Related Factors. Even though the type of source memory feature appeared as a candidate moderator of the overall effect, none of the levels of this moderator reached the threshold of statistical significance (see Table 3). Notwithstanding, subgroup analyses showed that the manipulation of the encoding task and of the identity led to worse source memory for positive relative to neutral information ($d_{unb} = -.18$, $SE = .06$, $z = -3.16$, $p = .002$; $d_{unb} = -.47$, $SE = .16$, $z = -2.88$, $p = .004$, respectively). No other statistically significant effects emerged in these analyses (background: $z = -1.30$, $p = .194$; color: $z = .62$, $p = .538$; location: $z = .35$, $p = .730$; modality of presentation: $z = 1.38$, $p = .168$; temporal: $z = .57$, $p = .571$).

Similar to what was described for the negative versus neutral comparison, exploratory subgroup analyses were run on the type of source memory feature: whereas no statistically significant effect was observed in the subgroup of studies testing intrinsic features ($z = -1.43$, $p = .153$), a very small effect was found in the subgroup of studies testing extrinsic features ($d_{unb} = -.12$, $SE = .05$, $z = -2.42$, $p = .016$). Thus, the impairment effect reported for extrinsic features in the case of the negative items was extended to the positive items.

In addition, we used exploratory subgroup analyses regarding the control of stimulus arousal (stimulus-related factor). Nonetheless, the results were also nonsignificant (subgroup that controlled for arousal: $z = -1.82$, $p = .069$; subgroup that did not control for arousal: $z = -1.07$, $p = .286$).

Negative Versus Positive Valence. When comparing these two valence categories, the overall effect was not statistically significant, $d_{unb} = -.04$, $SE = .05$, $z = -.90$, $p = .371$, 95% CI [−.13, .05], $Q(80) = 319.40$, $p < .001$ (see Supplemental Materials Figure S2C and Table S4).

Moderator and Subgroup Analyses. One stimulus-related factor, the type of study material ($Q_M(2) = 25.05$, $p < .001$), and two task-related factors, the type of source memory task ($Q_M(2) = 16.68$, $p < .001$) and the type of source memory feature ($Q_M(6) = 25.81$, $p < .001$), influenced the effect of stimulus valence on source memory. No other factors emerged as moderators of the effect of emotion on source memory (see Table 4).

Table 3
Summary of the Results of the Moderator Analyses for Positive Versus Neutral Valence

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UB]	Q_E, *p*	Q_M, *p*
Pos-Neu									
N of women	70	36	−0.006	0.004	−1.58	.115	[−0.01, 0.001]	156.07, <.001	2.48, .115
M_{age}	67	34	−0.01	0.01	−1.62	.104	[−0.02, 0.002]	154.49, <.001	2.64, .104
Type of study material	70	36						165.24, <.001	3.22, .200
Pictures			−0.25	0.19	−1.36	.173	[−0.62, 0.11]		
Words			−0.32	0.18	−1.76	.079	[−0.67, 0.04]		
Both (reference)			0.16	0.17	0.95	.341	[−0.17, 0.50]		
Control of arousal	62	33						157.58, <.001	1.62, .203
Yes			−0.13	0.10	−1.27	.203	[−0.32, 0.07]		
No (reference)			−0.06	0.05	−1.10	.273	[−0.17, 0.05]		
Type of SM task	70	36						155.67, <.001	4.24, .120
External (reference)			−0.02	0.06	−0.25	.802	[−0.13, 0.10]		
Internal			−0.17	0.10	−1.89	.059	[−0.34, 0.01]		
Reality			−0.18	0.14	−1.36	.173	[−0.45, 0.10]		
Type of SM feature	70	36						135.10, <.001	21.91, .001
Color			0.26	0.17	1.54	.124	[−0.07, 0.59]		
Encoding task			0.05	0.15	0.32	.746	[−0.25, 0.35]		
Identity			−0.22	0.19	−1.13	.257	[−0.26, 0.16]		
Location			0.27	0.18	1.55	.121	[−0.07, 0.61]		
Modality of presentation			0.40	0.21	1.92	.055	[−0.01, 0.81]		
Temporal			0.33	0.17	1.94	.053	[−0.004, 0.67]		
Background (reference)			−0.23	0.14	−1.61	.107	[−0.51, 0.05]		
Extrinsic versus intrinsic feature	70	36						170.94, <.001	0.12, .725
Extrinsic (reference)			−0.13	0.06	−2.01	.045	[−0.25, 0.003]		
Intrinsic			0.03	0.09	0.35	.725	[−0.15, 0.21]		
Encoding intentionality	68	35						163.87, <.001	0.004, .952
Intentional			0.006	0.09	0.06	.952	[−0.18, 0.19]		
Incidental (reference)			−0.10	0.07	−1.38	.168	[−0.25, 0.04]		
Retention interval	61	29	7.85×10^{-6}	1.23×10^{-5}	0.64	.521	[-1.62×10^{-5}, 3.19×10^{-5}]	149.93, <.001	0.41, .521

Note. CI = confidence interval; LB = lower bound; UB = upper bound; Neu = neutral; Pos = positive; SM = source memory.

Stimulus-Related Factors. Regarding the type of study material, the mean effect for both pictures (.01) and words (−.18) was lower than the mean effect of including both types of stimuli (.48; $z = -3.35$, $p < .001$; $z = -4.90$, $p < .001$, respectively). Subgroup analyses revealed that negative words hampered source memory performance compared with positive words ($d_{unb} = -.18$, $SE = .06$, $z = -2.80$, $p = .005$). In contrast, for studies that used both pictures and words, the effect was in the opposite direction and of medium magnitude ($d_{unb} = .49$, $SE = .11$, $z = 4.34$, $p < .001$). Additionally, the overall effect for the subset of studies that used pictures did not deviate statistically from zero ($d_{unb} = -.03$, $SE = .06$, $z = -.44$, $p = .657$).

Exploratory subgroup analyses were conducted considering the control of stimulus arousal, yielding no statistically significant effects (subgroup that controlled for arousal: $z = -1.85$, $p = .065$; subgroup that did not control for arousal: $z = -.94$, $p = .346$).

Task-Related Factors. With respect to the type of source memory task, the mean effect for reality monitoring tasks (.38) was higher than the mean effect for external source memory tasks (−.09; $z = 3.71$, $p < .001$). The difference in the mean effect for external and internal source memory tasks was nonsignificant ($z = -.35$, $p = .730$). Moreover, the results from the subgroup analyses showed that whereas the overall effect for internal source memory remained nonsignificant ($z = -1.76$, $p = .079$), the overall effect concerning reality monitoring was statistically significant, indicating improved source memory accuracy for negative compared with positive stimuli ($d_{unb} = .38$, $SE = .12$, $z = 3.23$, $p = .001$). Also, the overall effect was statistically significant in the case of external source memory ($d_{unb} = -.17$, $SE = .08$, $z = -2.04$, $p = .042$), suggesting an impairment effect for negative (vs. positive) items.

Even though the moderator analysis focusing on the type of source memory feature was statistically significant, none of the comparisons between the mean effect of background versus other features reached statistical significance (see Table 4). The subgroup analyses revealed that only studies manipulating the modality of presentation yielded a significant overall effect, which showed better source memory performance for negative (vs. positive) information ($d_{unb} = .49$, $SE = .11$, $z = 4.34$, $p < .001$). The other subgroup analyses remained nonsignificant (background: $z = 1.25$, $p = .212$; color: $z = -1.47$, $p = .141$; encoding task: $z = -1.79$, $p = .073$; identity: $z = .58$, $p = .561$; location: $z = -1.71$, $p = .087$; temporal features: $z = .12$, $p = .903$).

Exploratory subgroup analyses were also conducted considering the type of source memory feature. The results showed that whereas no effect was observed in the subgroup of studies including extrinsic features ($z = .24$, $p = .812$), there was a very small statistically significant effect in the subgroup of studies including intrinsic features ($d_{unb} = -.19$, $SE = .09$, $z = -2.05$, $p = .041$). As such, source memory performance was better for positive than for negative stimuli when the source feature was integrated in the study item.

Table 4
Summary of the Results of the Moderator Analyses for Negative Versus Positive Valence

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UB]	Q_E, *p*	Q_M, *p*
Neg-Pos									
N of women	81	45	−0.002	0.01	−0.35	.724	[−0.01, 0.01]	318.99, <.001	0.13, .724
M_{age}	74	40	0.01	0.01	0.88	.380	[−0.01, 0.02]	251.49, <.001	0.77, .380
Type of study material	81	45						229.48, <.001	25.05, <.001
Pictures			−0.46	0.14	−3.35	<.001	[−0.73, −0.19]		
Words			−0.65	0.13	−4.90	<.001	[−0.92, −0.39]		
Both (reference)			0.48	0.12	3.91	<.001	[0.24, 0.72]		
Control of arousal	69	39						239.51, <.001	0.62, .430
Yes			0.12	0.16	0.79	.430	[−0.18, 0.43]		
No (reference)			−0.21	0.15	−1.40	.162	[−0.50, 0.08]		
Type of SM task	81	45						252.11, <.001	16.00, <.001
External (reference)			−0.09	0.06	−1.58	.114	[−0.21, 0.02]		
Internal			−0.03	0.09	−0.35	.730	[−0.21, 0.15]		
Reality			0.47	0.13	3.71	<.001	[0.22, 0.72]		
Type of SM feature	81	45						227.21, <.001	25.73, <.001
Color			−0.38	0.20	−1.95	.052	[−0.76, 0.003		
Encoding task			−0.26	0.18	−1.42	.156	[−0.61, 0.10]		
Identity			−0.23	0.25	−0.91	.363	[−0.72, 0.26]		
Location			−0.25	0.20	−1.26	.209	[−0.65, 0.14]		
Modality of presentation			0.34	0.21	1.61	.107	[−0.07, 0.75]		
Temporal			−0.05	0.21	−0.24	.811	[−0.45, 0.36]		
Background (reference)			0.14	0.17	0.81	.417	[−0.20, 0.47]		
Extrinsic versus intrinsic feature	81	45						305.37, <.001	2.07, .150
Extrinsic (reference)			0.02	0.06	0.27	.785	[−0.11, 0.14]		
Intrinsic			−0.14	0.10	−1.44	.150	[−0.33, 0.05]		
Encoding intentionality	77	43						308.26, <.001	1.08, .299
Intentional			−0.10	0.10	−1.04	.299	[−0.30, 0.09]		
Incidental (reference)			0.02	0.08	0.19	.846	[−0.14, 0.17]		
Retention interval	70	36	1.42×10^{-5}	1.39×10^{-5}	1.02	.309	[-1.31×10^{-5}, 4.14×10^{-5}]	214.37, <.001	1.04, .309

Note. CI = confidence interval; LB = lower bound; UB = upper bound; Neg = negative; Pos = positive; SM = source memory.

Arousal-Based Effects

High Versus Low Arousal. Considering studies that manipulated stimulus arousal, a statistically significant small effect of arousal on source memory was observed, indicating improved performance for stimuli with high versus low arousal, d_{unb} = .27, *SE* = .11, *z* = 2.42, *p* = .016, 95% CI [.05, .48], *Q*(19) = 150.35, *p* < .001 (see Supplemental Materials Figure S2D and Table S4).

Moderator and Subgroup Analyses. Two factors were found to moderate the effect of stimulus arousal on source memory, namely: control for valence across arousal levels ($Q_M(1)$ = 6.06, *p* = .014) and encoding intentionality ($Q_M(1)$ = 5.59, *p* = .018). No other factors, including those that were found for the valence-based comparisons, emerged as significant moderators (see Table 5).

Stimulus-Related Factors. The mean effect was lower when valence was controlled across stimuli differing in arousal (−.20) than when no control was ensured (.41; *z* = −2.46, *p* = .014). Exploratory subgroup analyses revealed that for the subgroup of studies controlling for valence the overall effect was not statistically significant (*z* = −1.05, *p* = .296). In contrast, in the subgroup of studies not controlling for valence, the overall effect was small and significant, and in the same direction of the main analysis (d_{unb} = .38, *SE* = .12, *z* = 3.21, *p* = .001).

Task-Related Factors. The moderator analysis revealed that the effect of arousal on source memory was diminished when intentional learning instructions (−.06) were used compared with incidental learning conditions (.41; *z* = −2.36, *p* = .018). Similar to what was done in prior comparisons, subsidiary subgroup analyses were conducted regarding the type of source memory feature: extrinsic or intrinsic. A small and statistically significant overall effect was found considering the subgroup of studies that manipulated intrinsic features (d_{unb} = .32, *SE* = .11, *z* = 3.00, *p* = .003). Regarding the subgroup of studies examining extrinsic features, no statistically significant effect emerged (*z* = −.24, *p* = .812).

Medium Versus Low Arousal. Similarly to the high versus low arousal comparison, source memory performance was improved for stimuli with medium compared with low arousal levels, d_{unb} = .49, *SE* = .18, *z* = 2.65, *p* = .008, 95% CI [.13, .85], *Q*(4) = 15.04, *p* = .005 (see Supplemental Materials Figure S2E and Table S4). Due to the reduced number of experiments, no moderator or subgroup analyses were conducted in this case.

High Versus Medium Arousal. The overall effect did not diverge from zero, d_{unb} = −.12, *SE* = .15, *z* = −.82, *p* = .414, 95% CI [−.41, .17], *Q*(4) = 16.16, *p* = .003 (see Supplemental Materials Figure S2F and Table S4). For the same reason described for the high versus medium arousal comparison, no moderator or subgroup analyses were pursued.

Discussion

The current study aimed to provide an up-to-date meta-analytic review of behavioral studies probing the effects of emotional

Table 5
Summary of the Results of the Moderator Analyses for High Versus Low Arousal

Moderators	*N* of effects (*k*)	*N* of study samples (clusters)	Estimate	*SE*	*z*	*p*	95% CI [LB, UB]	Q_E, *p*	Q_M, *p*
High-Low									
N of women	17	11	−0.01	0.003	−1.40	.163	[−0.01, 0.002]	130.47, <.001	1.95, .163
M_{age}	16	12	−0.01	0.05	−0.24	.810	[−0.11, 0.09]	45.80, <.001	0.06, .810
Type of study material	20	14						149.99, <.001	0.09, .957
Pictures			0.05	0.45	0.10	.919	[−0.84, 0.93]		
Words			−0.04	0.51	−0.08	.934	[−1.04, 0.95]		
Both (reference)			0.24	0.43	0.56	.575	[−0.60, 1.08]		
Control of valence	17	12						83.51, <.001	6.06, .014
Yes			−0.61	0.25	−2.46	.014	[−1.10, −0.13]		
No (reference)			0.41	0.12	3.45	<.001	[0.18, 0.65]		
Type of SM task	20	14						150.11, <.001	0.004, .949
External (reference)			0.27	0.12	2.26	.024	[0.04, 0.50]		
Internal	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Reality			−0.03	0.43	−0.06	.949	[−0.87, 0.82]		
Type of SM feature	20	14						98.15, <.001	4.57, .334
Color			0.35	0.46	0.76	.447	[−0.55, 1.25]		
Encoding task	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Identity			0.87	0.55	1.57	.116	[−0.21, 1.94]		
Location			0.70	0.37	1.87	.061	[−0.03, 1.42]		
Modality of presentation			0.57	0.52	1.09	.275	[−0.45, 1.58]		
Temporal	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Background (reference)			−0.33	0.35	−0.93	.350	[−1.01, 0.36]		
Extrinsic versus intrinsic feature	20	14						111.63, <.001	2.10, .148
Extrinsic (reference)			−0.07	0.25	−0.28	.780	[−0.57, 0.43]		
Intrinsic			0.40	0.28	1.45	.148	[−0.14, 0.95]		
Encoding intentionality	20	14						92.22, <.001	5.59, .018
Intentional			−0.47	0.20	−2.36	.018	[−0.86, −0.08]		
Incidental (reference)			0.41	0.11	3.68	<.001	[0.19, 0.63]		
Retention interval	20	14	−0.00002	0.0002	−0.07	.943	[−0.0004, 0.0004]	150.16, <.001	0.005, .943

Note. CI = confidence interval; LB = lower bound; UB = upper bound; n/a = not applicable; SM = source memory.

arousal and valence on source memory. More specifically, we included studies that manipulated stimulus valence and arousal, as well as studies examining arousal effects only. With respect to valence-based studies, the meta-analytic results indicated that emotional stimuli, negative or positive, had a deleterious effect on source memory accuracy compared with neutral stimuli. The accuracy was similar considering the comparison between negative and positive items. In the case of arousal-based studies, source memory performance was improved for both high and medium compared with low arousing stimuli with no statistically significant differences between high and medium arousal. While the magnitude of the overall effect sizes ranged from small to medium for the arousal-based studies, the magnitude of the effects was consistently very small considering the valence-based studies. Even considering that small effects obtained in psychological research may be likely significant at some point in the future and should not be disregarded (Funder & Ozer, 2019), especially when evidence has the potential to amass over time, the interpretation of these meta-analytic findings warrant a word of caution. Although the experiments included in this review were methodologically sound (i.e., most of them showed high methodological quality), supporting an overall low risk of bias, the same cannot be said for publication bias. Specifically, it was not possible to compute publication bias for two of the arousal-based comparisons, and there was evidence of bias for the negative-neutral comparison. Furthermore, it is important to acknowledge the role of methodological heterogeneity and, consequently, the role of moderator variables (see Figure 4). In this context, the following discussion subsections will address the role of specific task-related (type of source memory task, type of source memory feature), stimulus-related (control for the affective properties of the stimuli, type of study material) and sociodemographic (sex, age) factors identified throughout this article, as well as possible theoretical implications and main caveats associated with this review.

The Role of Task-Related Factors in the Effects of Emotion on Source Memory

Regarding studies that used internal source memory tasks, it was observed that accuracy was improved for neutral compared with negative stimuli. Additionally, as there was an overlap between internal source memory tasks and the manipulation of the encoding task as a source feature, it was not surprising that source memory performance for the encoding task was improved for neutral compared with both negative and positive stimuli. Taken together, these findings suggest that emotional stimuli, especially the negative ones, are likely to impair internal source memory performance (Cook et al., 2007; Ferré et al., 2019, Experiments 2 and 3; Mao et al., 2015; Newsome et al., 2012; Otani, Jaffa, et al., 2012; Otani, Libkuman, et al., 2012).

The analyses including negative stimuli also showed that the mean effect for reality monitoring tasks was larger compared with the mean effect of external source memory tasks. Remarkably, there was a source memory advantage for negative compared with positive stimuli in reality monitoring tasks, which contrasted with the direction of the overall effects and the effects for external and internal source memory tasks. A closer look into the characteristics of the reality monitoring studies indicated that there were two main manipulations: the identity of who provided the information (Le Bigot et al., 2018; McKague et al., 2012; Morrison & Haddock, 1997); and the modality of presentation involving the discrimination between seen and imagined information (Kensinger, O'Brien, et al., 2007; Kensinger & Schacter, 2006b; Mathews et al., 2013; Robin et al., 2018). Concerning this last discrimination, it is well known that the implementation of visual imagery strategies during encoding benefits memory (e.g., Oliver et al., 2016), and this is why it is a recommended compensatory strategy in clinical settings (e.g., Cicerone et al., 2019). Nonetheless, it is unclear why these reality monitoring findings were observed in the case of negative but not positive stimuli. One tentative explanation can be derived from previous studies demonstrating that when participants receive a visual imagery instruction (vs. a verbal processing instruction to focus on the stimulus meaning) to process a negative event, they are more likely to report increased state anxiety, whereas the opposite seems to occur when the event is positive (Holmes & Mathews, 2005, 2010; Holmes et al., 2008). Moreover, descriptions of the generated images were more likely to incorporate a personal account of the events, including more specific and emotion-related details (Holmes et al., 2008). Thus, imagining negative events seems to generate distinct details that may later facilitate the discrimination between imagined and seen events in source memory decisions, as indicated by previous evidence (Kensinger, O'Brien, et al., 2007; Kensinger & Schacter, 2006b; Mathews et al., 2013).

Given the overlap between the reality monitoring studies and the studies manipulating the modality of presentation (i.e., seen vs. imagine), it is plausible to extend this same explanation to the source memory benefit observed for negative (vs. neutral and positive) stimuli in these later investigations. For the same reason, this explanation may also be suitable to understand the moderator effects found for the type of study material (i.e., the superior mean effect documented for studies using both pictures and words compared with studies using pictures or words alone), as well as the subgroup analyses showing a better source memory accuracy for negative (vs. neutral and positive) items in the case of studies that included both pictures and words.

However, not all reality monitoring discriminations were benefited by negative valence. In fact, subgroup analyses indicated that studies whose source manipulation was based on the identity of who provided the information reported a deleterious effect of emotion (negative and positive) on source memory (Le Bigot et al., 2018; Morrison & Haddock, 1997). The ABC theory (Mather & Sutherland, 2011) may grant a useful account to conciliate these enhancement and impairment effects in the context of reality monitoring. Specifically, when participants were prompted to mentally visualize a specific item, especially a negative one, they were more likely to be aroused and to incorporate self-related information into the mental image (Holmes & Mathews, 2005; Holmes et al., 2008). Thereby, items and respective mental images were potentially tagged as highly relevant during encoding and later used as diagnostic cues to discriminate between sources; thereby, improving source memory. In contrast, when identity was implemented as a source manipulation, the presence of emotion most likely led to the prioritization of other features over identity, hence resulting in worse source memory performance.

Besides identity, stimulus background was another type of source memory feature that, in the subgroup analysis, was related to worse source memory performance for negative compared with neutral stimuli. Such effect was expected given the converging evidence from studies that manipulated this feature (Jia et al., 2018; MacKenzie et al., 2015; Matsumoto & Kawaguchi, 2020; Mao et al., 2017; Rimmele et al., 2011). Moreover, this finding dovetails with the object-based binding theory (Mather, 2007) and the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959), in which stimulus background is posited as an extrinsic/peripheral detail.

Regarding other types of source memory features, including stimulus color, location, and temporal features, the analyses yielded a similar source memory performance for emotional and neutral stimuli. It is worth noting, however, that overall the findings were highly inconsistent, including impairment effects (color: MacKenzie et al., 2015, Experiment 2; location: Maddock & Frein, 2009; temporal: Aupée, 2007; Maddock & Frein, 2009, Experiment 3), enhancement effects (color: Kensinger & Corkin, 2003; Zhou et al., 2020; location: Novak & Mather, 2009; Rimmele et al., 2011, Experiment 2; Schmidt et al., 2011; temporal: D'Argembeau & Van der Linden, 2005; Rimmele et al., 2011, Experiment 3; Schmidt et al., 2011, Experiment 1; Yick et al., 2015), and even similar performances in the comparison between emotional and neutral stimuli (color: Dougal et al., 2007; Wang & Fu, 2011; Zhou et al., 2020; location: Koenig & Mecklinger, 2008; temporal: Koenig & Mecklinger, 2008; Minor & Herzmann, 2019; Schmidt et al., 2011, Experiment 2). This variability likely contributed to the lack of a detectable emotion effect on source memory.

With respect to encoding intentionality, previous studies suggested that the effect of emotional stimuli on source memory might be identified under incidental but not intentional learning instructions (D'Argembeau & Van der Linden, 2004; Ferré et al., 2019). Although encoding intentionality did not moderate the effect for valence-based comparisons, the mean effect for incidental conditions in the high versus low comparison was higher than the mean effect for intentional conditions. That is, intentional learning might dampen arousal effects on source memory. When participants are instructed to memorize the stimuli, knowing that their memory will be tested at some point, they are more likely to engage in effortful encoding and to prioritize both emotional and neutral items. As a result, source memory performance may be similar irrespective of the emotional quality of the stimuli. Notwithstanding, prior studies also noted that encoding intentionality may interact with other factors such as type of source memory features, being more relevant for some than others (e.g., D'Argembeau & Van der Linden, 2004; Kensinger & Schacter, 2006b). Thus, more studies manipulating encoding intentionality while testing different source memory features might be useful to understand these possible interaction effects. Even so, the fact that the high versus low arousal meta-analysis was composed of fewer studies and characterized by lower variability in terms of source features (see Supplemental Materials Table S8) might have contributed to the emergence of this moderation effect. By the same token, the great variability of source memory features evaluated in the valenced-based comparisons could have masked the influence of the encoding intentionality in the analyses.

The delay time between the study and the test phase was another task-related factor of interest but it did not influence the overall effects. This observation is in line with previous studies that investigated different delay times, showing that the effect of emotional stimuli on source memory was somehow stable across time (Sharot & Yonelinas, 2008; Wang & Fu, 2011). It should be noted, however, that most studies included in the current meta-analytic study relied on immediate recall or a short delay (see Supplemental Materials Tables S5–S8). Therefore, studies with longer delays and that systematically assess different delay intervals are necessary, similarly to what has been conducted for item memory research (e.g., Sharot & Phelps, 2004; Wirkner et al., 2018; Yonelinas & Ritchey, 2015).

The Role of Stimulus-Related Factors in the Effects of Emotion on Source Memory

Regarding stimulus-related factors, there were two observations of theoretical interest. First, the control of arousal across valence categories did not emerge as a candidate moderator in any of the valence-based comparisons. This finding was also supported by the lack of statistically significant differences documented in the exploratory subgroup analyses. In contrast, the control of stimulus valence in the high versus low comparison influenced the overall meta-analytic effects, indicating that mean effect was higher for experiments that did not control for stimulus valence compared with experiments that performed such control. Also, the subgroup analysis showed that the overall effect increased from .27 to .38 when considering studies that did not control for stimulus valence (see Figure 4).

As valence and arousal dimensions are commonly examined together, the codification regarding the control of the affective properties of the stimuli was an attempt to explore the contribution of each dimension. The results of this approach were not fully satisfactory especially considering the valence-based comparisons. This precludes the specification of differential roles of valence and arousal, and this should be acknowledged as a limitation of the current study. Additionally, the process of selecting and controlling the affective properties of stimuli can be a difficult task when using a dimensional approach. This happens as affective ratings of valence and arousal tend to fit a U-shape, boomerang distribution, meaning that the more positive or negative a certain stimulus is perceived, the more arousing it is; hence, stimuli with lower or higher valence scores are typically rated as more arousing than neutral ones (Kuppens et al., 2017; Lang et al., 1998). Future studies may disentangle the contributions of valence and arousal to source memory performance by testing different levels of arousal (e.g., Vieitez et al., 2021) or including orthogonal manipulations, as it was done in previous memory studies (e.g., Kensinger et al., 2011; Schmidt et al., 2011).

Despite this word of caution, a noteworthy observation was that the direction of the overall effect of emotion on source memory was not convergent when considering valence-based and arousal-based studies: emotional stimuli had a deleterious effect on source memory in valence-based studies, whereas they had an advantageous effect in arousal-based studies. Even though it is important to acknowledge that the nature of the relationship between valence and arousal is still under debate (e.g., Kuppens et al., 2017; Petrolini & Viola, 2020), this observation agrees with the notion that valence and arousal modulate memory in different ways. If we consider the hypothesis that arousal is likely to increase attentional resources during the encoding phase while valence is likely to engage elaboration resources (Kensinger, 2004; Kensinger & Corkin, 2004), this may explain the enhancement effect in the case of arousal-based studies as medium and high (vs. low) arousal stimuli might increase the likelihood of attention allocation to different aspects of an episode, as previously postulated by Yerkes and Dodson (1908). Even though the overall effect was larger for the medium versus low comparison than for the high versus low comparison (see Figure 4), somehow fitting with the Yerkes-Dodson law notion that optimal performance seems to be achieved for moderate levels of arousal, there was no detrimental effect of high arousing stimuli on source memory compared with medium arousing stimuli. Relatedly, the impairment effect found on source memory in valence-based studies may be accounted for by the different elaborative processes that might occur during the encoding of emotional information. When these processes provide diagnostic cues that can be used to support later source memory decisions, such as in the case of reality monitoring tasks involving visual imagery, source memory is enhanced for emotional stimuli. But when these elaborative processes provide cues that are less diagnostic given certain source memory decisions, as it could be the case of the source features typically tested in internal and external source memory tasks, source memory performance for emotional stimuli may be worse or no different from neutral stimuli.

Second, in relation to the type of study material, subgroup analyses revealed an impairment effect of negative (vs. neutral) words on source memory, whereas no statistically significant emotion effects were found in the case of pictures. In this context, there was no clear evidence supporting a memory superiority effect related to emotional pictures (e.g., Paivio et al., 1968). At the same time, the lack of emotion effects may suggest a great variability of findings in the studies using pictures, whereas the findings for words and the combination of both pictures and words were more consistent. A possible explanation for this variability may lie in the databases used for the picture selection, which were more diverse compared with the word stimuli (see Supplemental Materials Table S5 and Table S7). In this scenario, the process of controlling for picture properties (e.g., brightness, spatial frequency, color complexity) that might interfere with the affective response (e.g., Cano et al., 2009; Lakens et al., 2013) was likely more challenging.

The Role of Sociodemographic Factors in the Effects of Emotion on Source Memory

Other two potential moderators explored in the current meta-analytic study were the number of female subjects and the average age of the participants. Specifically, we did not expect any of these factors to influence the emotion effect on source memory. We held this expectation given that most studies were composed of samples of young adults (see Supplemental Materials Tables S5 and S7) and that samples of children, adolescents, and older adults were not included in the review. Future systematic reviews might be conducted to study these different age groups. Moreover, no major sex differences were reported in previous studies comparing female and male participants (D'Argembeau & Van der Linden, 2005; Wang, 2012). Although age did not emerge as a moderator in any of the analyses, it was observed that as the number of females increased, the advantage of neutral over negative stimuli on source memory accuracy was enhanced. Previous studies revealed that females tend to present increased activation for negative information in brain regions involved in emotional processing (see Stevens & Hamann, 2012). In addition, females tend to present an advantage in episodic memory tasks implying verbal abilities (Asperholm et al., 2019). However, few studies have addressed sex differences in emotional source memory, and future studies could address the role of this factor.

Theoretical Considerations

Regarding the distinction between intrinsic/central and extrinsic/peripheral features, which was tested as a moderator variable following the object-based binding theory (Mather, 2007) and the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959), this factor did not emerge as a statistically significant moderator in any of the analyses. Nevertheless, exploratory subgroup analyses indicated that whereas emotional (vs. neutral) stimuli impaired source memory performance for extrinsic/peripheral features, as postulated by the former theoretical frameworks, no beneficial effects were found for negative/positive (vs. neutral) stimuli in the case of intrinsic/central features. Furthermore, for specific source features such as stimulus color and spatiotemporal details, which are usually regarded as intrinsic and expected to be better remembered in the presence of emotion, the current meta-analyses failed to corroborate a beneficial effect. An enhancement effect considering intrinsic/central features was rather observed for positive compared with negative stimuli. Considering that positive (vs. negative) information has more associative potential and promotes semantically meaningful integration of elements (Unkelbach et al., 2020), positive information might also favor the within-item binding; thus, resulting in better source memory performance for intrinsic/central features. This benefit for intrinsic features also emerged in the high versus low comparison. However, most of the experiments in this case were coded as testing intrinsic features, with only two including extrinsic features. As a result, there was a substantial overlap between the experiments composing the main analysis and the experiments included in the subgroup analysis. Also, no reliable information can be drawn for extrinsic features due to the reduced number of experiments.

Taken together, the predictions of the object-based binding theory (Mather, 2007) and the Easterbrook's cue-utilization hypothesis (Easterbrook, 1959) regarding intrinsic/central and extrinsic/peripheral features were partially supported. Nonetheless, the findings also highlighted the relevance of incorporating differences in the processing of positive and negative information in current theoretical accounts. In this context, the ABC theory (Mather & Sutherland, 2011) might accommodate both arousal- and valence-related effects by placing emphasis in the distinction between low and high priority features or between goal-relevant or goal-irrelevant features, which is also in agreement with the notion that motivational goals modulate source memory decisions as proposed by the original source monitoring framework (Johnson et al., 1993). Even though the former distinction is not straightforward, it surely deserves further testing in forthcoming experiments (e.g., Mao et al., 2017).

With respect to the source monitoring framework (Johnson et al., 1993), we also aimed to investigate whether the emotion effects on source memory would vary as a function of different source memory tasks (i.e., external, internal, and reality monitoring). Indeed, the type of source memory task appeared not only as a moderator but also as a relevant factor in the subgroup analyses. Specifically, whereas the difference in the mean effect between internal and external source memory tasks did not achieve statistical significance, reality monitoring tasks had a superior mean effect relative to external source memory tasks in the analyses involving negative items. In addition, whereas internal/external source memory tasks and features were more often in the same direction of the overall main effects, indicating a deleterious effect of stimulus emotion on source memory performance, the reality monitoring tasks and specifically the modality of presentation were the only instances where an advantageous effect of negative information on source memory was observed. These findings support the view that internal and external source memory tasks are more demanding than reality monitoring tasks (Johnson & Foley, 1984; Raye & Johnson, 1980), and that the effects of stimulus emotion on source memory might differ in magnitude and even direction as a function of the type of source memory task and source memory feature. In accordance, this review confirmed the contribution of these factors to the explanation of the diverse findings encountered in the emotion and source memory research.

Limitations

First, it is worth noting some factors that might have contributed to publication bias: search conducted in English databases; inclusion of published studies letting out gray literature; it was not possible to incorporate all the eligible studies due to missing data (see Supplemental Materials Tables S5–S8); publication bias tests were not performed for two of the arousal-based comparisons (medium vs. low; high vs. medium) due to the small number of experiments; there was some evidence of bias in the case of the negative versus neutral comparison, even though there were few indicators of publication bias for the remaining comparisons. At the same time, the publication bias methods used are not without limitations. Specifically, the performance of these tests is affected when the true effect sizes are heterogeneous (Ioannidis & Trikalinos, 2007; Song et al., 2013; van Aert et al., 2019).

Second, considering the heterogeneity among experiments, the moderator and subgroup analyses were of the upmost relevance to understand the findings of this systematic review. However, these analyses require a reasonable number of experiments and moderator-related data to be robust. In the current study, there was a limited number of experiments and incomplete information for the selected moderator variables. Furthermore, conducting multiple moderator and subgroup analyses imply multiple statistical comparisons and increased probability of committing Type I error. Thus, even though these analyses were useful to pinpoint specific variables that modulate the relationship between emotion and source memory, they need to be interpreted with caution.

Third, the experiment categorization for some of the moderators was not trivial, being the distinction between intrinsic and extrinsic features a good example. Accordingly, it is not possible to rule out the putative influence of the subjectivity underpinning some of the coding procedures adopted in this study.

Conclusion

The current study aimed to provide a systematic overview on the effects of emotional stimuli on source memory and to identify factors that modulate the magnitude and direction of such effects. For the valence-based analyses (i.e., negative vs. neutral; positive vs. neutral; negative vs. positive), an impairment effect of negative and positive (vs. neutral) stimuli on source memory was observed. Notwithstanding, stimulus- (type of study material) and task-related factors (type of source memory task; type of source memory features) emerged as candidates affecting the former effects. On the contrary, for the arousal-based analyses (i.e., high vs. low; medium vs. low; high vs. medium), the results revealed that high and medium arousal stimuli granted a source memory advantage compared with low arousal stimuli. No difference was observed between high and medium arousal levels. Nonetheless, caution regarding these arousal-based findings is required due to the reduced number of experiments that were included in each analysis.

Overall, this study demonstrated that mixed findings are likely to be expected when probing the effects of emotional stimuli on source memory, which highlights the need to consider the role of methodological factors, such as the type of source memory task and the type of source memory features. In addition, more research is necessary to support theoretical claims, to determine the role of valence, and to reach a broader understanding of the effects of emotion on source memory. This includes studying the effects of emotion based not only on dimensional approaches (Bradley & Lang, 1994; Lang et al., 1993; Russell, 1980) but also on categorical approaches (Levine & Pizarro, 2004), as well as the effects derived from mood or emotional encoding contexts (e.g., Pereira et al., 2021a; Ventura-Bort et al., 2020) for which the extant evidence is still limited.

References

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

Asperholm, M., Högman, N., Rafi, J., & Herlitz, A. (2019). What did you do yesterday? A meta-analysis of sex differences in episodic memory. *Psychological Bulletin*, *145*(8), 785–821. https://doi.org/10.1037/bul0000197

Assink, M., & Wibbelink, C. J. M. (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

*Aupée, A. M. (2007). A detrimental effect of emotion on picture recollection. *Scandinavian Journal of Psychology*, *48*(1), 7–11. https://doi.org/10.1111/j.1467-9450.2006.00533.x

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088–1101. https://doi.org/10.2307/2533446

Bell, R., Mieth, L., & Buchner, A. (2017). Emotional memory: No source memory without old-new recognition. *Emotion*, *17*(1), 120–130. https://doi.org/10.1037/emo0000211

*Blix, I., & Brennen, T. (2011). Intentional forgetting of emotional words after trauma: A study with victims of sexual assault. *Frontiers in Psychology*, *2*, 235. https://doi.org/10.3389/fpsyg.2011.00235

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Routledge. https://doi.org/10.1002/9780470743386

Borg, C., Leroy, N., Favre, E., Laurent, B., & Thomas-Antérion, C. (2011). How emotional pictures influence visuospatial binding in short-term memory in ageing and Alzheimer's disease? *Brain and Cognition*, *76*(1), 20–25. https://doi.org/10.1016/j.bandc.2011.03.008

Bowen, H. J., Kark, S. M., & Kensinger, E. A. (2018). NEVER forget: Negative emotional valence enhances recapitulation. *Psychonomic Bulletin & Review*, *25*(3), 870–891. https://doi.org/10.3758/s13423-017-1313-9

*Boywitt, C. D. (2015). Non-monotonic relationships between emotional arousal and memory for color and location. *Cognition and Emotion*, *29*(8), 1335–1349. https://doi.org/10.1080/02699931.2014.977850

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Brébion, G., David, A. S., Jones, H. M., Ohlsen, R., & Pilowsky, L. S. (2007). Temporal context discrimination in patients with schizophrenia: Associations with auditory hallucinations and negative symptoms. *Neuropsychologia*, *45*(4), 817–823. https://doi.org/10.1016/j.neuropsychologia.2006.08.009

Brookwell, M. L., Bentall, R. P., & Varese, F. (2013). Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: A meta-analytic review. *Psychological Medicine*, *43*(12), 2465–2475. https://doi.org/10.1017/S0033291712002760

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory & Cognition*, *20*(3), 277–290. https://doi.org/10.3758/BF03199665

Cano, M. E., Class, Q. A., & Polich, J. (2009). Affective valence, stimulus attributes, and P300: Color vs. black/white and normal vs. scrambled images. *International Journal of Psychophysiology*, *71*(1), 17–24. https://doi.org/10.1016/j.ijpsycho.2008.07.016

Christianson, S.-Å. (1984). The relationship between induced emotional arousal and amnesia. *Scandinavian Journal of Psychology*, *25*(2), 147–160. https://doi.org/10.1111/j.1467-9450.1984.tb01007.x

Christianson, S.-Å. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, *112*(2), 284–309. https://doi.org/10.1037/0033-2909.112.2.284

Christianson, S.-Å., & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology*, *1*(4), 225–239. https://doi.org/10.1002/acp.2350010402

Christianson, S.-Å., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition and Emotion*, *5*(2), 81–108. https://doi.org/10.1080/02699939108411027

Cicerone, K. D., Goldin, Y., Ganci, K., Rosenbaum, A., Wethe, J. V., Langenbahn, D. M., Malec, J. F., Bergquist, T. F., Kingsley, K., Nagele, D., Trexler, L., Fraas, M., Bogdanova, Y., & Harley, J. P. (2019). Evidence-based cognitive rehabilitation: Systematic review of the literature from 2009 through 2014. *Archives of Physical Medicine and Rehabilitation*, *100*(8), 1515–1533. https://doi.org/10.1016/j.apmr.2019.02.011

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, *122*(3), 211–226. https://doi.org/10.1016/j.bandl.2011.12.007

Cook, G. I., Hicks, J. L., & Marsh, R. L. (2007). Source monitoring is not always enhanced for valenced material. *Memory & Cognition*, *35*(2), 222–230. https://doi.org/10.3758/BF03193443

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.

D'Argembeau, A., & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, *4*(2), 173–188. https://doi.org/10.1037/1528-3542.4.2.173

*D'Argembeau, A., & Van der Linden, M. (2005). Influence of emotion on memory for temporal information. *Emotion*, *5*(4), 503–507. https://doi.org/10.1037/1528-3542.5.4.503

*Davidson, P. S. R., McFarland, C. P., & Glisky, E. L. (2006). Effects of emotion on item and source memory in young and older adults. *Cognitive, Affective & Behavioral Neuroscience*, *6*(4), 306–322. https://doi.org/10.3758/CABN.6.4.306

Del Re, A. C. (2015). A practical tutorial on conducting meta-analysis in R. *The Quantitative Methods for Psychology*, *11*(1), 37–50. https://doi.org/10.20982/tqmp.11.1.p037

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, *1*(1), 5–11. https://doi.org/10.1037/1528-3542.1.1.5

Dolcos, F., Katsumi, Y., Weymar, M., Moore, M., Tsukiura, T., & Dolcos, S. (2017). Emerging directions in emotional episodic memory. *Frontiers in Psychology*, *8*, 1867. https://doi.org/10.3389/fpsyg.2017.01867

*Dougal, S., Phelps, E. A., & Davachi, L. (2007). The role of medial temporal lobe in item recognition and source recollection of emotional stimuli. *Cognitive, Affective & Behavioral Neuroscience*, *7*(3), 233–242. https://doi.org/10.3758/CABN.7.3.233

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177. https://doi.org/10.1037/1082-989X.1.2.170

*Durbin, K. A., Mitchell, K. J., & Johnson, M. K. (2017). Source memory that encoding was self-referential: The influence of stimulus characteristics. *Memory*, *25*(9), 1191–1200. https://doi.org/10.1080/09658211.2017.1282517

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

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3), 183–201. https://doi.org/10.1037/h0047707

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Fairfield, B., Altamura, M., Padalino, F. A., Balzotti, A., Di Domenico, A., & Mammarella, N. (2016). False memories for affective information in schizophrenia. *Frontiers in Psychiatry*, *7*, 191. https://doi.org/10.3389/fpsyt.2016.00191

Ferguson, S. A., Hashtroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging*, *7*(3), 443–452. https://doi.org/10.1037/0882-7974.7.3.443

*Ferré, P., Comesaña, M., & Guasch, M. (2019). Emotional content and source memory for language: Impairment in an incidental encoding task. *Frontiers in Psychology*, *10*, 65. https://doi.org/10.3389/fpsyg.2019.00065

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical & Statistical Psychology*, *63*(3), 665–694. https://doi.org/10.1348/000711010X502733

Ford, J. H., & Kensinger, E. A. (2019). Age-by-emotion interactions in memory retrieval processes: An event-related potential study. *The Journals of Gerontology: Series B*, *74*(7), 1101–1110. https://doi.org/10.1093/geronb/gbx098

Ford, J. H., Morris, J. A., & Kensinger, E. A. (2014). Effects of emotion and emotional valence on the neural correlates of episodic memory search and elaboration. *Journal of Cognitive Neuroscience*, *26*(4), 825–839. https://doi.org/10.1162/jocn_a_00529

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

*Gallant, S. N., & Yang, L. (2014). Positivity effect in source attributions of arousal-matched emotional and non-emotional words during item-based directed forgetting. *Frontiers in Psychology*, *5*, 1334. https://doi.org/10.3389/fpsyg.2014.01334

*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. https://doi.org/10.1037/pag0000352

Galli, G., Vadillo, M. A., Sirota, M., Feurra, M., & Medvedeva, A. (2019). A systematic review and meta-analysis of the effects of transcranial direct current stimulation (tDCS) on episodic memory. *Brain Stimulation*, *12*(2), 231–241. https://doi.org/10.1016/j.brs.2018.11.008

*Gallo, D. A., Korthauer, L. E., McDonough, I. M., Teshale, S., & Johnson, E. L. (2011). Age-related positivity effects and autobiographical memory detail: Evidence from a past/future source memory task. *Memory*, *19*(6), 641–652. https://doi.org/10.1080/09658211.2011.595723

Ghetti, S., & Bunge, S. A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, *2*(4), 381–395. https://doi.org/10.1016/j.dcn.2012.05.002

Hadley, C. B., & Mackay, D. G. (2006). Does emotion help or hinder immediate memory? Arousal versus priority-binding mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 79–88. https://doi.org/10.1037/0278-7393.32.1.79

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

Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Henmi, M., & Copas, J. B. (2010). Confidence intervals for random effects meta-analysis and robustness to publication bias. *Statistics in Medicine*, *29*(29), 2969–2983. https://doi.org/10.1002/sim.4029

Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., & Rockstroh, B. (2006). Processing of emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology*, *43*(2), 197–206. https://doi.org/10.1111/j.1469-8986.2006.00385.x

Higgins, J., & Altman, D. G. (2008). Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions: Cochrane book series* (pp. 187–241). Wiley, Ltd. https://doi.org/10.1002/9780470712184.ch8

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

Holmes, E. A., & Mathews, A. (2005). Mental imagery and emotion: A special relationship? *Emotion*, *5*(4), 489–497. https://doi.org/10.1037/1528-3542.5.4.489

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, *30*(3), 349–362. https://doi.org/10.1016/j.cpr.2010.01.001

Holmes, E. A., Mathews, A., Mackintosh, B., & Dalgleish, T. (2008). The causal effect of mental imagery on emotion assessed using picture-word cues. *Emotion*, *8*(3), 395–409. https://doi.org/10.1037/1528-3542.8.3.395

Hostler, T. J., Wood, C., & Armitage, C. J. (2018). The influence of emotional cues on prospective memory: A systematic review with meta-analyses. *Cognition and Emotion*, *32*(8), 1578–1596. https://doi.org/10.1080/02699931.2017.1423280

Ioannidis, J. P. A., & Trikalinos, T. A. (2007). The appropriateness of asymmetry tests for publication bias in meta-analyses: A large survey. *CMAJ*, *176*(8), 1091–1096. https://doi.org/10.1503/cmaj.060410

*Jia, X., Gao, C., Cui, L., & Guo, C. (2018). Does emotion arousal influence the benefit received from testing: Insights from neural correlates of retrieval mode effect. *NeuroReport*, *29*(17), 1449–1455. https://doi.org/10.1097/WNR.0000000000001130

Johnson, M. K. (1988). Reality monitoring: An experimental phenomenological approach. *Journal of Experimental Psychology: General*, *117*(4), 390–394. https://doi.org/10.1037/0096-3445.117.4.390

Johnson, M. K., & Foley, M. A. (1984). Differentiating fact from fantasy: The reliability of children's memory. *Journal of Social Issues*, *40*(2), 33–50. https://doi.org/10.1111/j.1540-4560.1984.tb01092.x

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3–28. https://doi.org/10.1037/0033-2909.114.1.3

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*(1), 67–85. https://doi.org/10.1037/0033-295X.88.1.67

Jurica, P. J., & Shimamura, A. P. (1999). Monitoring item and source information: Evidence for a negative generation effect in source memory. *Memory & Cognition*, *27*(4), 648–656. https://doi.org/10.3758/BF03211558

*Kaida, K., Niki, K., & Born, J. (2015). Role of sleep for encoding of emotional memory. *Neurobiology of Learning and Memory*, *121*, 72–79. https://doi.org/10.1016/j.nlm.2015.04.002

Kensinger, E. A., Addis, D. R., & Atapattu, R. K. (2011). Amygdala activity at encoding corresponds with memory vividness and with memory for select episodic details. *Neuropsychologia*, *49*(4), 663–673. https://doi.org/10.1016/j.neuropsychologia.2011.01.017

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*(4), 241–251. https://doi.org/10.1515/REVNEURO.2004.15.4.241

Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science*, *16*(4), 213–218. https://doi.org/10.1111/j.1467-8721.2007.00506.x

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, *1*(2), 99–113. https://doi.org/10.1177/1754073908100432

*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

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 3310–3315. https://doi.org/10.1073/pnas.0306408101

Kensinger, E. A., & Kark, S. M. (2018). Emotion and memory. In J. T. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed., pp. 1–26). Wiley. https://doi.org/10.1002/9781119170174.epcn101

Kensinger, E. A., Garoff-Eaton, R. J., & Schacter, D. L. (2007). Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. *Journal of Memory and Language*, *56*(4), 575–591. https://doi.org/10.1016/j.jml.2006.05.004

*Kensinger, E. A., O'Brien, J. L., Swanberg, K., Garoff-Eaton, R. J., & Schacter, D. L. (2007). The effects of emotional content on reality-monitoring performance in young and older adults. *Psychology and Aging*, *22*(4), 752–764. https://doi.org/10.1037/0882-7974.22.4.752

*Kensinger, E. A., & Schacter, D. L. (2005). Emotional content and reality-monitoring ability: fMRI evidence for the influences of encoding processes. *Neuropsychologia*, *43*(10), 1429–1443. https://doi.org/10.1016/j.neuropsychologia.2005.01.004

*Kensinger, E. A., & Schacter, D. L. (2006a). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *The Journal of Neuroscience*, *26*(9), 2564–2570. https://doi.org/10.1523/JNEUROSCI.5241-05.2006

*Kensinger, E. A., & Schacter, D. L. (2006b). Reality monitoring and memory distortion: Effects of negative, arousing content. *Memory & Cognition*, *34*(2), 251–260. https://doi.org/10.3758/BF03193403

Kensinger, E. A., & Schacter, D. L. (2016). Memory and emotion. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (4th ed., pp. 564–578). Guilford Press Publications.

*Koenig, S., & Mecklinger, A. (2008). Electrophysiological correlates of encoding and retrieving emotional events. *Emotion*, *8*(2), 162–173. https://doi.org/10.1037/1528-3542.8.2.162

Kuppens, P., Tuerlinckx, F., Yik, M., Koval, P., Coosemans, J., Zeng, K. J., & Russell, J. A. (2017). The relation between valence and arousal in subjective experience varies with personality and culture. *Journal of Personality*, *85*(4), 530–542. https://doi.org/10.1111/jopy.12258

*Kurinec, C. A., Whitney, P., Hinson, J. M., Hansen, D. A., & Van Dongen, H. P. A. (2021). Sleep deprivation impairs binding of information with its context. *Sleep*, *44*(8), zsab113. https://doi.org/10.1093/sleep/zsab113

Ladouceur, C. D. (2012). Neural systems supporting cognitive-affective interactions in adolescence: The role of puberty and implications for affective disorders. *Frontiers in Integrative Neuroscience*, *6*, 65. https://doi.org/10.3389/fnint.2012.00065

Laier, C., Schulte, F. P., & Brand, M. (2013). Pornographic picture processing interferes with working memory performance. *Journal of Sex Research*, *50*(7), 642–652. https://doi.org/10.1080/00224499.2012.716873

Lakens, D., Fockenberg, D. A., Lemmens, K. P., Ham, J., & Midden, C. J. (2013). Brightness differences influence the evaluation of affective pictures. *Cognition and Emotion*, *27*(7), 1225–1246. https://doi.org/10.1080/02699931.2013.781501

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*(12), 1248–1263. https://doi.org/10.1016/S0006-3223(98)00275-3

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

*Le Bigot, L., Knutsen, D., & Gil, S. (2018). I remember emotional content better, but I'm struggling to remember who said it!. *Cognition*, *180*, 52–58. https://doi.org/10.1016/j.cognition.2018.07.001

*Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Levine, L. J., & Edelstein, R. S. (2009). Emotion and memory narrowing: A review and goal-relevance approach. *Cognition and Emotion*, *23*(5), 833–875. https://doi.org/10.1080/02699930902738863

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition*, *22*(5), 530–554. https://doi.org/10.1521/soco.22.5.530.50767

Lindsay, D. S., & Johnson, M. K. (2000). False memories and the source monitoring framework: Reply to Reyna and Lloyd (1997). *Learning and Individual Differences*, *12*(2), 145–161. https://doi.org/10.1016/S1041-6080(01)00035-8

Lipinska, G., Stuart, B., Thomas, K. G. F., Baldwin, D. S., & Bolinger, E. (2019). Preferential consolidation of emotional memory during sleep: A meta-analysis. *Frontiers in Psychology*, *10*, 1014. https://doi.org/10.3389/fpsyg.2019.01014

*MacKenzie, G., Powell, T. F., & Donaldson, D. I. (2015). Positive emotion can protect against source memory impairment. *Cognition and Emotion*, *29*(2), 236–250. https://doi.org/10.1080/02699931.2014.911145

Madan, C. R., Shafer, A. T., Chan, M., & Singhal, A. (2017). Shock and awe: Distinct effects of taboo words on lexical decision and free recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *70*(4), 793–810. https://doi.org/10.1080/17470218.2016.1167925

*Maddock, R. J., & Frein, S. T. (2009). Reduced memory for the spatial and temporal context of unpleasant words. *Cognition and Emotion*, *23*(1), 96–117. https://doi.org/10.1080/02699930801948977

Mancuso, L. E., Ilieva, I. P., Hamilton, R. H., & Farah, M. J. (2016). Does transcranial direct current stimulation improve healthy working memory?: A meta-analytic review. *Journal of Cognitive Neuroscience*, *28*(8), 1063–1089. https://doi.org/10.1162/jocn_a_00956

*Mao, W. B., An, S., & Yang, X. F. (2017). The effects of goal relevance and perceptual features on emotional items and associative memory. *Frontiers in Psychology*, *8*, 1223. https://doi.org/10.3389/fpsyg.2017.01223

*Mao, X., You, Y., Li, W., & Guo, C. (2015). Emotion impairs extrinsic source memory—An ERP study. *Biological Psychology*, *110*, 182–189. https://doi.org/10.1016/j.biopsycho.2015.07.005

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, *2*(1), 33–52. https://doi.org/10.1111/j.1745-6916.2007.00028.x

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, *16*(3–5), 437–458. https://doi.org/10.1080/026432999380870

*Mather, M., & Knight, M. (2008). The emotional harbinger effect: Poor context memory for cues that previously predicted something arousing. *Emotion*, *8*(6), 850–860. https://doi.org/10.1037/a0014087

*Mather, M., & Nesmith, K. (2008). Arousal-enhanced location memory for pictures. *Journal of Memory and Language*, *58*(2), 449–464. https://doi.org/10.1016/j.jml.2007.01.004

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133. https://doi.org/10.1177/1745691611400234

*Mathews, A., Ridgeway, V., & Holmes, E. A. (2013). Feels like the real thing: Imagery is both more realistic and emotional than verbal thought. *Cognition and Emotion*, *27*(2), 217–229. https://doi.org/10.1080/02699931.2012.698252

*Matsumoto, N., & Kawaguchi, J. (2020). Negative item memory and associative memory: Influences of working memory capacity, anxiety sensitivity, and looming cognition. *Journal of Behavior Therapy and Experimental Psychiatry*, *68*, 101569. https://doi.org/10.1016/j.jbtep.2020.101569

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://doi.org/10.11613/BM.2012.031

*McKague, M., McAnally, K. I., Skovron, M., Bendall, S., & Jackson, H. J. (2012). Source monitoring and proneness to auditory-verbal hallucinations: A signal detection analysis. *Cognitive Neuropsychiatry*, *17*(6), 544–562. https://doi.org/10.1080/13546805.2012.676311

Mickley Steinmetz, K. R., & Kensinger, E. A. (2009). The effects of valence and arousal on the neural activity leading to subsequent memory. *Psychophysiology*, *46*(6), 1190–1199. https://doi.org/10.1111/j.1469-8986.2009.00868.x

*Minor, G., & Herzmann, G. (2019). Effects of negative emotion on neural correlates of item and source memory during encoding and retrieval. *Brain Research*, *1718*, 32–45. https://doi.org/10.1016/j.brainres.2019.05.001

Mitchell, K. J., Mather, M., Johnson, M. K., Raye, C. L., & Greene, E. J. (2006). A functional magnetic resonance imaging investigation of short-term source and item memory for negative pictures. *NeuroReport*, *17*(14), 1543–1547. https://doi.org/10.1097/01.wnr.0000234743.50442.e5

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

*Morrison, A. P., & Haddock, G. (1997). Cognitive factors in source monitoring and auditory hallucinations. *Psychological Medicine*, *27*(3), 669–679. https://doi.org/10.1017/S003329179700487X

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

*Nashiro, K., & Mather, M. (2010). How arousal affects younger and older adults' memory binding. *Experimental Aging Research*, *37*(1), 108–128. https://doi.org/10.1080/0361073X.2011.536746

*Nashiro, K., & Mather, M. (2011). Effects of emotional arousal on memory binding in normal aging and Alzheimer's disease. *The American Journal of Psychology*, *124*(3), 301–312. https://doi.org/10.5406/amerjpsyc.124.3.0301

*Newsome, R. N., Dulas, M. R., & Duarte, A. (2012). The effects of aging on emotion-induced modulations of source retrieval ERPs: Evidence for valence biases. *Neuropsychologia*, *50*(14), 3370–3384. https://doi.org/10.1016/j.neuropsychologia.2012.09.024

*Novak, D. L., & Mather, M. (2009). The tenacious nature of memory binding for arousing negative items. *Memory & Cognition*, *37*(7), 945–952. https://doi.org/10.3758/MC.37.7.945

Ochsner, K. N. (2000). Are affective events richly recollected or simply familiar? The experience and process of recognizing feelings past. *Journal of Experimental Psychology: General*, *129*(2), 242–261. https://doi.org/10.1037/0096-3445.129.2.242

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

Oliver, M. C., Bays, R. B., & Zabrucky, K. M. (2016). False memories and the DRM paradigm: Effects of imagery, list, and test type. *The Journal of General Psychology*, *143*(1), 33–48. https://doi.org/10.1080/00221309.2015.1110558

*Otani, H., Jaffa, M., Libkuman, T. M., Goernert, P. N., & Kato, K. (2012). Does source memory impairment associated with emotionally arousing stimuli occur at encoding or retrieval? *Journal of Cognitive Psychology*, *24*(7), 796–801. https://doi.org/10.1080/20445911.2012.693468

*Otani, H., Libkuman, T. M., Goernert, P. N., Kato, K., Migita, M., Freehafer, S. E., & Landow, M. P. (2012). Emotion, directed forgetting, and source memory. *British Journal of Psychology*, *103*(3), 343–358. https://doi.org/10.1111/j.2044-8295.2011.02078.x

Paivio, A., Rogers, T. B., & Smythe, P. C. (1968). Why are pictures easier to recall than words? *Psychonomic Science*, *11*(4), 137–138. https://doi.org/10.3758/BF03331011

*Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2019). Differential effects of valence and encoding strategy on internal source memory and judgments of source: Exploring the production and the self-reference effect. *Frontiers in Psychology*, *10*, 1326. https://doi.org/10.3389/fpsyg.2019.01326

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2021a). Is internal source memory recognition modulated by emotional encoding contexts? *Psychological Research*, *85*(3), 958–979. https://doi.org/10.1007/s00426-020-01294-4

*Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2021b). Interactions of emotion and self-reference in source memory: An ERP study. *Cognitive, Affective & Behavioral Neuroscience*, *21*(1), 172–190. https://doi.org/10.3758/s13415-020-00858-6

Pereira, D. R., Teixeira-Santos, A. C., Sampaio, A., & Pinheiro, A. P. (2022). *Examining the effects of emotional valence and arousal on source memory: A meta-analysis of behavioral evidence*. https://doi.org/10.17605/OSF.IO/F8S5T

Petrolini, V., & Viola, M. (2020). Core affect dynamics: Arousal as a modulator of valence. *Review of Philosophy and Psychology*, *11*(4), 783–801. https://doi.org/10.1007/s13164-020-00474-w

Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*, *90*(1), 24–46. https://doi.org/10.3102/0034654319877153

Pinheiro, A. P., Rezaii, N., Rauber, A., & Niznikiewicz, M. (2016). Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia. *Cognitive Neuropsychiatry*, *21*(4), 335–353. https://doi.org/10.1080/13546805.2016.1208611

Quintana, D. S. (2015). From pre-registration to publication: A non-technical primer for conducting a meta-analysis to synthesize correlational data. *Frontiers in Psychology*, *6*, 1549. https://doi.org/10.3389/fpsyg.2015.01549

Raye, C. L., & Johnson, M. K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of the Psychonomic Society*, *15*(6), 405–408. https://doi.org/10.3758/BF03334572

*Rimmele, U., Davachi, L., Petrov, R., Dougal, S., & Phelps, E. A. (2011). Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details. *Emotion*, *11*(3), 553–562. https://doi.org/10.1037/a0024246

*Rimmele, U., Davachi, L., & Phelps, E. A. (2012). Memory for time and place contributes to enhanced confidence in memories for emotional events. *Emotion*, *12*(4), 834–846. https://doi.org/10.1037/a0028003

*Ritchey, M., Wang, S. F., Yonelinas, A. P., & Ranganath, C. (2019). Dissociable medial temporal pathways for encoding emotional item and context information. *Neuropsychologia*, *124*, 66–78. https://doi.org/10.1016/j.neuropsychologia.2018.12.015

*Robin, F., Cébron, T., Letellier, M., & Nizard, J. (2018). The effects of emotional content on source monitoring in fibromyalgia patients. *The Journal of General Psychology*, *145*(4), 392–414. https://doi.org/10.1080/00221309.2018.1494536

Rohatgi, A. (2021). WebPlotDigitizer (Version 4.5) [Computer software]. https://automeris.io/WebPlotDigitizer/index.html

RStudio Team. (2020). *RStudio: Integrated development for R*. RStudio, Inc. http://www.rstudio.com/

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Sala, G., Aksayli, N. D., Tatlidil, K. S., Gondo, Y., & Gobet, F. (2019). Working memory training does not enhance older adults' cognitive skills: A comprehensive meta-analysis. *Intelligence*, *77*, 101386. https://doi.org/10.1016/j.intell.2019.101386

Schaefer, A., Pottage, C. L., & Rickart, A. J. (2011). Electrophysiological correlates of remembering emotional pictures. *NeuroImage*, *54*(1), 714–724. https://doi.org/10.1016/j.neuroimage.2010.07.030

*Schmidt, K., Patnaik, P., & Kensinger, E. A. (2011). Emotion's influence on memory for spatial and temporal context. *Cognition and Emotion*, *25*(2), 229–243. https://doi.org/10.1080/02699931.2010.483123

Schweizer, S., Satpute, A. B., Atzil, S., Field, A. P., Hitchcock, C., Black, M., Barrett, L. F., & Dalgleish, T. (2019). The impact of affective information on working memory: A pair of meta-analytic reviews of behavioral and neuroimaging evidence. *Psychological Bulletin*, *145*(6), 566–609. https://doi.org/10.1037/bul0000193

Sharot, T., & Phelps, E. A. (2004). How arousal modulates memory: Disentangling the effects of attention and retention. *Cognitive, Affective & Behavioral Neuroscience*, *4*(3), 294–306. https://doi.org/10.3758/CABN.4.3.294

*Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106*(1), 538–547. https://doi.org/10.1016/j.cognition.2007.03.002

Shields, G. S., Sazma, M. A., McCullough, A. M., & Yonelinas, A. P. (2017). The effects of acute stress on episodic memory: A meta-analysis and integrative review. *Psychological Bulletin*, *143*(6), 636–675. https://doi.org/10.1037/bul0000100

Song, F., Hooper, L., & Loke, Y. K. (2013). Publication bias: What is it? How do we measure it? How do we avoid it? *Open Access Journal of Clinical Trials*, *5*(1), 71–81. https://doi.org/10.2147/OAJCT.S34419

Spencer, W. D., & Raz, N. (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging*, *10*(4), 527–539. https://doi.org/10.1037/0882-7974.10.4.527

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *343*(1), d4002. https://doi.org/10.1136/bmj.d4002

Stevens, J. S., & Hamann, S. (2012). Sex differences in brain activation to emotional stimuli: A meta-analysis of neuroimaging studies. *Neuropsychologia*, *50*(7), 1578–1593. https://doi.org/10.1016/j.neuropsychologia.2012.03.011

Subramaniam, K., Luks, T. L., Fisher, M., Simpson, G. V., Nagarajan, S., & Vinogradov, S. (2012). Computerized cognitive training restores neural activity within the reality monitoring network in schizophrenia. *Neuron*, *73*(4), 842–853. https://doi.org/10.1016/j.neuron.2011.12.024

Taneja, P., Olausson, H., Trulsson, M., Svensson, P., & Baad-Hansen, L. (2021). Defining pleasant touch stimuli: A systematic review and meta-analysis. *Psychological Research*, *85*(1), 20–35. https://doi.org/10.1007/s00426-019-01253-8

Teixeira-Santos, A. C., Moreira, C. S., Magalhães, R., Magalhães, C., Pereira, D. R., Leite, J., Carvalho, S., & Sampaio, A. (2019). Reviewing working memory training gains in healthy older adults: A meta-analytic review of transfer for cognitive outcomes. *Neuroscience and Biobehavioral Reviews*, *103*, 163–177. https://doi.org/10.1016/j.neubiorev.2019.05.009

Thompson, T., Oram, C., Correll, C. U., Tsermentseli, S., & Stubbs, B. (2017). Analgesic effects of alcohol: A systematic review and meta-analysis of controlled experimental studies in healthy participants. *The Journal of Pain*, *18*(5), 499–510. https://doi.org/10.1016/j.jpain.2016.11.009

Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, *8*, 1454. https://doi.org/10.3389/fpsyg.2017.01454

Unkelbach, C., Alves, H., & Koch, A. (2020). Negativity bias, positivity bias, and valence asymmetries: Explaining the differential processing of positive and negative information. In B. Gawronski (Ed.), *Advances in experimental social psychology* (pp. 115–187). Academic Press. https://doi.org/10.1016/bs.aesp.2020.04.005

Urban, E. J., Charles, S. T., Levine, L. J., & Almeida, D. M. (2018). Depression history and memory bias for specific daily emotions. *PLoS ONE*, *13*(9), e0203574. https://doi.org/10.1371/journal.pone.0203574

van Aert, R. C. M., Wicherts, J. M., & van Assen, M. A. L. M. (2019). Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis. *PLoS ONE*, *14*(4), e0215052. https://doi.org/10.1371/journal.pone.0215052

Ventura-Bort, C., Dolcos, F., Wendt, J., Wirkner, J., Hamm, A. O., & Weymar, M. (2020). Item and source memory for emotional associates is mediated by different retrieval processes. *Neuropsychologia*, *145*, 106606. https://doi.org/10.1016/j.neuropsychologia.2017.12.015

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

Viechtbauer, W., & Cheung, M. W. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Vieitez, L., Haro, J., Ferré, P., Padrón, I., & Fraga, I. (2021). Unraveling the mystery about the negative valence bias: Does arousal account for processing differences in unpleasant words? *Frontiers in Psychology*, *12*, 748726. https://doi.org/10.3389/fpsyg.2021.748726

*Wang, B. (2012). Females' superiority in item memory, but not source memory for neutral and emotional Chinese words. *Personality and Individual Differences*, *52*(8), 925–929. https://doi.org/10.1016/j.paid.2012.02.005

*Wang, B., & Fu, X. (2011). Time course of effects of emotion on item memory and source memory for Chinese words. *Neurobiology of Learning and Memory*, *95*(4), 415–424. https://doi.org/10.1016/j.nlm.2011.02.001

Waring, J. D., & Kensinger, E. A. (2009). Effects of emotional valence and arousal upon memory trade-offs with aging. *Psychology and Aging*, *24*(2), 412–422. https://doi.org/10.1037/a0015526

*Weigl, M., Pham, H. H., Mecklinger, A., & Rosburg, T. (2020). The effect of shared distinctiveness on source memory: An event-related potential study. *Cognitive, Affective & Behavioral Neuroscience*, *20*(5), 1027–1040. https://doi.org/10.3758/s13415-020-00817-1

Weymar, M., Löw, A., & Hamm, A. O. (2011). Emotional memories are resilient to time: Evidence from the parietal ERP old/new effect. *Human Brain Mapping*, *32*(4), 632–640. https://doi.org/10.1002/hbm.21051

Weymar, M., Löw, A., Schwabe, L., & Hamm, A. O. (2010). Brain dynamics associated with recollective experiences of emotional events. *NeuroReport*, *21*(12), 827–831. https://doi.org/10.1097/WNR.0b013e32833d180a

Whitfield, T., Barnhofer, T., Acabchuk, R., Cohen, A., Lee, M., Schlosser, M., Arenaza-Urquijo, E. M., Böttcher, A., Britton, W., Coll-Padros, N., Collette, F., Chételat, G., Dautricourt, S., Demnitz-King, H., Dumais, T., Klimecki, O., Meiberth, D., Moulinet, I., Müller, T., . . . Marchant, N. L. (2022). The effect of mindfulness-based programs on cognitive function in adults: A systematic review and meta-analysis. *Neuropsychology Review*, *32*, 677–702. https://doi.org/10.1007/s11065-021-09519-y

Wirkner, J., Ventura-Bort, C., Schulz, P., Hamm, A. O., & Weymar, M. (2018). Event-related potentials of emotional and neutral memories: The role of encoding position and delayed testing. *Psychophysiology*, *55*(7), e13069. https://doi.org/10.1111/psyp.13069

Wulff, A. N., & Thomas, A. K. (2021). The dynamic and fragile nature of eyewitness memory formation: Considering stress and attention. *Frontiers in Psychology*, *12*, 666724. https://doi.org/10.3389/fpsyg.2021.666724

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *The Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

*Yick, Y. Y., Buratto, L. G., & Schaefer, A. (2015). The effects of negative emotion on encoding-related neural activity predicting item and source recognition. *Neuropsychologia*, *73*, 48–59. https://doi.org/10.1016/j.neuropsychologia.2015.04.030

Yonelinas, A. P., & Ritchey, M. (2015). The slow forgetting of emotional episodic memories: An emotional binding account. *Trends in Cognitive Sciences*, *19*(5), 259–267. https://doi.org/10.1016/j.tics.2015.02.009

*Zhang, Y., Pan, Z., Li, K., & Guo, Y. (2018). Self-Serving bias in memories: Selectively forgetting the connection between negative information and the self. *Experimental Psychology*, *65*(4), 236–244. https://doi.org/10.1027/1618-3169/a000409

*Zhou, W., Nie, A., Xiao, Y., Liu, S., & Deng, C. (2020). Is color source retrieval sensitive to emotion? Electrophysiological evidence from old/new effects. *Acta Psychologica*, *210*, 103156. https://doi.org/10.1016/j.actpsy.2020.103156

Received April 12, 2022
Revision received September 14, 2022
Accepted September 20, 2022 ■