Emotional Source Memory:
(When) Are Emotional Sources Remembered Better?

NIKOLETTA SYMEONIDOU

Inauguraldissertation

zur Erlangung des akademischen Grades einer Doktorin

der Sozialwissenschaften der Universität Mannheim

Tag der Disputation:
15. August 2022

Betreuerinnen:
Prof. Dr. Beatrice G. Kuhlmann
Prof. Dr. Mandy Hütter

Dekan der Fakultät für Sozialwissenschaften:
Prof. Dr. Michael Diehl

*Gutachter*innen:*
Prof. Dr. Edgar Erdfelder
Prof. Dr. Meike Kroneisen

For mom

Для мамы

♡

As knowledge increases, wonder deepens.

– Charles Morgan

Table of Contents

Summary

Emotional information benefits memory. This phenomenon of emotion-enhanced memory (EEM) has been well established for item memory (i.e., memory for central information) but poorly investigated for source memory (i.e., memory for the context of information). Filling this research gap, I examined in the following dissertation whether and under which conditions source memory is better for emotional (versus neutral) sources by focusing on three potential influencing factors: Valence and arousal of sources, aging, and encoding instructions. In all three manuscripts, source stimuli were selected based on normative valence and arousal ratings, thus ensuring an effective emotionality manipulation. Item stimuli were neutral and unrelated to the source material. In all manuscripts, the methodological approach followed the standard source-monitoring paradigm, and analyses of source-monitoring data were based on multinomial modeling.

Manuscript 1 revealed that there is no beneficial effect of source valence or source arousal on source memory. Manuscript 2 indicated that only younger but not older adults show enhanced source memory for emotional (i.e., positive and negative) compared to neutral sources. Thus, Manuscript 2 showed a valence effect in source memory which, however, was absent in Manuscript 1. Clarifying this inconsistent result pattern, Manuscript 3 unveiled that EEM effects in source memory depend on the encoding instructions: EEM effects robustly occur if an affective orienting, item-focused task is used during item-source encoding (as in Manuscript 2) but do not occur if no such orienting task is used (as in Manuscript 1). In sum, the overall results clearly indicate that emotional sources per se are not remembered better. Instead, an affective item-source processing seems crucial for establishing EEM effects in source memory. With this, my thesis identifies important boundary conditions that foster versus hinder EEM effects and thus contributes to a better understanding of how emotion influences episodic memory.

Manuscripts

This dissertation answers the question of whether and when emotional sources benefit source memory by focusing on three different factors: Valence and arousal of sources (Manuscript 1), aging (Manuscript 2), and encoding instructions (Manuscript 3). Manuscript 1 is published, Manuscript 2 is accepted for publication, and Manuscript 3 is submitted for publication in *Cognition and Emotion*. The research conducted in this dissertation has been supported by the Research Training Group "Statistical Modeling in Psychology" (SMiP), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation).

In the main text of this thesis, I provide a brief review of previous research on emotional source memory, describe the methodological approach used in this dissertation, give an overview of the three manuscripts, and conclude by discussing the strengths and weaknesses of this research as well as potential future directions. For specifics about the experimental procedures and statistical analyses used in the manuscripts, please refer to the original manuscripts appended to this thesis.

Manuscript 1

Symeonidou, N., & Kuhlmann, B. G. (2022). Better memory for emotional sources? A systematic evaluation of source valence and arousal in source memory. *Cognition and Emotion*, *36*(2), 300-316. https://doi.org/10.1080/02699931.2021.2008323

Manuscript 2

Symeonidou, N., Hassan, A., Porstein, I. & Kuhlmann, B. G. (in press). Is there an emotionality effect in older adults' source memory? *Aging, Neuropsychology and Cognition*.

Manuscript 3

Symeonidou, N., & Kuhlmann, B. G. (2022). *Enhanced source memory for emotional sources: Does an affective orienting task make the difference?* Manuscript submitted for publication.

1 Introduction

If you take a moment to remember the first time you drove a car or your last day at school, you might be surprised by the vividness of your memory. Emotional memories are almost like snapshots of previous live events, characterized by a high richness of details, vividness, and accuracy (Kensinger, 2009; Kensinger & Schacter, 2008; Rimmele et al., 2011). Their fascinating nature has stimulated a great bulk of research investigating how emotional memories shape our autobiographic memory (Brown & Kulik, 1977), contribute to trauma development (Holmes & Bourne, 2008), bias eyewitness testimonies (Loftus et al., 1987), influence false memories (Pesta et al., 2001) and many more. To systematically investigate the basic mechanisms behind the phenomenon of emotion-enhanced memory (EEM), researchers have used emotional (versus neutral) items (e.g., words or pictures) in their studies and tested whether and why memory is enhanced for these emotional (versus neutral) items (Kang et al., 2014; Libkuman et al., 2004; Phelps, 2004; Talmi & McGarry, 2012). Locating emotional stimuli on two emotionality dimensions, valence (negative versus positive) and arousal (calming versus activating; see circumplex model by Russell, 1980), these studies have identified important cognitive-behavioral and neural-affective mechanisms underlying valence-based and arousal-based EEM effects, respectively (see Kensinger & Schacter, 2008; Mather, 2007; Talmi, 2013 for reviews).

Interestingly, this research has reliably established EEM effects in item memory, that is, memory for centrally presented stimuli (Glisky et al., 1995), such as pictures or words, but neglected to investigate EEM effects in source memory in the same systematic manner. Source memory refers to remembering the contextual details of an experienced event, for example, its location, its time of day, other persons involved, and so on (Johnson et al., 1993; Mitchell & Johnson, 2009). It is so far unclear whether such contextual features are remembered better if they have an emotional value. Put differently, there has been no or very little systematic research on whether and when source memory is enhanced for emotional (versus neutral) source features (but see Bell & Buchner, 2012). This research gap is surprising considering that episodic memories are often marked by such emotional context features, for example, when we remember receiving information from a likable or dislikable person or when we remember walking home in a severe storm or spectacular sunset. Such emotional context features can determine the emotionality of the whole experience, thereby shaping how the event is represented in episodic memory. Thus, a shift in research focus from item emotionality to source emotionality would contribute to a more holistic understanding of how emotion influences episodic memory. Although some researchers have already begun to investigate memory for emotional source features (Bell & Buchner, 2012; May et al., 2005; Ventura-Bort et al., 2017), this research mostly lacks a common methodological ground and has yielded inconsistent results.

Considering all this, the goal of this dissertation was to systematically investigate whether and under which conditions source memory is enhanced for (inherently) emotional compared to neutral sources. All studies encompassed in this dissertation relied on the same methodological approach, which was specifically tailored to the investigation of emotional sources. I will first give a brief overview of the research conducted so far on emotionality and source memory before describing the methodological approach. I will then turn to the central findings of the three manuscripts and conclude by discussing their implications for future research whilst considering the strength and limitations of this dissertation.

2 Emotionality and Source Memory

If you again recall your memory of your first car drive, you might realize that its high vividness is the result of recollecting the small contextual details of this event (e.g., the time of day, the color of the car, your driving instructor). The ability to remember all these (e.g., temporal, spatial, social) details of an experience has been termed source memory (Johnson et al., 1993). While previous research has established a robust emotion-enhanced memory effect for item memory (i.e., memory for central information; e.g., Kensinger, 2007; Talmi, 2013), the research on emotion and source memory has been rather inconclusive. When reviewing this literature, it is important to differentiate between two lines of research (cf., Bell, Buchner, Erdfelder et al., 2012):

1) Studies that investigated the effects of *emotional items* on source memory. These studies used neutral sources and manipulated *item emotionality*.
2) Studies that investigated the effects of *emotional sources* on source memory. These studies used neutral items and manipulated *source emotionality*.

In the following review, I will briefly summarize the findings on item emotionality (research line 1) but mostly focus on source emotionality (research line 2), as this research mainly motivated the rationale of this dissertation.

2.1 Effects of Item Emotionality on Source Memory

Most of the research on emotionality and source memory can be classified under research line 1, that is, it has focused on the effects of emotional (versus neutral) items on source memory for neutral source features. Taken together, the results of these studies suggest that source memory is enhanced for intrinsic source features of emotional items (e.g., the font color of emotional word items, Doerksen & Shimamura, 2001) versus reduced for external source features of emotional items (e.g., frame color of emotional picture items, Boywitt, 2015; see Chiu et al., 2013 for a review of studies). This is in line with the prevalent belief that emotional items draw focused attention. This attentional bias leads to enhanced memory for the emotional item and its central/intrinsic features (i.e., EEM effect) but reduced memory for all other peripheral/external information (so-called emotion-induced memory trade-off; see Kensinger, 2009; Levine & Edelstein, 2009; Mather, 2007). Although there are still findings that are not in line with this influential central-peripheral trade-off account (see

Chiu et al., 2013; Mather & Sutherland, 2011 for overviews), the question of how item emotionality influences source memory (and associative memory more generally) has enjoyed a continuous research interest, leading to comprehensive, nuanced, still to-be-tested accounts on emotion and memory, which go beyond the scope of this dissertation (e.g., Bisby & Burgess, 2013; Mather & Sutherland, 2011). In contrast and somewhat surprisingly, the research on whether and how emotional sources influence source memory (research line 2) is rather sparse and unsystematic.

2.2 Effects of Source Emotionality on Source Memory

The few studies that manipulated source emotionality considerably vary in their research goals and thus in their methodological approach. When reviewing these studies, I will first focus on whether they show an *EEM effect in source memory*, that is, enhanced source memory for emotional over neutral sources (also referred to as *source emotionality effect* in the following). I will then highlight the main differences across these studies and thus derive the rationale behind the three manuscripts which constitute this dissertation.

One of the first who studied emotional sources was the research group around Bell et al. (starting with Buchner et al., 2009; see Bell & Buchner, 2012 for a review). They investigated whether source memory is enhanced for (contextual) behavioral information that signals cheating (versus trustworthy) behavior. For example, Bell and Buchner (2010) presented neutral faces (=items) with descriptions of cheating versus trustworthy behavior (=sources) to participants and instructed them to rate the likability of the face items during encoding (i.e., affective orienting task). Across several studies (Bell & Buchner, 2010, 2011; Bell, Buchner, Erdfelder et al., 2012; Buchner et al., 2009), the authors found that participants' source memory was enhanced for cheating (i.e., socially threatening) compared to trustworthy sources. Extending these results to descriptions of other negative (non-cheating) behavior, the authors later argued that the negative valence of sources and expectancy violation (instead of social threat specifically) underlies these source-memory enhancements (Bell & Buchner, 2010; Bell, Buchner, Kroneisen et al., 2012). However, the unique contribution of (negative) valence versus arousal to source emotionality effects remained an open question.

Emotional sources have also been examined in aging research. For example, focusing on age-related changes in socio-emotional processing, May et al. (2005) and Rahhal et al. (2002) showed that older adults benefit more from emotional sources compared to younger adults. In these studies, the authors used neutral sources and, via instructions, related these neutral sources to the concept of threat (May et al., 2005) or falsehood (Rahhal et al., 2002).

For example, participants were told that food (=items) presented left versus right (=source) were safe versus dangerous (May et al., 2005) or that statements (=items) spoken by voice A versus voice B (=source) were lies or true statements (Rahhal et al., 2002). Results suggested that older adults' source memory was better if the source was tied to an emotional concept (e.g., items' safety) while younger adults did not show such enhancements. In contrast, using sentences (=items) spoken by voices with a neutral or emotional tone (=sources), Davidson et al. (2006) showed that older adults benefit *less* (instead of more) from emotional sources than younger adults. Thus, results are overall inconclusive, and it remains unclear whether and how source emotionality effects differ between older and younger adults.

Finally, emotional sources have been also used in neuropsychological studies to investigate the neural dynamics underlying memory for (neutral) items that occur in such emotional (versus neutral) contexts. Interestingly, these studies often applied perceptual (instead of conceptual) emotional material as sources. For example, Ventura-Bort, Löw, Wendt, Moltó et al. (2016) presented neutral objects superimposed on emotional (positive and negative) or neutral scene pictures and instructed participants to imagine the objects as part of the scene (i.e., mental imagery instructions). Results indicated that source memory was better for emotional compared to neutral source pictures (see also Smith et al., 2004; Smith et al., 2005, for similar procedures and results). In contrast, Schellhaas et al. (2020) presented neutral faces (=items) in different background colors (=sources), which either signaled threat of electric shock or safety. Although neural processes at retrieval differentiated between faces encoded in a threatening versus safe context, participants' source memory did not differ between threat-of-shock versus safe contexts (see also Arnold et al., 2021, for similar behavioral results). Importantly, these neurological studies do not only differ in how source emotionality was manipulated (i.e., emotional pictures versus threat-of-shock instructions), but they also differ in their encoding instructions. While Ventura-Bort, Löw, Wendt, Moltó et al. (and Smith et al., 2004) told participants to imagine a link between items and sources, no such instructions were used in Schellhaas et al. (or Arnold et al., 2021). Notably, such variations in encoding instructions also occur in the above-reviewed behavioral research. For example, Bell and Buchner have typically used an affective orienting task during encoding (e.g., likability ratings, Bell & Buchner, 2010), while May et al. (2005) and Rahhal et al. (2002) used intentional item and source encoding instructions. As encoding instructions can substantially alter how items and sources are linked and stored in memory (e.g., Diana et al., 2008), they might also modulate source emotionality effects. Put simply, encoding instructions might be another important factor that needs to be considered to understand the inconclusive findings on EEM effects in source memory.

This short review already shows the considerable variation across studies in the main substantive focus (e.g., cognitive aging versus neural dynamics), in the material used for the source emotionality manipulation (e.g., conceptually versus perceptually emotional stimuli), and in the instructions used for item-source encoding (e.g., affective orienting tasks versus mental imagery instructions). Building on this, this dissertation systematically tackled these differences as they might explain the diverging results of previous studies. More specifically, the three manuscripts answer the following questions:

Manuscript 1: Is source memory indeed enhanced for emotional sources? How do source valence and source arousal contribute to such enhancement effects?

Manuscript 2: Do source emotionality effects differ between older versus younger adults? That is, do older adults profit more or less from emotional compared to neutral sources?

Manuscript 3: Can encoding instructions influence source emotionality effects? That is, do source emotionality effects occur only when certain types of instructions are used?

Of note, while both emotionality dimensions (source valence and arousal) were manipulated in Manuscript 1, the research questions in the other two manuscripts required a focus on source valence only (holding source arousal constant); see section 4 for more details. Further crucially, to exclude that variations in results are confounded with variations in method, it was first important to set up a joint methodological approach, which was tailored to the investigation of emotional sources. More specifically, all studies relied on the standard experimental paradigm to investigate source monitoring (Johnson et al., 1993; see next section). Additionally, great care was taken to select stimuli for the source emotionality manipulation. Details of this methodological approach are described next.

3 General Methodological Approach

All studies included in this thesis share three main methodological aspects: First, the experimental procedure followed the standard source-monitoring paradigm (Johnson et al., 1993), which is specifically tailored to the investigation of source-monitoring processes. Second, multinomial modeling (Bayen et al., 1996) was used to derive source memory (and item memory) measures that are corrected for guessing bias. Third, perceptually emotional material was used to manipulate source emotionality. The material was carefully selected based on normative valence and arousal ratings.

3.1 The Standard Source-Monitoring Paradigm as General Procedure

Source monitoring encompasses all processes that are involved when we reconstruct the source (i.e., origin) of an experience, including source memory (Johnson et al., 1993; Kuhlmann et al., 2021). That is, to make a source attribution, people do not only rely on the actual recollection of source features but additionally use their general knowledge and beliefs. For example, if you recall your driving-memory again, you might remember regretting that you put on your flip-flops that day. From this, you might then reconstruct that the driving event must have taken place on a warm, sunny day.

These processes of source memory and reconstruction are also at play when we look at source attributions in the experimental setting. In the standard source-monitoring paradigm, which was used in this dissertation, participants first study multiple items that are presented with either one of two (or three) sources. For example, participants might study words (=items) that are paired with one of three pictures (=sources), see Figure 1. Then in the test phase, all studied items plus some new items (i.e., distractors) are presented, and participants are asked to make a source judgment. For example, as can be seen in Figure 1, participants have to decide whether the word was originally presented with the negative, positive, or neutral picture or whether the word is new. Thus, in the standard paradigm, the item information varies from trial to trial, while the sources repeat across trials, meaning that one source is paired with several items (so-called *many-to-few mapping* of items to sources; Glisky et al., 2001; Kuhlmann et al., 2021). Crucially, a correct answer in the source-monitoring test can be based on actual recollection (i.e., source memory) or a lucky guess (i.e., source guessing).

To disentangle memory and guessing processes in this standard paradigm and thus derive separate measures for source memory and source guessing, Bayen et al. (1996) formulated and empirically validated the so-called *two-high-threshold multinomial model of* *source monitoring* (2HTSM). Generally speaking, multinomial models are binary stochastic models for discrete categorical data (Erdfelder et al., 2009). They are often applied in cognitive psychology to dissociate different cognitive processes that lead to the same empirical observation (e.g., correct assignment of an item to its source) by estimating the probability of each of these underlying processes. The 2HTSM builds on the assumption that item and source memory are discrete (all-or-none) processes (Bayen et al., 1996). This assumption is opposed to the view that memory is a graded process and relies on a continuous strength signal (Wixted, 2007). Notably, the discrete versus continuous account result in alternative approaches to model item and source memory (e.g., 2HTSM versus [bivariate] signal detection model; DeCarlo, 2003). In fact, there is an ongoing debate on whether memory can be best described as a discrete or continuous process. However, the current state of evidence suggests that source memory is a discrete, threshold-like process (Zhou et al., 2021), whereas item memory relies on a continuous signal (Kellen et al., 2021; see also Yonelinas, 2002). As the focus of my dissertation is on source memory, the 2HTSM was a reasonable choice for modeling the source-monitoring data.

The 2HTSM assumes that source judgments in the standard source-monitoring paradigm are driven by four processes: Item recognition (parameter D), source memory (parameter d), item old/new guessing (parameter b), and source guessing (parameter g). The source-memory results reported in all three manuscripts refer to the source-memory parameter d. Note that the original model was designed for a paradigm that implements two sources in the study phase (Bayen et al., 1996). As, however, in all studies of this thesis, three source types were used, an extended version of this model was applied for data analysis (Keefe et al., 2002). This extended version is illustrated in Figure 2. The software multiTree (Moshagen, 2010) was used to estimate model parameters based on the aggregated observed response frequencies in the source-monitoring test (aggregated across participants and items). MultiTree was also used to evaluate model fit via maximum likelihood estimation methods.

Further note that the 2HTSM formed the basis for conducting a priori power analyses in all studies of this dissertation. More specifically, differences across source memory parameters d (e.g., the difference between source memory for emotional sources $d_{emotional}$ and source memory for neutral sources $d_{neutral}$) entered the power analysis as effects of interest. Thus, the sample size in each experiment was a priori tailored to reliably detecting source memory differences of a certain size (with $\alpha = .05$ and $1-\beta = .80$). Details on these power analyses can be found in the original manuscripts.

3.2 Selection of Emotional Stimuli

Normed emotional stimuli were used to manipulate source emotionality in all studies of this dissertation. More specifically, building on the above-described neuropsychological studies (e.g., Smith et al., 2004), we opted for perceptually emotional material (e.g., pictures) because it has been shown to be more emotionally charged compared to semantically emotional material (e.g., words; Bayer & Schacht, 2014; Kensinger & Schacter, 2006). With this, we provided a stronger source emotionality manipulation compared to most of the behavioral studies on source emotionality, which typically applied semantically (e.g., Buchner et al., 2009) or conceptually emotional (e.g., May et al., 2005) material. The use of normed material did not only ensure a certain effectiveness of the emotionality manipulation but also allowed for a systematic variation and/or control of the two prevalent emotionality dimensions, valence and arousal. For Experiment 1 of Manuscript 1, sounds were drawn from the *International Affective Digitized Sounds* (IADS) database (Bradley & Lang, 2007) and used as source stimuli (e.g., the sound of a siren, a train, or rock & roll music). For all remaining experiments, pictures drawn from the *Open Affective Standardized Image Set* (OASIS; Kurdi et al., 2017) were used as sources (e.g., pictures of a garbage dump, a car race, or a lake; see Figure 1).

Following the standard source-monitoring paradigm, only a small number of stimuli (usually three stimuli) were selected to function as sources (see Experiment 2 of Manuscript 2 for an exception). For example, the negative source was operationalized via one or two negative pictures. These stimuli were selected based on their original norm ratings. Notably, to ensure that the stimuli were indeed emotionally effective, their original ratings were additionally checked in one of the two following ways: 1) Valence and arousal ratings for the used source stimuli were either post-hoc collected at the end of the respective experiment (i.e., manipulation check) or 2) a pre-study was conducted in which valence and arousal ratings for a reasonable pre-selection of potentially suitable stimuli were collected. Then, based on these pre-study ratings, the final source stimuli for the main study were chosen. Either way, it was ensured that the source stimuli had the intended emotionality in all studies.

Unlike the source material, items (pictures in Experiment 1 of Manuscript 1, words in all remaining experiments) were neutral in valence and low in arousal. Further importantly, sources and items were chosen in such a way that there was no inherent relation between both (e.g., words as items and unrelated pictures as sources). Thus, sources were unlikely to be processed as an intrinsic feature of the item. This ensured that items and sources were clearly distinguishable and emotionality effects on item versus source memory could be separated. Note that this was not always the case in previous studies outlined above, which have often used highly relatable item-source material (e.g., faces [=items] with descriptions of cheating behavior [=sources] can be processed as cheaters; Bell & Buchner, 2010), or instructed participants to process items and sources as a unit (e.g., imagine the object [=item] as part of the scene image [=source]; Ventura-Bort, Löw, Wendt, Moltó et al., 2016). Such material and instructions blur the distinction between item and source (Diana et al., 2008) and thus make it difficult to disentangle source versus item memory effects. Further importantly, incidental source learning was applied in all studies of this dissertation. This means that participants' attention was not explicitly guided towards the sources, and thus EEM effects in source memory could unfold rather spontaneously. Both the use of external (item-unrelated) sources and the use of incidental source learning served the goal of investigating whether emotional sources *per se* (independent of the item) influence source memory.

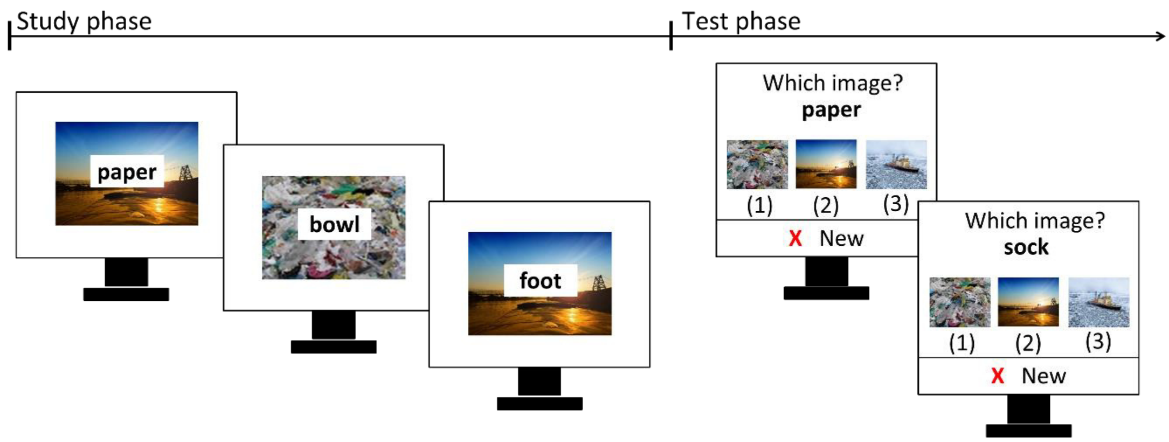


Figure 1. Schematic illustration of the standard source-monitoring paradigm with words used as items and three pictures used as sources (two emotional and one neutral picture, drawn from the *Open Affective Standardized Image Set* [OASIS]; Kurdi et al., 2017). Items vary from trial to trial whereas sources repeat across trials, resulting in a many-to-few mapping of items to sources. Note that this type of item material (i.e., words) and source material (i.e., pictures) was used in Experiment 2 of Manuscript 1 and in both experiments of Manuscript 2 and 3.

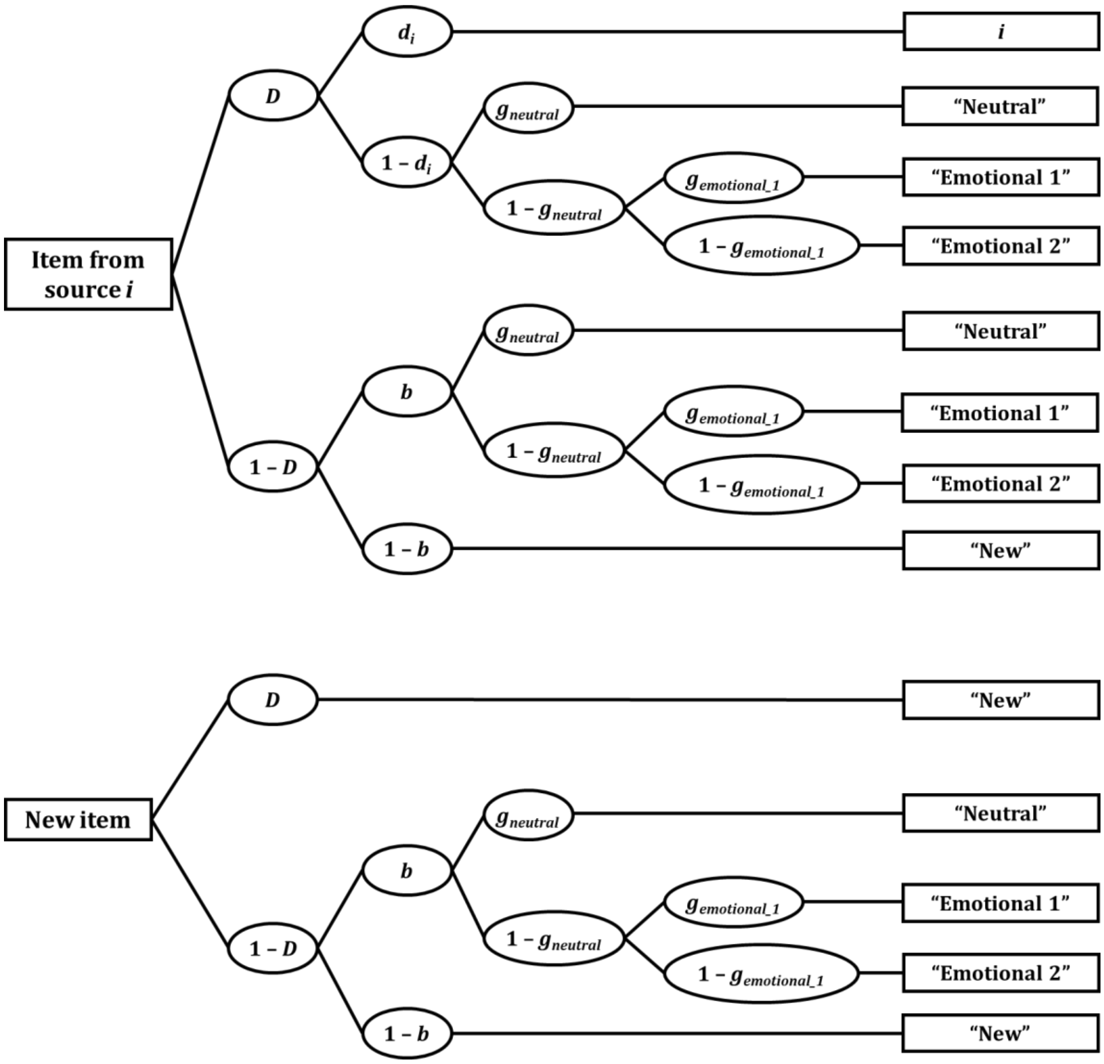


Figure 2. Graphical representation of the *two-high-threshold multinomial model of source monitoring* (2HTSM; Bayen et al., 1996) for three sources, adapted from Keefe et al. (2002). *i* denotes the emotionality of the source with which the item was originally paired. *D* = probability of detecting an item as previously presented or not presented; d_i = probability of correctly recalling the source of a recognized item; *b* = probability of guessing that an item was previously presented; $g_{neutral}$ = probability of guessing the neutral source for a detected or undetected item; $g_{emotional\_1}$ = probability of guessing the first (versus second) emotional source for a detected or undetected item when the neutral source was not guessed. Dependent on the research question, emotional sources varied within participants either in their arousal (Manuscript 1) or their valence (Manuscript 2 and 3).

4 Identifying Influencing Factors of Emotional Source Memory

In three manuscripts, we investigated whether and when source memory is enhanced for emotional sources by focusing on the following factors: Valence and arousal of sources (Manuscript 1), aging (Manuscript 2), and encoding instructions (Manuscript 3). All conducted experiments rely on the above-described methodological approach. In the following, I will briefly outline each manuscript's substantive focus, its methodological specifics, and its main results.

4.1 Manuscript 1: Source Valence Versus Source Arousal

In the first manuscript, we investigated whether source memory is generally enhanced for perceptually (and thus inherently) emotional sources compared to neutral sources and specifically looked at the contribution of valence and arousal to this effect. As reviewed above, research is inconclusive on whether EEM effects occur in source memory, as some studies find such effects (e.g., Bell & Buchner, 2012; Ventura-Bort, Löw, Wendt, Moltó et al., 2016) and others do not (e.g., Arnold et al., 2021; Bell et al., 2017). Notably, the literature on EEM effects in *item memory* emphasizes the importance to separate valence- from arousal-based EEM effects as they seem to rely on different mechanisms (Dolcos et al., 2017; Kensinger & Corkin, 2004). In general, this research suggests that both valence and arousal contribute to EEM effects independently from each other. That is, positive and especially negative items are remembered better compared to neutral items if matched on arousal; and high-arousing items are remembered better than low-arousing items if matched on valence (Kang et al., 2014). Importantly, however, findings additionally suggest that the arousal-based EEM effect might be more robust because it relies on automatic, resource-independent attentional processes (mediated via an amygdala—hippocampus network; Kang et al., 2014; Kensinger & Corkin, 2004; Kern et al., 2005). In contrast, the valence-based EEM effect rather draws on controlled, resource-dependent processes (mediated via a prefrontal-cortex—hippocampus network). As such, it seemed promising to consider and systematically manipulate the valence and arousal of sources as potential factors that contribute to the previous inconclusive findings.

We conducted two experiments in Manuscript 1. In both experiments, we manipulated valence between participants and arousal within participants. That is, we implemented two experimental groups: In the negative group, we used negative sources of high versus low arousal, whereas, in the positive group, we used positive sources of high versus low arousal to manipulate source emotionality. Neutral sources were additionally used in both groups as the baseline. In both experiments, participants were told to learn the items only, without any (explicit) reference to the sources (i.e., incidental source learning). After a three-minute retention interval, participants were presented with all studied items plus some new items and were asked to make a source-monitoring judgment. The two experiments mainly differed in their item and source material: In Experiment 1, participants learned neutral objects as items, which were presented with a high-arousing, low-arousing, or neutral sound as the source. Sounds were selected based on a pre-study conducted with a German student sample. In Experiment 2, participants learned neutral words as items, which were superimposed on emotional or neutral scenery pictures as sources. Crucially, two picture stimuli were used for each valence-arousal combination in this experiment. That is, one source type (e.g., negative high-arousing source) consisted of two pictures. Further crucially, in Experiment 2, we asked participants to rate the valence and arousal of all source pictures at the end of the study. For a more effective source emotionality manipulation, we included only those participants in our main analysis who perceived the source pictures as intended in terms of valence and arousal.

Results were somewhat surprising: Across both experiments, we did not find any beneficial effects of source valence or source arousal on source memory. That is, source memory was not better for high-arousing (versus low-arousing) sources, and also not better for negative or positive (versus neutral) sources. Interestingly, source memory was reduced for negative high- (versus low-) arousing sources in Experiment 1. This might support research showing that high negative arousal has detrimental effects on hippocampus-dependent memory binding and associative memory (Bisby & Burgess, 2017), considering that source memory is a special case of associative memory (Old & Naveh-Benjamin, 2008). However, this detrimental effect of high negative arousal in Experiment 1 did not hold against Bonferroni-Holm adjustment and did not occur in Experiment 2, even when we inspected a subgroup of people with a particularly high rating difference between the high-arousing and low-arousing negative pictures. This rather suggests that the detrimental arousal effect in Experiment 1 was a false positive.

In total, these first two experiments provide conclusive evidence that source memory is not per se enhanced for emotional compared to neutral sources. This suggests that the EEM effects in source memory observed in other studies rely on (methodological) specifics or factors other than valence and arousal. We thus focused on other factors in the next two manuscripts.

4.2 Manuscript 2: Aging

The goal of the second manuscript of this thesis was to investigate whether source emotionality effects differ between older and younger adults. Although the experiments of manuscript 1 suggested no effects of source emotionality on source memory with a young sample, such effects might more readily manifest in older adults (see May et al., 2005; Rahhal et al., 2002). Notably, in this second manuscript, we systematically varied the valence of sources within participants and kept arousal constant at a low-to-medium level. We focused on valence because the literature on emotionality effects in item memory has shown valence-dependent differences between older and younger adults. That is, while younger adults tend to show a negativity bias in item memory (i.e., better memory for negative than for positive items; e.g., Grühn et al., 2007; Spaniol et al., 2008; see Baumeister et al., 2001, for a review), older adults show a positivity bias (i.e., better memory for positive than for negative items) or a reduced negativity bias relative to younger adults (i.e., memory enhancement for negative compared to positive items is weaker in older versus younger adults; e.g., Charles et al., 2003; Kwon et al., 2009). This phenomenon is called the age-related *positivity effect* and has been robustly shown in attention and item memory (see Reed et al., 2014 for a meta-analysis). This effect is theoretically underpinned by the influential *socio-emotional selectivity theory* (SST) of Carstensen (Carstensen et al., 1999), which suggests that, as we age, our motivational priorities shift from future-oriented goals (e.g., knowledge acquisition) to present-oriented goals (e.g., emotional satisfaction). From this perspective, it makes sense that older compared to younger adults more strongly prioritize positive over negative information to maximize their goal of emotional satisfaction. As these effects rely on motivational (goal-related) processes, it is important to keep arousal at a low level because, as discussed above, high-arousing material captures attention automatically and thus counteracts the unfolding of motivational (controlled) processes (see also Kensinger, 2008). Notably, such considerations were lacking in previous studies on source emotionality effects in older versus younger adults, potentially contributing to their inconclusive results.

Across both experiments, the chosen material was similar to Experiment 2 of Manuscript 1 (i.e., neutral words as items superimposed on either emotional or neutral scenery pictures as sources). However, an important difference was that we implemented incidental learning not only for the sources (as in Manuscript 1) but also for the items. This was motivated by Reed et al.'s (2014) meta-analysis, which showed that incidental instructions boost the age-related positivity effect, presumably because an incidental, unconstrained way of processing supports older and younger adults' inherent processing preferences. More specifically, in both experiments, we implemented an affective, item-focused orienting task during item-source encoding: We asked participants to rate the pleasantness of the neutral item, which was presented with either an emotional (positive or negative) or neutral source. A surprise source-monitoring test was administered after a three-minute retention interval. In Experiment 1, we applied the standard many-to-few mapping of items to sources: Neutral items were presented with either the positive, the negative, or the neutral source picture, resulting in a repeated presentation of the three chosen source pictures. To eliminate habituation effects, which might have occurred due to this repeated source presentation, we applied a one-to-one mapping of items to sources in Experiment 2. That is, each item was presented with a unique source picture (of either positive, negative, or neutral valence) during item-source-encoding, meaning that each picture was presented only once.

Results were highly consistent across both experiments: Younger and older participants incorporated source valence into their pleasantness ratings of the neutral items. That is, items paired with positive sources were rated more pleasant than items paired with neutral sources, which in turn were rated more pleasant than items paired with negative sources (i.e., positive > neutral > negative). Of note, an age-related positivity effect additionally occurred in these pleasantness ratings (see Figure 3): In Experiment 1, older adults rated items paired with the negative source as less unpleasant compared to younger adults (i.e., reduced negativity bias); in Experiment 2, older adults rated items more pleasant than younger adults for all three source types. Importantly, this age-related positivity effect in the pleasantness ratings did not transfer to source memory: While younger adults showed better source memory for emotional (and especially positive) compared to neutral sources, indicating an EEM effect in source memory, older adults' source memory did not differ across source types (Figure 4). This suggests that older adults' source memory did not benefit from emotional sources, as younger adults' source memory did, supporting the findings of Davidson et al. (2006), however contradicting those of May et al. (2005) and Rahhal et al. (2002).

Further notably, despite the common methodological ground across the first and second manuscript, their results were somewhat inconsistent: While no valence-based EEM effects in source memory occurred in Manuscript 1, such effects occurred in Manuscript 2 (for younger adults). However, one important methodological difference was that we used different encoding instructions across manuscripts (i.e., participants were told to learn the items in Manuscript 1 versus to rate the pleasantness of the items in Manuscript 2). To further clarify whether these inconsistent results can be explained by the variation in encoding instructions, we focused on the role of encoding instructions in source emotionality effects in the third manuscript.

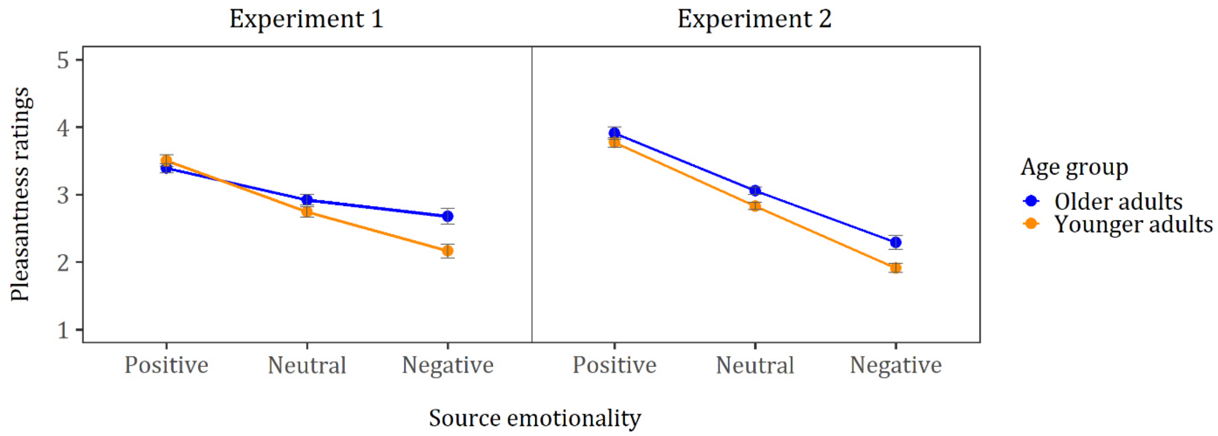


Figure 3. Older and younger adults' pleasantness ratings for neutral words during encoding in Experiment 1 (left-hand plot) and Experiment 2 (right-hand plot) of Manuscript 2. Neutral words were presented with either the positive, neutral or negative source. Error bars indicate one standard error of the mean. The pleasantness rating scale ranged from 1 = very unpleasant to 5 = very pleasant.

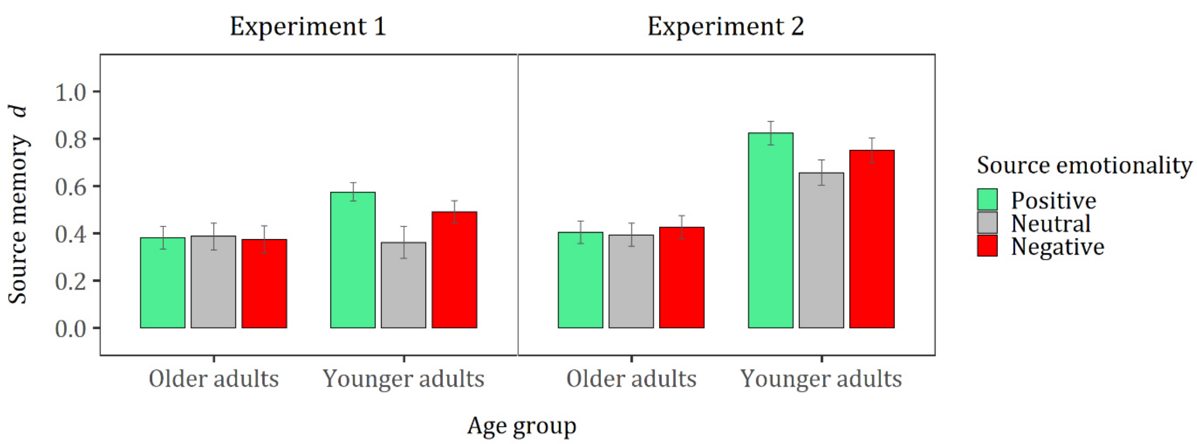


Figure 4. Older and younger adults' source memory for the positive, negative, and neutral sources in Experiment 1 (left-hand plot) and Experiment 2 (right-hand plot) of Manuscript 2. Error bars indicate one standard error of the estimate. Note that *d* = 0 denotes chance performance while *d* = 1 means perfect source memory.

4.3 Manuscript 3: Encoding Instructions

The goal of the third manuscript was to investigate whether the type of encoding instructions influences EEM effects in (younger adults') source memory. With this, we not only aimed at clarifying our own inconclusive results (see previous two manuscripts) but also at clarifying the generally inconsistent findings on source emotionality effects reported in the literature (see *Introduction* section). Interestingly, many studies that reported rather robust EEM effects in source memory applied an affective, item-focused orienting task during item-source encoding (e.g., likeability ratings, see Bell & Buchner, 2012). Combined with the EEM effect found in Manuscript 2, in which we similarly used an affective, item-focused orienting task (i.e., pleasantness ratings), it seems that such instructions might foster EEM effects in source memory. We initially deemed that integrative item-source processing might drive these effects. More specifically, we hypothesized that affective judgments during encoding potentially boost integrative item-source processing because participants can use the sources to inform their judgment about the neutral item. An integrative item-source processing, in turn, benefits source memory, leading to the observed EEM effects. Notably, this idea is further in line with research showing that an EEM effect in source memory can also be established with integrative (non-affective) encoding instructions (Smith et al., 2004; Ventura-Bort, Löw, Wendt, Moltó et al., 2016). Before testing our proposition, we first wanted to replicate the source emotionality effect found in Manuscript 2. Thus, in Experiment 1 of Manuscript 3, we used emotional versus neutral pictures as sources (one per source type), neutral words as items, and an affective orienting task (i.e., pleasantness ratings) for an incidental item-source encoding (cf., Manuscript 2 for more details). We found better source memory for emotional compared to neutral sources, thus replicating the results of Manuscript 2 (for younger adults). In Experiment 2, we aimed at systematically testing under which encoding conditions EEM effects in source memory occur. More specifically, we used the same type of material as in Experiment 1 (neutral words as items superimposed on emotional or neutral pictures as sources), but encoding instructions differed across the four implemented conditions: In the affective orienting task (OT) condition, participants judged the pleasantness of the neutral items (cf., Experiment 1); in the integrative OT condition, participants judged how well the item fits to the source; in the non-integrative OT condition, participants indicated whether the item represents something living or something non-living; and finally, in the no-OT condition, participants (intentionally) learned the items without any (explicit) reference to the sources. Note that in the no-OT condition, we applied the same encoding instructions (intentional item and incidental source learning) as in Manuscript 1. Further note that across all four conditions, sources were encoded incidentally. Memory was tested in a standard source-monitoring test briefly after encoding. We expected to replicate the source emotionality effect on source memory found in Experiment 1 (and Manuscript 2) in the affective OT condition. Notably, based on the above reasoning, EEM effects should occur whenever an integrative item-source encoding is fostered. Thus, we also expected to find an EEM effect in the integrative OT condition, which explicitly encouraged integrative processing. No EEM effects in source memory were expected in the non-integrative and no-OT conditions, as these conditions fostered a rather segregated item-source encoding. Replicating our previous results, we found better source memory for the emotional (and especially positive) sources compared to the neutral source in the affective OT condition (see Figure 5). Surprisingly, no EEM effects in source memory could be established in the integrative OT condition. Finally, no EEM effects occurred in the non-integrative condition and the no-OT condition, as expected (see Figure 5). However, in the no-OT condition, source memory was higher for the positive compared to the negative source (significantly) and neutral source (descriptively). This was presumably because participants found it easier to relate the items to the positive (compared to the negative and neutral) source, as their item-source-fit judgments from the integrative OT condition indicated (see *General Discussion*).

Taken together, across two experiments, we robustly found an EEM effect in source memory when using an affective orienting task during item-source encoding. However, no such effect occurred in the integrative OT condition, which explicitly encouraged integrative processing. This contradicts our idea that EEM effects in source memory occur whenever an integrative item-source encoding is fostered. Instead, it seems that there is something special about the affective orienting task. Possibly, the affective orienting task made sources' valence more salient and meaningful during encoding, thus resulting in the observed EEM effect (see *General Discussion* for a more detailed discussion). In sum, Manuscript 3 clearly shows that EEM effects in source memory are fostered by an affective item-source encoding. With this, our research contributes to a better understanding of the conditions that foster versus hinder source emotionality effects.

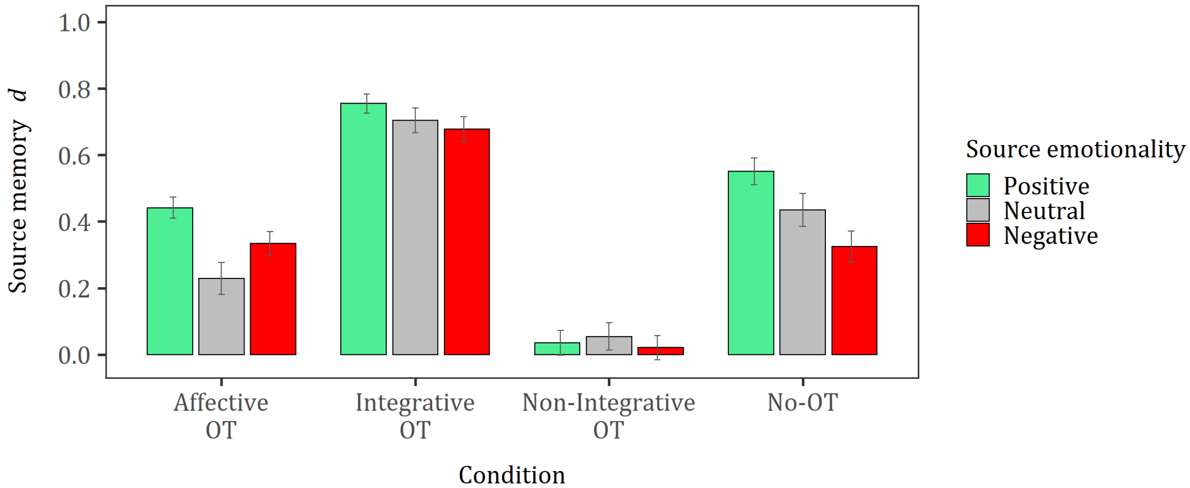


Figure 5. Source memory for the positive, negative, and neutral sources in the affective orienting task (OT), integrative OT, non-integrative OT, and no-OT condition of Experiment 2 of Manuscript 3. Error bars indicate one standard error of estimate. Note that d = 0 denotes chance performance while d = 1 means perfect source memory.

5 General Discussion

This dissertation aimed to systematically investigate whether and under which conditions source memory is enhanced for emotional compared to neutral sources. By shifting the research focus from emotional items to emotional sources, this thesis significantly extends the literature on emotion-enhanced memory (EEM) and thus contributes to a broader view on how emotion influences episodic memory. The developed research program carefully considered previous, inconclusive findings on EEM effects in source memory and, based on this, identified three important factors that might have contributed to such result inconsistencies: valence and arousal of sources, aging, and encoding instructions. Using a joint experimental approach, these factors were systematically investigated in three manuscripts.

In the first manuscript, we investigated whether source memory was enhanced for external, perceptually emotional sources and whether valence versus arousal of sources would independently contribute to this (potential) memory benefit. To ensure a natural, unforced source processing, we applied incidental instructions for source learning but intentional instructions for item learning. Somewhat surprisingly, no beneficial effects of source valence or source arousal on source memory could be established across the two experiments. These findings suggest that emotional sources are not "by default" remembered better and that additional factors might be necessary to promote EEM effects.

In the second manuscript, we investigated whether EEM effects in source memory differ between younger and older adults (see May et al., 2005). We further tested whether older adults specifically benefit from positive compared to negative sources, akin to the seemingly robust age-related positivity effect found in item memory (Reed et al., 2014). We used an affective, item-focused orienting task (i.e., item-pleasantness ratings) to ensure incidental item and source learning. Contrary to our expectations, older adults' source memory did not benefit from emotional (or specifically positive) sources. In contrast, younger adults showed better source memory for emotional compared to neutral sources, indicating a valence-based EEM effect in source memory. Combining the results of Manuscripts 1 and 2, it seemed that the presence versus absence of EEM effects in source memory partially depended on the type of instructions used during item-source encoding. This idea was investigated in the third manuscript.

In Manuscript 3, we first successfully replicated the EEM effect in younger adults' source memory observed in the second manuscript, thus verifying the robustness of this effect. In a second experiment, we systematically varied the type of item-source encoding by either applying an affective orienting task (OT; as in Manuscript 2), an integrative but non-affective OT (new), a non-integrative OT (new), or intentional item encoding instructions (no-OT condition; as in Manuscript 1). Source memory was enhanced for emotional compared to neutral sources in the affective orienting task condition only, emphasizing the importance of affective encoding instructions for source emotionality effects. Interestingly, in the no-OT condition, source memory was higher for the positive compared to the negative source, which was presumably driven by the higher relatedness of the neutral items to the positive source. Manuscript 3 illustrates that spontaneous EEM effects in source memory seem to occur only when source emotionality is salient and meaningful during item-source processing.

Altogether, this thesis contributes to clarifying previous inconsistent results on EEM effects in source memory by specifying conditions under which such EEM effects are present versus absent. On a broader level, the thesis shows that source emotionality per se does not benefit source memory. This implies that the robust EEM effect found for item memory does not simply transfer to source memory, thus underpinning the theoretical distinction of (recollection-based) source memory from (familiarity-supported) item memory (Kuhlmann et al., 2021; Mitchell & Johnson, 2009).

5.1 Strengths and Limitations

This dissertation stands out from previous research particularly due to its methodological soundness. In all studies, the standard source-monitoring paradigm and the two-high-threshold multinomial model of source monitoring (2HTSM) were applied, thus following a well-established approach to investigate and measure source memory (Bayen et al., 1996). Also, to achieve good statistical power, the sample size was always determined via a priori power analyses. Estimates for the effect of interest (i.e., the difference in source memory parameters *d* of the 2HTSM) were carefully derived based on the current state of evidence. Further crucially, perceptually emotional stimuli were used as sources to ensure an effective emotionality manipulation. Additionally, great efforts were undertaken when selecting the source material: Valence and arousal were systematically considered by either varying both independent of each other (Manuscript 1) or by varying valence and keeping arousal constant (Manuscripts 2 and 3). Crucially, the emotionality of the selected source material was additionally checked by collecting valence and arousal ratings either post-hoc (after the experiment) or a priory (in a pre-study). To the best of my knowledge, no previous work has done this in such a careful and thorough manner. Another shortcoming of previous studies was that they used material or instructions which facilitated the processing of items and sources as one joint unit (e.g., Bell & Buchner, 2010; Ventura-Bort, Löw, Wendt, Moltó et al., 2016).

This considerably blurs the distinction between source and item memory. To avoid such a confound, the sources used in this dissertation were external and unrelated to the items. In addition, source learning was always incidental, and instructions never put focus on the item-source relation (with the exception of the integrative OT condition in Manuscript 3, which was deliberate and aimed at investigating the impact of such relational item-source encoding). At large, the experimental approach used in this dissertation was specifically tailored to the investigation of emotional source memory and future studies could continue to use this approach to add to this research line.

Having said that, it is also important to acknowledge the limitations of this dissertation. One limitation was that only a small number of stimuli were used for the source emotionality manipulation. More specifically, following the standard approach of many-to-few mapping of items to sources (i.e., one source presents many items; Glisky et al., 1995), each source type (e.g., negative source) typically consisted of one stimulus (e.g., one negative picture). As this implied a repeated presentation of sources across study trials, participants might have habituated to the emotional material. Another problem with using only one stimulus per source is the resulting confound between emotionality and the specific content of the respective stimulus. For example, in Experiment 1 of Manuscript 2, it is unclear whether the enhanced source memory for the positive source picture was due to its positive valence or due to its specific content (i.e., depicted lake). However, these problems were considered and addressed within each manuscript. That is, in Manuscript 1, we used two pictures for each source type in Experiment 2 to reduce habituation and counteract stimulus-specific, idiosyncratic effects. Similarly, in Manuscript 2, we opted for a one-to-one mapping of items to sources in Experiment 2, presenting each item with a different, unique source picture, thus eliminating the risk of habituation and stimulus-specific effects. In Manuscript 3, we observed EEM effects in source memory only when using an affective orienting task during encoding. As habituation (or stimulus-specific) effects should have been similarly pronounced in all conditions, they cannot sufficiently account for the observed condition-sensitive result patterns. To conclude, although habituation effects and stimulus-specific effects are generally valid concerns, such effects do not constrain the findings and conclusions of this dissertation. Future research on emotional source memory could consider to consistently use a one-to-one mapping of items to sources (i.e., pairing each item with a unique source), as the standard many-to-few mapping approach comes with the risk of habituation effects. Note, however, that having unique sources complicates the differentiation between what is the source and what is the item, as a typical feature of sources is their recurring nature (Kuhlmann et al., 2021). Thus, a one-to-one mapping might rather tap into item-to-item binding instead of item-to-source binding, and emotionality effects potentially differ between these two binding types (see next section). Another way to reduce habituation effects but keep a many-to-few mapping procedure is to use emotional source themes (e.g., pollution as the negative theme), which contain several (e.g., two or three) stimuli per theme, as was done in Manuscript 1, Experiment 2. In this case, the source theme is recurring and thus not unique (tapping into item-source binding), but at the same time consists of several instead of only one stimulus (reducing habituation effects).

5.2 Future Directions

While fostering the understanding of EEM effects in source memory, the findings of this dissertation also prompt new questions and highlight potential future directions. In the following, I will discuss some of these current research gaps, first with regard to source arousal, and then with regard to source valence.

5.2.1 Source Arousal

Although effects of source arousal were investigated in both studies of Manuscript 1, these studies did not use an affective orienting task during item-source-encoding but rather instructed participants to memorize the items only. For source valence, we now know that such affective encoding instructions can foster EEM effects in source memory (see Manuscript 3). Future studies could investigate whether high-arousing sources similarly affect source memory if an affective orienting task is used during encoding. It is noteworthy, however, that the investigation of arousal effects comes with two major challenges. From a theoretical perceptive, the empirical evidence on how arousal affects associative memory binding, including item-to-source binding (i.e., source memory), has been highly inconclusive (see Bisby et al., 2016; Cook et al., 2007; Pierce & Kensinger, 2011, for detrimental effects; Doerksen & Shimamura, 2001; Guillet & Arndt, 2009; Nadarevic, 2017, for beneficial effects; and Meyer et al., 2015; Naveh-Benjamin et al., 2012, for null-effects), prompting an ongoing debate and comprehensive accounts on this issue (e.g., Bisby & Burgess, 2017; Chiu et al., 2013; Levine & Edelstein, 2009; Mather, 2007; Mather & Sutherland, 2011). These accounts mostly rely on studies that have used high-arousing items (not sources) to investigate the impact of arousal on associative binding, but some accounts also enable predictions for the effects of high-arousing sources. One such account is the dual-representation theory by Bisby and colleagues (Bisby et al., 2016; Bisby & Burgess, 2017). It suggests that (negative) arousal benefits amygdala-dependent memory representations, such as memory for emotional items, but disrupts hippocampally-dependent memory representations, such as memory for associations. As memory for item-source associations (i.e., source memory) has been shown to rely on hippocampal activity (Mitchell & Johnson, 2009), this account would predict reduced (instead of enhanced) source memory for (negative) high-arousing compared to low-arousing sources. Although the studies of Manuscript 1 did not support this prediction, more research is needed to identify potential boundary conditions for the occurrence of such disruptive effects. For example, Bisby et al. typically present two pictures (each being either emotional or neutral) and use associative imagery instructions during encoding, (e.g., asking participants to create a mental image that includes all to-be-bound elements; Bisby et al., 2018). Such instructions foster the binding of separate elements into a coherent memory representation. Put differently, disruptive effects of arousal on the hippocampus and thus on associative memory might become apparent only when binding processes are explicitly encouraged via instructions, which was not the case in Manuscript 1, as instructions focused only on the item, not on item-source binding.

Note, however, that such imagery instructions have also been successfully used to facilitate item-source-unitization (Diana et al., 2008; Murray & Kensinger, 2012), fostering the representation of separate elements as one bound unit in memory (instead of distinct, related elements). Such bound units, in turn, have been shown to rely *less* on the hippocampus (Diana et al., 2007; Murray & Kensinger, 2013) and thus should be less affected by hippocampal disruptions caused by (negative) arousal. Considering this, the question arises why imagery instructions sometimes seem to foster hippocampus-dependent associative binding (as in Bisby & Burgess, 2017) and sometimes lead to hippocampus-independent unitization (as in Diana et al., 2008). This might partially depend on the type of binding. Unitization might be easier to induce for item-to-source (compared to item-to-item) associations because items and sources are often perceptually or semantically linked in source-monitoring experiments (e.g., font color [source] of words [items], Doerksen & Schimamura, 2001; location [source] signals safety of food [item], May et al., 2005). Future experiments could test whether unitization difficulty and success systematically vary across item-source versus item-item associations by applying both behavioral (Murray & Kensinger, 2012) and neurological measures (Diana et al., 2007) of unitization. In sum, future research on arousal and binding needs to take into account that arousal effects might depend on the type of binding (item-source versus item-item), the type of encoding instructions (e.g., incidental versus mental imagery), or an interaction between both.

Further notably, Bisby et al. (2016) do not specify whether hippocampal disruptions rely on high arousal, negative valence, or a combination of both (see also Bisby & Burgess, 2013). They use these terms interchangeably and typically contrast (associative) memory for negative high-arousing stimuli against neutral low-arousing stimuli in their studies, thus confounding the effects of high arousal and negative valence. This leads us to the second, methodological challenge associated with investigating the effects of arousal on (associative) memory. In many established normative databases of perceptually emotional stimuli (e.g., International Affective Picture System [IAPS], Lang et al., 2008; IADS, Bradley & Lang, 2007; Geneva affective picture database [GAPED], Dan-Glauser & Scherer, 2011), the relation between valence and arousal ratings typically follows an asymmetrical *V*-shape (see also Kurdi et al., 2017). Put simply, there are no (or very few) negative and positive stimuli with low arousal levels and no (or very few) neutral stimuli with high arousal levels. This makes it difficult to investigate arousal effects independent of valence. Although Russell (1980) assumed independence of arousal and valence in his pioneering work of the circumplex model, there are different views on what the relation between valence and arousal might look like. In a comprehensive analysis, Kuppens et al. (2013) confirmed that the empirical relation seems to follow an asymmetrical *V*-shape. However, as there are large individual differences in the shape of this relation, the authors conclude that the *V*-shaped relation is weak and that "(…) affective experiences of all combinations of valence and arousal can occur (e.g., low arousal but highly positive or negative affect states do occur, although less frequently)" (p. 933). Building on this, Kurdi et al. (2017) have stressed the need to add negative and positive low-arousing stimuli to emotional databases. Of note, such databases would also profit from adding age norms as valence and arousal perception might vary between younger and older adults (Grühn & Scheibe, 2008; Kurdi et al., 2017).

5.2.2 Source Valence

Manuscripts 2 and 3 substantially contribute to clarifying the effects of source valence on source memory: Beneficial effects can be robustly established when an affective, item-focused orienting task (i.e., item-pleasantness ratings) is used during item-source-encoding, suggesting that the valence effect is tied to the affective encoding instructions. However, the exact mechanisms remain rather unclear. After careful consideration of the full result pattern, we deem that the affective encoding instructions stand out from the other instructions used in our experiments in two main ways. First, the pleasantness judgments made the valence of the source more salient because pleasantness directly maps onto valence. Second, as the affective task focused on the items only, it did not constrain participants to process the sources in a certain way. Put differently, participants were free to pursue their goals and preferences when processing the sources. They were thus inclined to process the emotional (over the neutral) sources because emotional stimuli are, in general, more salient and goal-relevant than neutral stimuli (Levine & Edelstein, 2009; Mather & Sutherland, 2011). This ultimately resulted in the observed EEM effect. Put simply, the general dominance of emotional over neutral sources more clearly comes through if the experimental instructions increase the salience of the source's emotionality but at the same time put minimal constraints on participants' source processing. This can also explain why EEM effects were absent in the integrative OT condition. Here, participants were strongly constrained to engage in an integrative item-source-processing, thus increasing source memory for all three sources, not only for the emotional ones. However, more research is needed to investigate whether salience of source emotionality and experimental constraints on source processing indeed determine the unfolding of EEM effects.

Relatedly, future studies could examine whether EEM effects in source memory occur if sources are learned intentionally (rather than incidentally). Although intentional encoding instructions prompt participants towards an integrative item-source-processing (cf., integrative OT condition), the occurrence of EEM effects might strongly depend on the strategies participants use (or are instructed to use) during item-source-encoding. For example, EEM effects might more readily manifest if participants use a mediator to connect items to emotional versus neutral sources (e.g., a mental image that contains both item and source; see Ventura-Bort, Löw, Wendt, Moltó et al., 2016). This could be investigated more systematically in future studies by manipulating participants' strategy use (e.g., mediator-based versus spontaneous; cf., Kuhlmann & Touron, 2012, during encoding).

Notably, such strategies could also help in establishing an EEM effect in older adults' source memory. To reiterate, in Manuscript 2, EEM effects in source memory manifested only in younger but not in older adults. This was surprising because source valence affected older adults' word-pleasantness ratings in the expected way (i.e., age-related positivity effect), indicating that they not only considered the sources while processing the items, but they did it in a way that matched their processing preferences. However, this was apparently insufficient to boost their source memory. As noted, we think that older adults potentially need an additional, explicit mediator during encoding (e.g., a sentence or image; see Kuhlmann & Touron, 2012) that links the item to the source. In fact, May et al. (2005) as well as Rahhal et al. (2002), who established EEM effects in older adults' source memory, provide such a mediator by linking item and source via an emotional concept (i.e., source signals threat/safety of the item in May et al.; and truth/falsehood in Rahhal et al.). Thus, strictly speaking, May et al. manipulated the emotionality of the item-source-*link*, not the emotionality of the source itself. Future studies could investigate whether older adults' source memory would benefit from perceptually emotional sources if participants are additionally provided with a mediator for the item-to-source link. Note, however, that such mediators (as used in May et al. and Rahhal et al.) might foster the storage of the item-source pair as one unit (i.e., item-source-unitization), which then blurs the distinction between item and source memory and their underlying processes (Diana et al., 2008). Thus, any observed emotionality effect or positivity effect could then rely on (familiarity-supported) item memory processes (i.e., remembering the emotional item-source-unit) instead of (recollection-based) source memory processes (i.e., remembering the emotional source).

More generally, the literature on age-related emotionality effects in memory, and the positivity effect, in particular, would profit from a more thorough investigation of the conditions that favor or moderate such effects. The meta-analysis of Reed et al. (2014) identified two important moderators of the age-related positivity effect in attention and item memory: the experimental constraints imposed on encoding (the fewer, the stronger the effect) and the age difference between the younger and older sample (the larger, the stronger the effect). However, the studies included in this meta-analysis considerably vary across several other (methodological) factors that might similarly moderate the age-related positivity effect. For example, it is unclear whether the effect differs in size across different types of stimuli (e.g., social stimuli such as faces versus non-social stimuli such as pictures) or different types of memory tests (e.g., recognition versus free recall). Our own review of the literature indicated that the latter factor (i.e., type of memory test) might be a promising moderator. More specifically, the positivity effect seems to manifest more robustly in studies applying a free recall test instead of a recognition test (e.g., Charles et al., 2003; Tomaszczyk et al., 2008). As free recall is more retrieval-demanding than recognition (Riefer & Rouder, 1992; Rouder & Batchelder, 1998), this might suggest that the age-related positivity effect relies on a retrieval rather than a storage advantage. Note, however, that free recall and recognition also put different demands on recollection-based processes, with free recall being fully dependent on recollection while recognition also relies on familiarity (Yonelinas et al., 2001). Yet, the idea of a primarily recollection-based positivity effect is rather disproved by Manuscript 2, as we do not find such an effect in (recollection-dependent) source memory (see also Kapucu et al., 2008). However, any conclusion about the underlying processes of the positivity effect would be premature at this point. Future research should first establish whether the effect is indeed moderated by the type of memory test before investigating underlying processes.

Another result pattern that merits further attention in future research is that in both studies of Manuscript 2 and Manuscript 3, source memory for positive sources was slightly but persistently higher than for negative (and neutral) sources. Interestingly, participants in Experiment 2 of Manuscript 3 rated the item-source fit (integrative condition) higher for the positive compared to the negative and neutral source. This indicates a higher relatedness between the neutral word items and the positive source picture, which potentially facilitated their binding and resulted in the observed higher source memory for the positive source. Of note, this pattern descriptively showed up across four experiments (Manuscript 2 and Manuscript 3). This suggests that the effect is tied to positive valence in general rather than the specific positive picture because different pictures constituted the positive source across experiments. In fact, Ventura-Bort, Löw, Wendt, Dolcos et al. (2016) similarly found that participants reported higher success in imagining neutral objects as part of positive (versus negative or neutral) sceneries, again pointing to a higher relatedness. Future studies could investigate why there is a higher relatedness between neutral and positive stimuli and more systematically test how it affects source memory and associative memory.

Finally, it is worthwhile mentioning that all our studies investigated whether source memory is enhanced for the general emotional tone of a source (e.g., was the source positive, negative, or neutral?). Thus, it remains unclear how *specific* source memory is for emotional sources. That is, future studies could investigate whether people are better at discriminating between three positive (or negative) sources than three neutral sources. A study by Bell, Buchner, Erdfelder et al. (2012) suggests that source memory is only better for the general, emotional category of the source (i.e., cheating versus trustworthy behavior) but not for specific source details (i.e., specific behavior). However, it still remains unclear whether participants would be better at differentiating between negative (or positive) source categories than between neutral source categories. Future studies could, for example, investigate this by applying three negative sources in one condition versus three neutral sources in the other condition and compare participants' source memory across these conditions. If emotion indeed boosts recollection, then participants' source memory should be better when sources are negative rather than neutral. Of note, encoding instructions might again play a crucial role here, as they significantly shape how and to what extent participants engage in relational item-source processing.

5.3 Conclusion

The goal of this thesis was to investigate whether and under which conditions emotional sources are remembered better. Overall, the findings of this research clearly show that the mere presence of emotional sources does *not* enhance source memory. Focusing on three influencing factors (source valence and source arousal, aging, and encoding instructions), I identified important boundary conditions that foster versus hinder EEM effects in source memory. With this, my dissertation significantly contributes to clarifying previous inconsistent results and provides a fruitful basis for future research. When all is said and done, it seems that emotion does not always benefit memory.

6 References

Arnold, N. R., González Cruz, H., Schellhaas, S., & Bublatzky, F. (2021). A multinomial modelling approach to face identity recognition during instructed threat. *Cognition & Emotion*, *35*(7), 1–18. https://doi.org/10.1080/02699931.2021.1951175

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037//1089-2680.5.4.323

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 197–215. https://doi.org/10.1037/0278-7393.22.1.197

Bayer, M., & Schacht, A. (2014). Event-related brain responses to emotional words, pictures, and faces - a cross-domain comparison. *Frontiers in Psychology*, *5*, 1106. https://doi.org/10.3389/fpsyg.2014.01106

Bell, R., & Buchner, A. (2010). Valence modulates source memory for faces. *Memory & Cognition*, *38*(1), 29–41. https://doi.org/10.3758/MC.38.1.29

Bell, R., & Buchner, A. (2011). Source memory for faces is determined by their emotional evaluation. *Emotion (Washington, D.C.)*, *11*(2), 249–261. https://doi.org/10.1037/a0022597

Bell, R., & Buchner, A. (2012). How adaptive is memory for cheaters? *Current Directions in Psychological Science*, *21*(6), 403–408. https://doi.org/10.1177/0963721412458525

Bell, R., Buchner, A., Erdfelder, E., Giang, T., Schain, C., & Riether, N. (2012). How specific is source memory for faces of cheaters? Evidence for categorical emotional tagging. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(2), 457–472. https://doi.org/10.1037/a0026017

Bell, R., Buchner, A., Kroneisen, M., & Giang, T. (2012). On the flexibility of social source memory: A test of the emotional incongruity hypothesis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(6), 1512–1529. https://doi.org/10.1037/a0028219

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

Bisby, J. A., & Burgess, N. (2013). Negative affect impairs associative memory but not item memory. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *21*(1), 21–27. https://doi.org/10.1101/lm.032409.113

Bisby, J. A., & Burgess, N. (2017). Differential effects of negative emotion on memory for items and associations, and their relationship to intrusive imagery. *Current Opinion in Behavioral Sciences*, *17*, 124–132. https://doi.org/10.1016/j.cobeha.2017.07.012

Bisby, J. A., Horner, A. J., Bush, D., & Burgess, N. (2018). Negative emotional content disrupts the coherence of episodic memories. *Journal of Experimental Psychology: General*, *147*(2), 243–256. https://doi.org/10.1037/xge0000356

Bisby, J. A., Horner, A. J., Hørlyck, L. D., & Burgess, N. (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. *Social Cognitive and Affective Neuroscience*, *11*(6), 981–990. https://doi.org/10.1093/scan/nsw028

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

Bradley, M. M., & Lang, P. J. (2007). *The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual. Technical report B-3.* University of Florida.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*(1), 73–99. https://doi.org/10.1016/0010-0277(77)90018-X

Buchner, A., Bell, R., Mehl, B., & Musch, J. (2009). No enhanced recognition memory, but better source memory for faces of cheaters. *Evolution and Human Behavior*, *30*(3), 212–224. https://doi.org/10.1016/j.evolhumbehav.2009.01.004

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously. A theory of socioemotional selectivity. *The American Psychologist*, *54*(3), 165–181. https://doi.org/10.1037//0003-066X.54.3.165

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

Chiu, Y.-C., Dolcos, F., Gonsalves, B. D., & Cohen, N. J. (2013). On opposing effects of emotion on contextual or relational memory. *Frontiers in Psychology*, *4*, 103. https://doi.org/10.3389/fpsyg.2013.00103

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

Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance. *Behavior Research Methods*, *43*(2), 468–477. https://doi.org/10.3758/s13428-011-0064-1

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

DeCarlo, L. T. (2003). Source monitoring and multivariate signal detection theory, with a model for selection. *Journal of Mathematical Psychology*, *47*(3), 292–303. https://doi.org/10.1016/S0022-2496(03)00005-1

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 730–740. https://doi.org/10.1037/0278-7393.34.4.730

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion (Washington, D.C.)*, *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

Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models. *Zeitschrift Für Psychologie / Journal of Psychology*, *217*(3), 108–124. https://doi.org/10.1027/0044-3409.217.3.108

Glisky, E. L., Polster, M. R., & Routhieaux, B. C. (1995). Double dissociation between item and source memory. *Neuropsychology*, *9*(2), 229–235. https://doi.org/10.1037/0894-4105.9.2.229

Glisky, E. L., Rubin, S. R., & Davidson, P. S. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1131–1146. https://doi.org/10.1037//0278-7393.27.5.1131

Grühn, D., & Scheibe, S. (2008). Age-related differences in valence and arousal ratings of pictures from the International Affective Picture System (IAPS): Do ratings become more extreme with age? *Behavior Research Methods*, *40*(2), 512–521. https://doi.org/10.3758/BRM.40.2.512

Grühn, D., Scheibe, S., & Baltes, P. B. (2007). Reduced negativity effect in older adults' memory for emotional pictures: The heterogeneity-homogeneity list paradigm. *Psychology and Aging*, *22*(3), 644–649. https://doi.org/10.1037/0882-7974.22.3.644

Guillet, R., & Arndt, J. (2009). Taboo words: The effect of emotion on memory for peripheral information. *Memory & Cognition*, *37*(6), 866–879. https://doi.org/10.3758/MC.37.6.866

Holmes, E. A., & Bourne, C. (2008). Inducing and modulating intrusive emotional memories: A review of the trauma film paradigm. *Acta Psychologica*, *127*(3), 553–566. https://doi.org/10.1016/j.actpsy.2007.11.002

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

Kang, C., Wang, Z., Surina, A., & Lü, W. (2014). Immediate emotion-enhanced memory dependent on arousal and valence: The role of automatic and controlled processing. *Acta Psychologica*, *150*, 153–160. https://doi.org/10.1016/j.actpsy.2014.05.008

Kapucu, A., Rotello, C. M., Ready, R. E., & Seidl, K. N. (2008). Response bias in "remembering" emotional stimuli: A new perspective on age differences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 703–711. https://doi.org/10.1037/0278-7393.34.3.703

Keefe, R. S.E., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source-monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources. *Schizophrenia Research*, *57*(1), 51–67. https://doi.org/10.1016/S0920-9964(01)00306-1

Kellen, D., Winiger, S., Dunn, J. C., & Singmann, H. (2021). Testing the foundations of signal detection theory in recognition memory. *Psychological Review*, *128*(6), 1022–1050. https://doi.org/10.1037/rev0000288.

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

Kensinger, E. A. (2008). Age differences in memory for arousing and nonarousing emotional words. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *63*(1), P13-18. https://doi.org/10.1093/geronb/63.1.P13

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. (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., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(2), 110–126. https://doi.org/10.3758/CABN.6.2.110

Kensinger, E. A., & Schacter, D. L. (2008). Memory and emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 601–617). The Guilford Press.

Kern, R. P., Libkuman, T. M., Otani, H., & Holmes, K. (2005). Emotional stimuli, divided attention, and memory. *Emotion (Washington, D.C.)*, *5*(4), 408–417. https://doi.org/10.1037/1528-3542.5.4.408

Kuhlmann, B. G., Symeonidou, N., Tanyas, H., & Wulff, L. (2021). Remembering and reconstructing episodic context: An overview of source monitoring methods and behavioral findings. In K. D. Federmeier & L. Sahakyan (Eds.), *Psychology of Learning and Motivation. The Context of Cognition: Emerging Perspectives* (Vol. 75, pp. 79–124). Academic Press. https://doi.org/10.1016/bs.plm.2021.06.002

Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(5), 1352–1364. https://doi.org/10.1037/a0027863

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940. https://doi.org/10.1037/a0030811

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods*, *49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Kwon, Y., Scheibe, S., Samanez-Larkin, G. R., Tsai, J. L., & Carstensen, L. L. (2009). Replicating the positivity effect in picture memory in Koreans: Evidence for cross-cultural generalizability. *Psychology and Aging*, *24*(3), 748–754. https://doi.org/10.1037/a0016054

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8.* University of Florida.

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

Libkuman, T. M., Stabler, C. L., & Otani, H. (2004). Arousal, valence, and memory for detail. *Memory (Hove, England)*, *12*(2), 237–247. https://doi.org/10.1080/09658210244000630

Loftus, E. F., Loftus, G. R., & Messo, J. (1987). Some facts about "weapon focus". *Law and Human Behavior*, *11*(1), 55–62. https://doi.org/10.1007/BF01044839

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., & 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

May, C. P., Rahhal, T., Berry, E. M., & Leighton, E. A. (2005). Aging, source memory, and emotion. *Psychology and Aging*, *20*(4), 571–578. https://doi.org/10.1037/0882-7974.20.4.571

Meyer, M. M., Bell, R., & Buchner, A. (2015). Remembering the snake in the grass: Threat enhances recognition but not source memory. *Emotion (Washington, D.C.)*, *15*(6), 721–730. https://doi.org/10.1037/emo0000065

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*(4), 638–677. https://doi.org/10.1037/a0015849

Moshagen, M. (2010). Multitree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, *42*(1), 42–54. https://doi.org/10.3758/BRM.42.1.42

Murray, B. D., & Kensinger, E. A. (2012). The effects of emotion and encoding strategy on associative memory. *Memory & Cognition*, *40*(7), 1056–1069. https://doi.org/10.3758/s13421-012-0215-3

Murray, B. D., & Kensinger, E. A. (2013). A review of the neural and behavioral consequences for unitizing emotional and neutral information. *Frontiers in Behavioral Neuroscience*, *7*, 42. https://doi.org/10.3389/fnbeh.2013.00042

Nadarevic, L. (2017). Emotionally enhanced memory for negatively arousing words: Storage or retrieval advantage? *Cognition & Emotion*, *31*(8), 1557–1570. https://doi.org/10.1080/02699931.2016.1242477

Naveh-Benjamin, M., Maddox, G. B., Jones, P., Old, S., & Kilb, A. (2012). The effects of emotional arousal and gender on the associative memory deficit of older adults. *Memory & Cognition*, *40*(4), 551–566. https://doi.org/10.3758/s13421-011-0169-x

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

Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally charged lures immune to false memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 328–338. https://doi.org/10.1037/0278-7393.27.2.328

Phelps, E. A. (2004). Human emotion and memory: Interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, *14*(2), 198–202. https://doi.org/10.1016/j.conb.2004.03.015

Pierce, B. H., & Kensinger, E. A. (2011). Effects of emotion on associative recognition: Valence and retention interval matter. *Emotion (Washington, D.C.)*, *11*(1), 139–144. https://doi.org/10.1037/a0021287

Rahhal, T. A., May, C. P., & Hasher, L. (2002). Truth and character: Sources that older adults can remember. *Psychological Science*, *13*(2), 101–105. https://doi.org/10.1111/1467-9280.00419

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging*, *29*(1), 1–15. https://doi.org/10.1037/a0035194

Riefer, D. M., & Rouder, J. N. (1992). A multinomial modeling analysis of the mnemonic benefits of bizarre imagery. *Memory & Cognition*, *20*(6), 601–611. https://doi.org/10.3758/BF03202710

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 (Washington, D.C.)*, *11*(3), 553–562. https://doi.org/10.1037/a0024246

Rouder, J. N., & Batchelder, W. H. (1998). Multinomial models for measuring storage and retrieval processes in paired associate learning. In C. E. Dowling, F. S. Roberts, & P. Theuns (Eds.), *Recent progress in mathematical psychology: Psychophysics, knowledge, representation, cognition, and measurement* (pp. 195–225). Psychology Press.

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

Schellhaas, S., Arnold, N., Schmahl, C., & Bublatzky, F. (2020). Contextual source information modulates neural face processing in the absence of conscious recognition: A threat-of-shock study. *Neurobiology of Learning and Memory*, *174*, 107280. https://doi.org/10.1016/j.nlm.2020.107280

Smith, A. P. R., Dolan, R. J., & Rugg, M. D. (2004). Event-related potential correlates of the retrieval of emotional and nonemotional context. *Journal of Cognitive Neuroscience*, *16*(5), 760–775. https://doi.org/10.1162/089892904970816

Smith, A. P. R., Henson, R. N. A., Rugg, M. D., & Dolan, R. J. (2005). Modulation of retrieval processing reflects accuracy of emotional source memory. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *12*(5), 472–479. https://doi.org/10.1101/lm.84305.

Spaniol, J., Voss, A., & Grady, C. L. (2008). Aging and emotional memory: Cognitive mechanisms underlying the positivity effect. *Psychology and Aging*, *23*(4), 859–872. https://doi.org/10.1037/a0014218

Talmi, D. (2013). Enhanced emotional memory. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Talmi, D., & McGarry, L. M. (2012). Accounting for immediate emotional memory enhancement. *Journal of Memory and Language*, *66*(1), 93–108. https://doi.org/10.1016/j.jml.2011.07.009

Tomaszczyk, J. C., Fernandes, M. A., & MacLeod, C. M. (2008). Personal relevance modulates the positivity bias in recall of emotional pictures in older adults. *Psychonomic Bulletin & Review*, *15*(1), 191–196. https://doi.org/10.3758/PBR.15.1.191

Ventura-Bort, C., Dolcos, F., Wendt, J., Wirkner, J., Hamm, A. O., & Weymar, M. (2017). Item and source memory for emotional associates is mediated by different retrieval processes. *Neuropsychologia*, *145*, 1–11. https://doi.org/10.1016/j.neuropsychologia.2017.12.015

Ventura-Bort, C., Löw, A., Wendt, J., Dolcos, F., Hamm, A. O., & Weymar, M. (2016). When neutral turns significant: Brain dynamics of rapidly formed associations between neutral stimuli and emotional contexts. *The European Journal of Neuroscience*, *44*(5), 2176–2183. https://doi.org/10.1111/ejn.13319

Ventura-Bort, C., Löw, A., Wendt, J., Moltó, J., Poy, R., Dolcos, F., Hamm, A. O., & Weymar, M. (2016). Binding neutral information to emotional contexts: Brain dynamics of long-term recognition memory. *Cognitive, Affective & Behavioral Neuroscience*, *16*(2), 234–247. https://doi.org/10.3758/s13415-015-0385-0

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176. https://doi.org/10.1037/0033-295X.114.1.152

Yonelinas, A. P., Hopfinger, J. B., Buonocore, M. H., Kroll, N. E., & Baynes, K. (2001). Hippocampal, parahippocampal and occipital-temporal contributions to associative and item recognition memory: An fMRI study. *Neuroreport*, *12*(2), 359–363. https://doi.org/10.1097/00001756-200102120-00035

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864

Zhou, J., Osth, A. F., Lilburn, S. D., & Smith, P. L. (2021). A circular diffusion model of continuous-outcome source memory retrieval: Contrasting continuous and threshold accounts. *Psychonomic Bulletin & Review*, *28*(4), 1112–1130. https://doi.org/10.3758/s13423-020-01862-0

You can do what you want,

but you cannot want what you want.

– From the series *Dark*,

inspired by Arthur Schopenhauer

A Acknowledgments

This thesis only marks the end of a four-year journey. During this journey, I got to know some unbelievably smart, kind, and simply amazing people. I am grateful to every single one of you for supporting, encouraging, and comforting me along the way. I would like to especially thank the people who accompanied me most persistently:

• Bea, my mentor and supervisor, who guided and supported me from start to finish. No matter how busy your schedule was, you always found time to answer my questions, give me the best advice possible, and resolve my concerns. Thank you for believing in me and my ideas, especially during the most challenging times. Your passionate and unbiased approach to research is remarkable and inspiring. Thank you for passing on these values and skills to me. I am incredibly proud of our joint work and excited about our future accomplishments.

• Lili, who was a second mentor to me from the very first day. You shared every snippet of code, every presentation slide, every piece of experience with me. Your humble but brilliant mind has inspired me as a researcher. Your friendship continues to shape me as a person every single day. I am so grateful for every minute we spent together – in and out of academia.

• Edgar, who was the first to see a researcher in me. I owe my fascination for cognitive research to you. I feel honored to have had you as a supervisor throughout all these years, from my Bachelor's studies until my Ph.D.

• Every single member of the SMiP group, the PIs, the Ph.D. students, the coordinating center, and every guest researcher. I feel privileged and humbled to be part of a network of the brightest scientists in psychology. My special thanks go to Arndt and Mandy for their ideas and thoughts on my projects during retreats and colloquia. I would also like to thank coach Silke, who boosted my self-confidence and encouraged me to aim high.

• My lovely lab ladies, Hilal and Caro, who are one of the deepest thinkers and most attentive listeners I have ever met. Our "therapy sessions" were truly therapeutic. Thank you.

• All our former and current student assistants, who endured my rather mediocre admin skills and last-minute requests. Thank you.

• Pascal, Felix, and Yury, who spent a huge amount of time teaching me OpenSesame, lab.js, and OpenLab. I'll be forever in your debt.

• Michi and Tine, whose mutual silliness has kept me sane throughout this journey.

- Michi, thank you for showing me how to be brave in the most nerve-racking moments, far beyond the academic context. Thank you for sharing your inexhaustible knowledge with me: food, sports, technology, music – the list is endless (almost like your open tabs).
- Tine, thank you for talking me up during the toughest times by framing them positively. Thank you for raising my awareness of social challenges, small ones (data protection) and big ones (discrimination). Most importantly, спасибо за русскую дружбу.

One of the largest gains from this journey is your friendship.

• My L13 4th (former 3rd) floor mates, Philipp, Vanessa, and Ira, who made sure that I was always well-hydrated with coffee. Thank you for broadening my academic and culinary horizon with your sharp thoughts and ideas on both.

• David, who many times saved my life with his expertise in R and lab.js. Thank you for comforting me with your kindness and infinite optimism. Our pink-red debate is still on, though.

• Kilian, who never got tired of listening to my repetitive stories and complaints about academia (and dating) in our 3-hours video calls. I appreciate your friendship very much.

• Bine, who has been a constant in my life for 9 (!) years throughout my entire academic path. I cannot put into words how grateful I am for all these years of friendship. Thank you for always having my back and accepting the good and not-so-good versions of me every single day, without any judgment. You have a special place in my heart.

• My lovely, big, Greek family, who was a distant yet strong emotional source of support. Even though some of you are far away, you're always close to my heart. I'd like to especially thank my Germany-based Greeks, Kostas and Eva, who always found time for me, να αράζουμε λιγάκι. Thank you for challenging my knowledge of economics and politics and reminding me what is actually important in life.

• Most importantly, my mom, who, with her little means yet unbelievably big heart, has given me the best possible education. Μαμά, σε ευχαριστώ μέσα από τα βάθη της ψυχής μου για την εμπιστοσύνη, την κατανόηση, την υποστήριξη και την αγάπη σου. Χωρίς εσένα δεν θα κατάφερνα όσα έχω καταφέρει. От всей души благодарю я тебя.

Nikoletta Symeonidou

Mannheim, May 2022

♡

B Statement of Originality

1. I hereby declare that the presented doctoral dissertation with the title *Emotional Source Memory: (When) Are Emotional Sources Remembered Better?* is my own work.
2. I did not seek unauthorized assistance of a third party and I have employed no other sources or means except the ones listed. I clearly marked any quotations derived from the works of others.
3. I did not present this doctoral dissertation or parts of it at any other higher education institution in Germany or abroad.
4. I hereby confirm the accuracy of the declaration above.
5. I am aware of the significance of this declaration and the legal consequences in case of untrue or incomplete statements.

I affirm in lieu of oath that the statements above are, to the best of my knowledge, true and complete.

Signature:

Date:

C Co-Authors' Statements

Co-Author: Beatrice G. Kuhlmann

I hereby confirm that the following manuscripts included in the thesis *Emotional Source Memory: (When) Are Emotional Sources Remembered Better?* were primarily conceived and written by Nikoletta Symeonidou, Ph.D. candidate at the University of Mannheim:

Symeonidou, N., & Kuhlmann, B. G. (2022). Better memory for emotional sources? A systematic evaluation of source valence and arousal in source memory. *Cognition and Emotion*, *36*(2), 300-316.

Symeonidou, N., Hassan, A., Porstein, I. & Kuhlmann, B. G. (in press). Is there an emotionality effect in older adults' source memory? *Aging, Neuropsychology and Cognition*.

Symeonidou, N., & Kuhlmann, B. G. (2022). *Enhanced source memory for emotional sources: Does an affective orienting task make the difference?* Manuscript submitted for publication.

I sign this statement to the effect that Nikoletta Symeonidou is credited as the primary source of the ideas and the main author of the three above-listed manuscripts. She derived the research questions for the first and third manuscripts and refined the question in the second manuscript, programmed all experiments, collected the data or monitored the data collection, conducted the data analyses, wrote the first drafts, and was responsible for revising all manuscripts. I derived the research question for the second manuscript, contributed to developing and refining the research questions in the other two manuscripts, performed the statistical analyses for Experiment 1 in the second manuscript, suggested ideas for additional analyses for all three manuscripts, provided recommendations for structuring the manuscripts, and contributed to refining and revising the drafts of all manuscripts.

Prof. Dr. Beatrice G. Kuhlmann

Mannheim, May 2022

Co-Author: Abdolaziz Hassan

I hereby confirm that the following manuscript included in the thesis *Emotional Source Memory: (When) Are Emotional Sources Remembered Better?* was primarily conceived and written by Nikoletta Symeonidou, Ph.D. candidate at the University of Mannheim:

Symeonidou, N., Hassan, A., Porstein, I. & Kuhlmann, B. G. (in press). Is there an emotionality effect in older adults' source memory? *Aging, Neuropsychology and Cognition*.

I sign this statement to the effect that Nikoletta Symeonidou is credited as the primary source of the ideas and the main author of the above-stated manuscript. She programmed the experiments, monitored the data collection for Experiment 2, performed the data analyses, wrote the first draft, and was responsible for revising and improving the manuscript. I contributed to developing the theoretical background, designing and refining Experiment 1 of the manuscript, collecting the data for Experiment 1, and improving the manuscript.

Abdolaziz Hassan
Mannheim, May 2022

Co-Author: Isabel Porstein

I hereby confirm that the following manuscript included in the thesis *Emotional Source Memory: (When) Are Emotional Sources Remembered Better?* was primarily conceived and written by Nikoletta Symeonidou, Ph.D. candidate at the University of Mannheim:

Symeonidou, N., Hassan, A., Porstein, I. & Kuhlmann, B. G. (in press). Is there an emotionality effect in older adults' source memory? *Aging, Neuropsychology and Cognition*.

I sign this statement to the effect that Nikoletta Symeonidou is credited as the primary source of the ideas and the main author of the above-stated manuscript. She programmed the experiments, monitored the data collection for Experiment 2, conducted the data analyses, wrote the first draft, and was responsible for revising and improving the manuscript. I contributed to refining the theoretical background, designing and refining Experiment 2 of the manuscript, and revising the second draft of the manuscript.

Isabel Porstein

Mannheim, May 2022

D Copies of Manuscript

https://doi.org/10.1080/02699931.2021.2008323

Better memory for emotional sources? A systematic evaluation of source valence and arousal in source memory*

Nikoletta Symeonidou and Beatrice G. Kuhlmann

Department of Psychology, University of Mannheim, Mannheim, Germany

ABSTRACT

Emotion-enhanced memory (EEM) describes the robust memory advantage of emotional over non-emotional stimuli. While extensively investigated with emotional items, it is unclear whether the EEM effect extends to source memory for a neutral item's emotional context. In two pre-registered studies, we systematically manipulated source valence (positive, negative) between participants and source arousal (high, low, neutral-low) within participants. In Experiment 1 (lab study, $N = 80$), we used emotional sound sources and presented them together with neutral pictures as items. In Experiment 2 (online study, $N = 172$), we used emotional background pictures with superimposed neutral item words to similarly manipulate source emotionality. Multinomial model-based analysis showed no general effects of valence or arousal on source memory across both experiments. Source memory was impaired for the negative high-arousing source in Experiment 1 but this did not replicate in Experiment 2. Altogether, we conclude that there are no memory-enhancing effects of source emotionality (valence, arousal, or any specific combination thereof) on source memory, dissociating emotionality effects between source and item memory. Additionally, we propose that material-dependent influences carry more weight if the used emotional material is limited in number, as is the case in the standard source-monitoring paradigm employing few sources only.

ARTICLE HISTORY
Received 2 September 2021
Revised 3 November 2021
Accepted 15 November 2021

KEYWORDS
Source memory; emotion-enhanced memory; valence; arousal; multinomial modelling

Do you remember the day you graduated from school? Typically our memory is better for emotional than non-emotional life events not only in terms of subjective vividness of the memory but also in its objective accurateness (Kensinger & Corkin, 2003; Kensinger & Schacter, 2006; Rimmele et al., 2012). This phenomenon is called *emotion-enhanced memory* (EEM) and has been comprehensively investigated and often replicated in *item memory*, that is memory for central information, such as emotional words or pictures (Kensinger, 2009; Mather, 2007; Talmi, 2013). However, it is less clear whether there is an EEM effect in *source memory*, which refers to remembering "the conditions under which a memory is acquired" (Johnson et al., 1993, p. 3). Thus, source memory includes, however is not limited to, memory for the (temporal, spatial, or social) context of a central item. Studies that have investigated emotionality and source memory have mostly focused on the effects of central emotional items on memory for emotionally neutral sources (for reviews, see Chiu et al., 2013; Dolcos et al., 2020; Mather, 2007; Mather & Sutherland, 2011), neglecting that the source can be emotional by itself (Bell & Buchner, 2012). For instance, we might receive (neutral) information in a dangerous context and

CONTACT Nikoletta Symeonidou nsymeoni@mail.uni-mannheim.de Department of Psychology, University of Mannheim, L13, 15, Mannheim D-68161, Germany
Supplemental data for this article can be accessed at https://doi.org/10.1080/02699931.2021.2008323.
*Parts of this research were presented at the 2021 Tagung Experimentell Arbeitender Psychologen (TeaP Meeting of Experimental Psychologists).

this might crucially influence how we remember the whole episode.

The studies that have indeed manipulated source emotionality and investigated its effects on source memory have yielded mixed results, with some finding EEM in source memory (e.g. Bell & Buchner, 2012; Smith et al., 2005; Ventura-Bort et al., 2016) but others not (e.g. Bell et al., 2017; May et al., 2005). However, those studies lack a systematic manipulation of the two most prevalent emotionality dimensions (Russell, 1980), that is, valence (i.e. pleasantness of a stimulus: positive vs. negative) and arousal (i.e. activating nature of a stimulus: calming vs. activating). Furthermore, as explained later in detail, the separate assessment of item and source memory was not ideal in many of these studies, leaving the question open of whether emotional sources per se are remembered better. Considering these limitations, the driving goal of our research was to systematically test if there is indeed a source memory enhancement for inherently emotional, external sources and to investigate the influence of source valence versus source arousal using normed emotional sounds (Experiment 1) and pictures (Experiment 2) as sources.

Emotionality and source memory

Most research on emotion and source memory has focused on how central, emotional items influence memory for non-central, neutral sources. This intense research has established that emotional (especially negative high-arousing) stimuli draw focused attention, which leads to prioritised processing of (and better memory for) the emotional stimulus along with its central/intrinsic aspects, at the expense of all other peripheral/external information (i.e. emotion-induced memory trade-off; see Kensinger, 2009; Levine & Edelstein, 2009; Mather, 2007). This idea is shared by several influential frameworks on emotion and memory (e.g. Kensinger, 2009; Mather, 2007) and, despite its ill-defined concept of centrality versus periphery, it has significantly contributed to explaining the seemingly inconsistent results in this research area. Related to this line of research are studies that have investigated emotionality effects on associative memory and memory binding, for example by manipulating the emotionality of the cue or target item of a to-be-learnt (word or image) pair. The inconsistent results of these studies (see e.g. Bisby et al., 2016; Touryan et al., 2007 for detrimental effects of emotionality on associative memory; and Guillet & Arndt, 2009; Nadarevic, 2017 for beneficial effects) cannot be satisfyingly explained by the central-peripheral memory trade-off and have thus given rise to alternative accounts (e.g. dual-representation account by Bisby et al., 2016; discussed below), triggering an on-going debate on how emotionality affects memory binding and fostering insightful, comprehensive research (see Chiu et al., 2013 for a review).

Acknowledging this, a limitation of this previous research is that it has only considered effects of *item emotionality* and has thus neglected the possibility that the source per se can carry emotional value (i.e. *source emotionality*). Only recently, researchers have begun to investigate memory for emotional sources. In a series of studies, Bell and colleagues have found a quite consistent enhancement of source memory for socially threatening or unpleasant (e.g. cheating, disgusting) sources over pleasant or neutral sources (see Bell & Buchner, 2012). In aging research employing other emotional sources (e.g. safety/danger of a neutral food item; May et al., 2005), however, this EEM effect in source memory was only found for older but *not* younger adults (see also Rahhal et al., 2002). While this may reflect interesting age-related changes in socio-emotional motivation (Carstensen et al., 1999), this goes against a general emotionality effect in source memory. Similarly, Bell et al. (2017) and Meyer et al. (2016) paired emotional items (i.e. snake pictures) with emotional source information (i.e. poisonousness vs. non-poisonousness) in their studies and did not find enhanced source memory for the emotional source feature. Notably, in all these studies the source material was conceptually instead of inherently emotional (i.e. the source's implied meaning elicited the emotionality, not the depicted source per se), which might engage different (i.e. more elaborate) processes and brain regions than the ones typically associated with emotional memory (Dolcos et al., 2017).

Interestingly, inherently emotional contexts (e.g. emotional background pictures) have been used in neuropsychological studies, but these studies often focus on item memory only and do not test participants' source memory for these emotional contexts (e.g. Jaeger & Rugg, 2012). Those that test source memory tend to find enhancement effects (Smith et al., 2004; Smith et al., 2005; Ventura-Bort et al., 2016), but see Bisby and Burgess (2013) for detrimental effects and Schellhaas et al. (2020) for null-effects.

However, most rely on source memory measures that are biased (cf., Bell & Buchner, 2010; but see Schellhaas et al., 2020; Ventura-Bort et al., 2020 for bias-corrected measures). Further, they typically use unitisation instructions during encoding (e.g. telling participants to form an integrated mental image of item and source; Ventura-Bort et al., 2016). When item and source are stored as a unit, it is difficult to measure emotionality effects on source memory independent of item memory (cf. Diana et al., 2008). Notably, this also applies to the socially pleasant/threatening sources studied by Bell and colleagues (reviewed in Bell & Buchner, 2012), in which the emotional source is always informative of the item and could thus be embedded as its intrinsic feature (e.g. face with a cheating description = cheater).

The influence of source valence versus source arousal

As evident in our review, the literature on emotional source memory has employed various different manipulations of (inherently) emotional sources. Considering the mixed results, it seems promising to consider the valence and arousal of these emotional source manipulations as a common ground that may explain the inconsistent results. Among those (few) studies drawing on normed emotional stimuli as sources, most compared negative and positive high-arousing sources to neutral sources thus confounding valence with high arousal (Jaeger & Rugg, 2012; Smith et al., 2005; Ventura-Bort et al., 2016). Some studies have demonstrated effects of source valence when controlling for arousal (Erk et al., 2005; Pereira et al., 2021), but source emotionality effects could not be explained with valence alone (Bell & Buchner, 2011). Further, to the best of our knowledge, there is yet no study, in which arousal of sources was systematically varied while controlling for valence. Notably, studies on emotional item memory typically emphasise the crucial and leading role of arousal in EEM effects as it more reliably leads to memory enhancing effects (Mather, 2007; Mather & Sutherland, 2011). More specifically, studies have shown that high-arousing items are typically remembered better compared to low-arousing items of the same (positive/negative) valence (Kang et al., 2014; Kensinger & Corkin, 2003). However, the same studies show that valence can additionally contribute to EEM effects in item memory, independent of arousal. Furthermore, research suggests that negative items are remembered better than positive items (so-called negativity bias, see Baumeister et al., 2001 for an overview). Accordingly, we argue that a comprehensive investigation of EEM effects in source memory requires a systematic dissociation of both emotionality dimensions. For this purpose, we manipulated source arousal and source valence independent of each other in both of our experiments.

Overview of the present experiments

The primary goal of our research was to systematically investigate whether there is a source memory enhancement for emotional sources and whether source valence versus source arousal independently contribute to this effect. To efficiently cross both emotionality dimensions in our experimental design, we manipulated arousal within and valence between participants in both experiments. Accordingly, one group of participants was presented with negative high- versus low-arousing sources and another group was presented with positive high- versus low-arousing sources. In each group, neutral low-arousing sources served as a baseline condition. To corroborate the generality of our results, we used two different manipulations of source emotionality (Experiment 1: sounds; Experiment 2: background pictures). The general procedure in both experiments followed the standard source-monitoring paradigm. This allowed us to apply the two-high-threshold multinomial model of source monitoring (2HTSM; Bayen et al., 1996) and derive bias-free measures for item, and crucially, source memory. To further ensure the dissociation of item and source memory, we did not instruct intentional source encoding to avoid source-item unitisation.

Based on the preliminary (albeit inconclusive) support of EEM in source memory and the robust EEM effects in item memory outlined above, we generally expected to find better memory for emotional compared to neutral sources and tentatively hypothesised that both, high arousal and negative valence contribute to this effect, independently from each other. This prediction is in line with the idea that people preferably focus their attention on emotional information (here: emotional sources) over neutral information. Admittedly, there are other accounts on emotion and memory with different predictions, which however we became aware of only after conducting our first experiment. Concretely, some researchers have argued that high emotionality disrupts associative binding and underlying neural circuits (Bisby & Burgess, 2017; Chiu et al., 2013; Levine & Edelstein, 2009). For example, in their dual-representation account, Bisby and colleagues (Bisby et al., 2016; Bisby & Burgess, 2013) propose and empirically show that negative high arousal has a beneficial effect on amygdala-dependent memory representations, fostering memory for emotional items, however a detrimental effect on hippocampally-dependent memory representations, disrupting memory for associations. As the hippocampus supports memory for item-source associations (Mitchell & Johnson, 2009), this account would predict reduced source memory for negative high-arousing sources due to their detrimental influence on binding. Another prevalent account on emotional arousal is the arousal-biased competition theory (Mather & Sutherland, 2011), which assumes that high arousal exacerbates the preference to process high priority stimuli (e.g. emotional or task-relevant stimuli) over low priority stimuli. Considering that the sources in our experiment were emotional, however task-irrelevant (i.e. incidental), we cannot derive clear predictions from this account. Rather, as our study most closely resembled the studies on emotional sources (i.e. manipulation of source emotionality rather than item emotionality; e.g. Bell & Buchner, 2012), we oriented ourselves on the results reported there at the time of planning and conducting Experiment 1 (see also preregistration). Crucially, independent of the direction of effect (i.e. memory enhancing versus impairing), our experimental design allows us to test whether arousal, negative valence, or a combination of both contribute to any observed emotionality effect on source memory.

Experiment 1

In our first experiment, we combined neutral pictures as items with a neutral or emotional (high- or low-arousing) sound as source. Depending on the experimental group, the emotional sounds were either of positive valence (i.e. "Positive Group") or negative valence (i.e. "Negative Group").

Method

The experiment was approved by the ethics board of the University of Mannheim and participants were informed about the potentially aversive nature of the used study material upfront (same was true for Experiment 2).

Design and participants

The design of our experiment was a two × three-mixed design with source valence (positive vs. negative) manipulated between participants and source arousal (high vs. low vs. neutral-low) manipulated within participants. Based on an a priori power analysis with $\alpha = .05$ and $1 - \beta = .80$, we aimed at recruiting 38 eligible participants per valence group (i.e. 38 × 92 items = 3496 observations) to allow detection of differences of .10 or larger between source-memory parameters in the expected direction (see online supplement for a more detailed description of the power analysis; power analysis was performed using the software *multiTree*; Moshagen, 2010).

In total, 84 students of the University of Mannheim participated in our study. Four participants were excluded from data analysis because they did not meet pre-defined eligibility requirements (German as native language [i.e. learned before age of six]; aged 18–30 years; no diagnosed depression and/or anxiety disorder within the past 6 months). Thus, final data analysis was based on 80 participants (66 women, $M = 21.91$ years, $SD = 2.70$ years), equally distributed across both valence groups.

Materials

Standardised emotional sounds were drawn from the *International Affective Digitized Sounds* (IADS; Bradley & Lang, 2007) and cut to three seconds such that their main (emotional) character was preserved. To ensure the validity of the arousal and valence ratings of the three-seconds IADS sounds with a German student population, as well as to exclude unrecognisable or unrealistic sounds, we conducted a pre-study with a separate sample of University of Mannheim students ($N = 10$). Based on pre-study valence and arousal ratings, we carefully selected two negative sounds (high- ["siren"] vs. low-arousing ["belch"]) for the Negative Group, and two positive sounds (high- ["RockNRoll"] vs. low- arousing ["applause"]) for the Positive Group, as well as a neutral (low-arousing ["train"]) sound to use in both groups. The valence extremity of the positive and negative sounds as well as their high and low arousal levels were matched, with the neutral sound's arousal matched to low arousal.

Non-emotional scenery or object pictures were taken from the standardised picture-database *Open Affective Standardized Image Set* (OASIS; Kurdi et al., 2017) to serve as items. Based on the original valence and arousal ratings (on a 7-point rating scale), we selected 96 pictures of neutral valence and low arousal. The pictures were grouped into four different sets in advance (3 sets for each of the 3 sources, 1 set for the distractor items) to make sure that thematically related pictures did not appear in the same set (i.e. were not presented with the same source). Thus, each set contained 24 (pre-defined) pictures in total. One picture of each set was randomly assigned to the list of primacy items (i.e. four items in total), from which three actually served as primacy items in the study phase (again, randomly selected for each participant anew). Ultimately, 69 pictures (+ 3 primacy pictures) were used in the study phase (23 per sound) and additionally 23 new pictures in the test phase (92 in total). See online supplement for further details on the selection of source and item material.

Procedure

All participants were tested in groups up to eight people in laboratory rooms with separated computer cubicles. First, participants provided written informed consent. Based on the order they came to the lab, they were then randomly assigned to the valence groups. The experiment was administered via the programming software OpenSesame (Mathôt et al., 2012). Except for the sound stimuli, the procedure was the same for both groups.

The sounds were presented via headphones, which participants wore throughout. The experiment started with a volume-regulation procedure: Participants were presented with the three sounds consecutively (order counterbalanced between participants) at medium volume. They could then adapt the volume (between a minimum pre-set volume and the PC sound card's maximum volume) for all three sounds simultaneously to their personal preference.

For the study phase, participants were instructed to memorise the neutral pictures for a later memory test and informed that each picture would appear together with one of three sounds. No explicit instructions on memorising the respective sounds of the pictures were given (i.e. incidental source learning). The pictures were presented centred on the screen in their original, standardised size (500 × 400 pixels, Kurdi et al., 2017). Each picture-sound combination was presented for three seconds in total (with the sound playing throughout), with each study trial initialised by a 750 ms fixation circle. Three of the four predefined picture sets (see Materials) were assigned as study items to one of the three sources each and the remaining set served as distractors in the test phase. This assignment was counterbalanced so that, across participants, each picture set was equally often presented with the negative (positive) high-arousing, the negative (positive) low-arousing, the neutral source or as a distractor (see test phase). Additionally, pictures were presented in a random order, with the constraint of maximally three immediate repetitions of the same sound source. Each sound was paired with 23 pictures, resulting in a total amount of 69 pictures in the study phase. In the beginning, three additional picture-sound pairs (one picture per sound) served as primacy items (not tested later).

After the study phase, participants performed a filler task of verifying simple mathematical equations for three minutes before turning to the test phase. In the following standard source-monitoring test, all 69 target pictures from the study phase (23 per source) and the 23 new pictures from the distractor set (i.e. 92 pictures in total) were presented consecutively on the top centre of the screen (picture size reduced to 75% of the original size) in a randomised order. Below, the labels of the sounds (e.g. "alarm", "belch", "train" for the Negative Group) and the option "new" were printed, rhombically arranged (the position of the sound labels were counterbalanced across participants, the option "new" was always printed at the centre bottom). Participants' task was to decide (self-paced) for each picture with which of the three sounds it was previously presented or whether it was not presented at all during the study phase by pressing the corresponding key ("D", "Z" and "K" for the sources, space key for new items). For exploratory reasons, we then administered a second test, which included source reinstatement to better tap into source storage specifically (cf., Symeonidou & Kuhlmann, 2021). As preregistered, our main interest and hypotheses pertained to the first, standard source-monitoring test only and we thus report analyses based on this test here. Finally, participants provided demographic information, were debriefed in detail and compensated (course credit or payment).

Results

We set $\alpha = .05$ for all analyses. For our main analysis, we applied the 2HTSM (Bayen et al., 1996), extended to three sources (Keefe et al., 2002). The 2HTSM models source-monitoring performance as jointly determined by memory and guessing processes, thus disentangling these processes and provides a purer measure of item and source memory (we additionally report standard performance measures based on hit and false alarm rates in the online supplement). Applied to this experiment, the following model parameters were estimated (see also Figure 1): The probability of item memory (i.e. detecting a target or distractor picture) is measured in parameter D (assumed to be equal across the sources/sounds). If a picture is recognised, the original source may also be correctly remembered, estimated separately by the emotional properties of the sound sources with probabilities d_{high}, d_{low} or $d_{neutral}$. If the source cannot be remembered (e.g. $1 - d_{\text{high}}$), guessing processes take place. Specifically, parameter g_{high} measures the probability to guess the high-arousing source, whereas parameter g_{low} measures the probability to guess low-arousing source (probability $1 - g_{low}$ for guessing the neutral source). As source memory for unrecognised items is at chance (Bell et al., 2017), participant's answers to unrecognised items ($1 - D$) are solely modelled by guessing processes: With probability b, participants guess that a picture was previously presented in the study phase (i.e. is "old"), followed by guessing that the picture was presented with the high-arousing (g_{high}), low-arousing (g_{low}) or neutral source ($1 - g_{low}$). With the complementary probability $1 - b$, participants guess that the picture is new.

We estimated the model specified as described based on the (across participants and items) aggregated observed response frequencies and evaluated model fit via maximum likelihood (ML) estimation methods implemented in the software *multiTree* (Moshagen, 2010). Note that this deviates from our preregistered proposal to use the Bayesian-hierarchical latent-trait approach (Klauer, 2010) for parameter estimation. As, however, we are interested in group rather than individual differences, the aggregative approach is preferable, because it estimates parameters on the group-level more precisely and allows us to use clear-cut inferential statistics (Chechile, 2009). Table 1 provides the parameter estimates based on the aggregated data. The model fit the data of both valence groups well, $G^2(5) = 7.19$, $p = .207$ in the Negative Group, $G^2(5) = 2.40$, $p = .792$ in the Positive Group. Note that this implies that there were no item memory differences across sources (i.e. item memory parameter D could be set equal across the sources). Also note that the Bayesian-hierarchical estimated parameters were very similar to these aggregated parameter estimates (however less precise) which speaks against a systematic bias due to participant heterogeneity (see online supplement).

Expectedly, source memory for the neutral train sound was comparable between both valence groups, $\Delta G^2(1) = 0.03$, $p = .863$. To test whether arousal had an effect on source memory, we set d_{high} (i.e. source memory for high-arousing source) and d_{low} (i.e. source memory for low-arousing sound) equal within each valence group. This led to a significantly worse model fit in the Negative Group, $\Delta G^2(1) = 4.99$, $p = .026$, but not in the Positive Group, $\Delta G^2(1) = 0.89$, $p = .344$. As evident in Table 1, source memory was worse for the negative high-arousing compared to the negative low-arousing source, violating our expectation. However, note this decrease in source memory was no longer significant when evaluated against Bonferroni–Holm adjusted $\alpha = .025$. Also note that the Bayesian-hierarchical analysis similarly suggests that the difference between d_{high} and d_{low} is not credible (judged based on their overlapping Bayesian credibility intervals, see online supplement). There was further no evidence for any arousal effect on source memory in the Positive Group. Moreover, we compared d_{high} to $d_{neutral}$ in each valence group. Although the comparison of d_{high} to $d_{neutral}$ is ambiguous in terms of the underlying dimension (arousal vs. valence), it can be still informative as it tests the combined effect of negative valence and high arousal, which usually proves to be strongest. Nonetheless, equating d_{high} to $d_{neutral}$ did not result in a significant model fit reduction in neither group, $\Delta G^2(1) = 1.72$, $p = .190$ for the Negative Group, $\Delta G^2(1) = 0.12$, $p = .726$ for the Positive Group. Thus, taken together, we conclude that there were no clear effects of arousal on source memory.

To test for valence effects, we set memory parameters of the same arousal level equal across both valence groups, that is $d_{negative\_high} = d_{positive\_high}$ for the high-arousing sources, and $d_{negative\_low} = d_{positive\_low}$ for the low-arousing sources. The only significant comparison suggested worse source memory for the negative compared to the positive high-arousing source, $\Delta G^2(1) = 3.89$, $p = .049$; however, this was no longer significant if tested against Bonferroni–Holm adjusted $\alpha = .025$. There was no valence difference between the low-arousing emotional sources, $\Delta G^2(1) = 1.95$, $p = .163$. Additionally, to comprehensively test for a valence effect, we set $d_{neutral}$ equal to $d_{negative\_low}$ and to $d_{positive\_low}$,

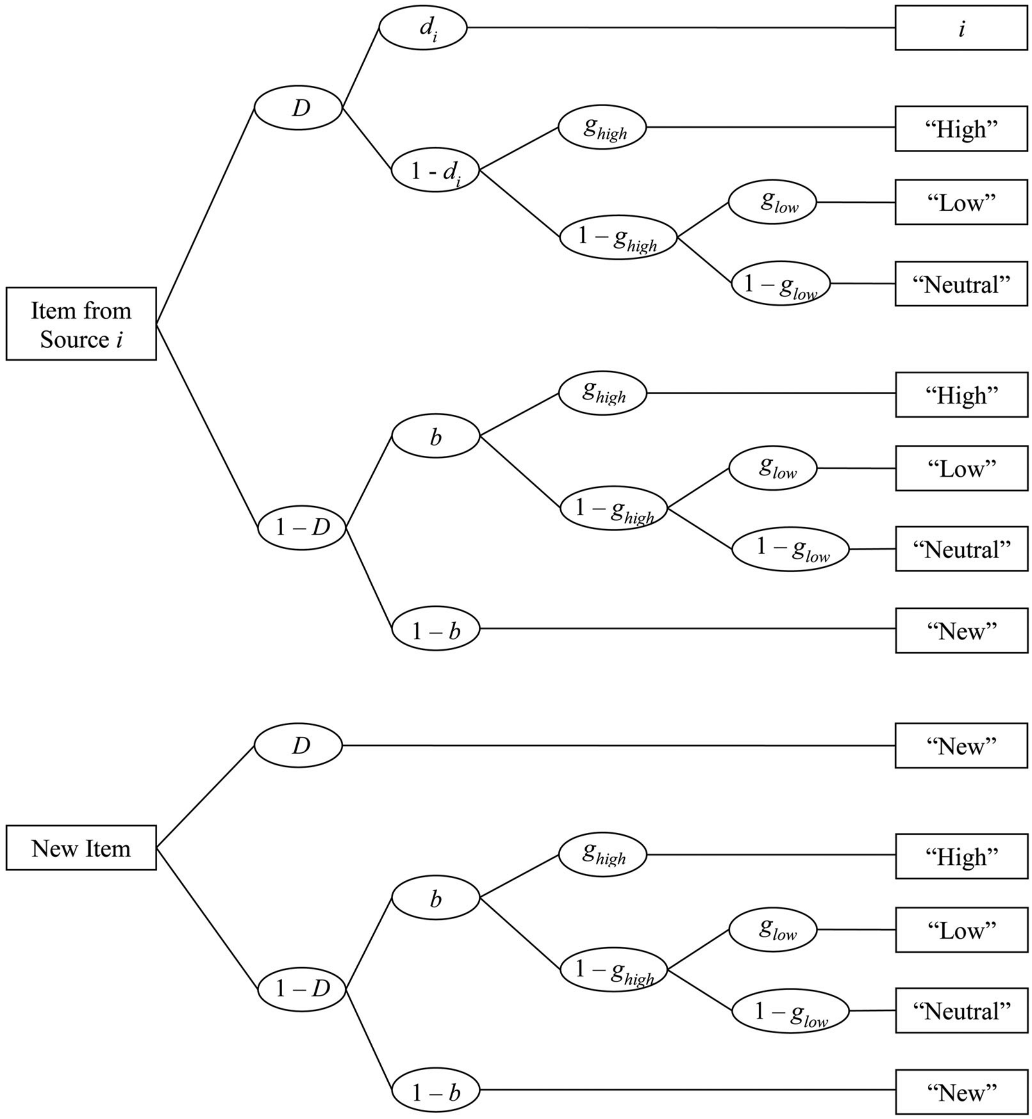


Figure 1. Graphical representation of the two-high-threshold multinomial model of source monitoring (2HTSM) for three sources.

Note: The figure shows sub-model 5d of the 2HTSM (Bayen et al., 1996) for target items (upper tree) and for new items (lower tree), extended to three sources. i denotes items from Source i , $i \in$ {High, Low, Neutral} with high and low referring to the arousal of the valenced sources. Boxes on the right represent participants' answers in the source memory test. D = probability of detecting an item as previously presented or not presented; d_i = probability of correctly recalling the source of a recognized item; b = probability of guessing that an item was previously presented; g_{high} = probability of guessing that a detected or undetected item was presented with the high-arousing source; g_{low} = probability of guessing that a detected or undetected item was presented with the low-arousing source. Adapted from "Source monitoring deficits for self generated stimuli in schizophrenia: multinomial modeling of data from three sources", by Keefe et al. (2002, p. 63).

respectively, within each valence group. None of these restrictions led to a significant model fit reduction, all ΔG^2s(1) ≤ 0.76, ps $\geq .383$. Taken together, there were no valence effects on source memory.

Discussion

The systematic manipulation of source valence and source arousal using IADS sounds in Experiment 1 yielded a somewhat surprising detrimental effect of source emotionality on source memory, which, however, was confined to negative high-arousing sources only (and not significant when tested against Bonferroni–Holm adjusted α). There were no effects of source arousal or valence per se. While this finding is in contrast to the emotion enhancement typically found for item memory (and also suggested for source memory; e.g. Bell & Buchner,

Table 1. Parameter estimates and model fit of the two-high-threshold multinomial model of source monitoring (2HTSM) extended to three sources for both experiments.

Experiment	Group	Model fit $G^2(5)$	Parameter estimates D	d_{high}	d_{low}	$d_{neutral}$	b	g_{high}	g_{low}
Experiment 1	Negative ($n = 40$)	7.19, $p = .207$	.58 [.55; 61]	.16 [.07; .25]	.31 [.23; .39]	.25 [.16; .34]	.38 [.34; .41]	.35 [.32; .38]	.48 [.44; .51]
	Positive ($n = 40$)	2.40, $p = .792$	.59 [.56; .62]	.28 [.20; .37]	.22 [.13; .31]	.26 [.18; .34]	.39 [.35; .43]	.33 [.30; .35]	.52 [.48; .56]
Experiment 2	Negative ($n = 86$)	1.30, $p = .934$	.52 [.50; .54]	.29 [.23; .36]	.28 [.21; .35]	.29 [.23; .36]	.47 [.45; .50]	.33 [.31; .35]	.52 [.49; .54]
	Positive ($n = 86$)	7.36, $p = .195$	.53 [.51; .56]	.30 [.24; .37]	.30 [.24; .37]	.32 [.25; .38]	.47 [.45; .50]	.33 [.31; .35]	.51 [.48; .54]

Notes: Brackets indicate 95% confidence intervals. D = probability of detecting an item as previously presented (equal across sources) or not presented; $d_{high/low/neutral}$ = probability of correctly recalling the (high-arousing/low-arousing/neutral) source of a recognised item; b = probability of guessing that an item was previously presented; g_{high} = probability of guessing that a detected or undetected item was presented with the high-arousing source; g_{low} = probability of guessing that a detected or undetected item was presented with the low-arousing source. Group "Negative" = high- and low-arousing sources were negative in valence. Group "Positive" = high- and low-arousing sources were positive in valence.

2012; Smith et al., 2005), it is in line with some research on the impairing effect of emotionality on associative binding (Bisby & Burgess, 2017; Chiu et al., 2013).

Alternatively, one might argue that the negative high-arousing source drew focused attention, as originally assumed, however to an extent that any surrounding information, including the item-source connection was neglected, resulting in poorer source memory. Regarding this, however, it is notable that item memory was not poorer for items paired with the negative high-arousing source suggesting that aside this emotional source the item was also focused on.

Further notably, the detrimental effect of negative high arousal on source memory might simply be a false positive. Indeed, this effect did not hold up robustly against a Bonferroni–Holm adjustment for the multiple tests, suggesting a mere chance effect. At the same time, our study was potentially not sensitive enough to reliably detect small differences in source memory via the d parameters. Indeed, a post hoc sensitivity power analysis using the observed parameter estimates, the total number of observations ($N_{obs} = 7360$ across valence groups) and a stricter power criterion ($\alpha = \beta = .05$) indicated that Experiment 1 was sensitive to medium-sized source memory differences only ($\Delta d = .23$ for the arousal comparison, $\Delta d = .22$ for the valence comparison).

Thus, before further engaging in any discussion on the underlying mechanisms of the found effect, we deemed it crucial to test the robustness of our results. We therefore conducted a second experiment with a different manipulation of source emotionality and different study material. To exclude the possibility that the found effect resulted from some idiosyncratic features of the material used or was simply a false positive (i.e. chance) effect rather than a true effect of negative high arousal per se, we considerably increased sample size (and thus sensitivity), and checked our source emotionality manipulation by collecting valance and arousal ratings, respectively.

Experiment 2

Availing ourselves of previous research (Pereira et al., 2021; Smith et al., 2005; Ventura-Bort et al., 2020), we used emotional background pictures to induce source emotionality and presented neutral words as items superimposed on these pictures. Similar to Experiment 1, we systematically varied valence versus arousal of the source (i.e. background pictures) between versus within participants, respectively, to disentangle their (potential) effects on source memory.

Building on the results from our first study, we updated our hypotheses and tentatively predicted to find lower source memory for negative high-arousing sources compared to low-arousing sources. In contrast, we did not expect an effect of high arousal on source memory paired with positive valence and also no effects of negative valence per se on source memory.

Additionally, we improved this experiment to address potential shortcomings of Experiment 1. For one, we used two different (instead of only one) stimuli per valence-arousal-combination (e.g. two different pictures for the negative high-arousing source) to counter habituation to the emotional sources. Further, we collected participants' valence and arousal ratings of the source stimuli at the end of the experiment and included only those participants in the data analysis, who perceived the sources as intended based on the norms.

Method

Design and participants

The design replicated that of Experiment 1, a two × three-mixed design with source valence (positive vs. negative) manipulated between participants and source arousal (high vs. low vs. neutral-low) manipulated within participants. A crucial difference of this experiment from the first is that we used background pictures as sources with each source (of a specific valence × arousal) made up of two pictures and ensured that the valence and arousal ratings for the source pictures were as intended for all participants included in the analysis.

We conducted a step-wise a priori power analysis on the comparison of d_{high} versus d_{low} (i.e. source-memory parameter from the 2HTSM) within the Negative Group: Population parameter values were first assumed as observed in Experiment 1 and then adapted to the new material based on the first recruited 64 eligible data sets (i.e. 32 eligible data sets per group; see online supplement for further details). The final power analysis with $\alpha = .05$ and $1 - \beta = .80$ yielded a total number of 7517 required observations to detect the difference of .15 between d_{high} versus d_{low} in the expected direction ($d_{high} < d_{low}$) as observed in the Negative Group of Experiment 1. Thus, we aimed at recruiting 86 eligible participants per valence group (i.e. 86*88 items = 7568 observations). Given the Covid-19 pandemic, this experiment had to be conducted online. In total, we recruited 172 eligible participants ($n = 86$ in the Negative Group and $n = 86$ in the Positive Group; 53 women, $M = 24.81$ years, $SD = 3.77$ years in the Negative Group; 49 women, 1 other, $M = 24.21$ years, $SD = 3.78$ in the Positive Group) via the online recruitment platform Prolific (https://www.prolific.co/). Based on our preregistered eligibility criteria, we excluded 113 additionally recruited, however non-eligible participants from data analysis (see online supplement for details; the reported results do not change if all participants are considered). Note that assignment of participants to valence groups was not random because we had to rerun the Negative Group after the first data collection due to a substantial model misfit (see online supplement). This misfit was presumably caused by semantic similarities across the two negative (high- and low-arousing) source categories, which is why we changed the material in our second recruitment. The reported sample characteristics refer to this second Negative Group, included in the main analysis.

Materials

Pictures to serve as emotional background sources were taken from the OASIS database (Kurdi et al., 2017). To ensure that the arousal and valence ratings of the OASIS norming sample generalise to Prolific participants and to cluster pictures to superordinate source categories, we pre-selected 10 categories (alcohol, car race, destruction, dogs, fire, fireworks, flowers, garbage, injury, lakes; based on Kurdi et al.'s categorisation) and seven pictures per category to conduct a pre-study ($N = 43$, via Prolific). Based on participants' valence and arousal ratings (on a 7-point rating scale, with 1: negative/low-arousing and 7: positive/high-arousing, for valence and arousal, respectively) and category assignments in the pre-study, we selected two negative (high- vs. low-arousing; categories: "fire" vs. "garbage"), two positive (high- vs. low-arousing; categories: "fireworks" vs. "flowers"), and one neutral category ("alcohol") for the main study, each consisting of two pictures (i.e. overall 10 pictures). Thus, within each valence group, six pictures à three categories (two pictures per category) alternated with a total of 11 presentations per picture. The high and low arousal of the negative and positive pictures was matched. For the neutral source, we used the same pictures (neutral valence, arousal matched to the low-arousing negative and positive pictures) in both groups. A total of 91 (3 primacy buffers, 66 study words, 22 distractors) neutral words as items to be superimposed on the pictures were taken from Janschewitz (2008). Words were randomly assigned to serve as study items (equally split between sources) versus distractor items for each participant anew. Details on the selection of source and item material are described in the online supplement.

Procedure

The experiment was built in lab.js (Henninger et al., 2021) and hosted on the server application OpenLab (https://open-lab.online/). Participants needed a PC or laptop to work on the study (completing the study with a smartphone or tablet was technically not possible). Participants received a description of the study and its requirements on Prolific and, after deciding to participate, were redirected to OpenLab to conduct the actual experiment. They were then randomly assigned to the valence groups by using OpenLab's function "customize parameters".

After providing informed consent, participants had to perform a scaling task in order to adapt the size of the background pictures to participants' screen size. This ensured that the source pictures covered a large part of participant's screen (providing their contextual nature) and, at the same time, had the same physical size across a certain range of different screen sizes (i.e. heights of 12 cm versus 15 cm versus 20 cm [and width 1.25* thereof] for estimated browser window heights [in cm] of [13, 17] versus [17, 22] versus ≥ 22, respectively). Participants with a browser window height smaller than 13 cm (in full screen) were precluded from further participation.

After passing the scaling task, participants started with the actual source-monitoring task, consisting of a study phase, a filler task and a test phase. In the study phase, participants were presented with neutral words shown in a black-framed, white box and superimposed on emotional (negative/positive, high-arousing/low-arousing) or non-emotional (neutral, low-arousing) source pictures. To additionally strengthen the contextual nature of the pictures, each picture was first presented on its own for 750 ms, then with the word superimposed on it for 3000 ms and finally again presented on its own for another 750 ms (without the word) before the next trial began. In total, 66 words were presented in the study phase in a random order (maximum three same-category words in direct succession), equally split between the three source categories (i.e. 22 words per source category, thus 11 words per picture). Further, three additional word-picture-pairs (one word per picture category; picture per category was randomly drawn for each participant anew) served as primacy items in the beginning of the study phase and thus were not included in the data analysis. After the study phase, and the filler task (as in Experiment 1), participants continued with the test phase. In this standard source-monitoring test, all 66 words from the study phase (22 per source) and additionally 22 new words from the remaining word pool (i.e. 88 words in total) were presented at the centre of the screen in a randomised order. Below, the three source category labels (e.g. "fireworks") were printed side by side on the screen. Beneath these labels, there was a horizontal black line and the option "new" was printed at the centre bottom underneath. Participants' task was to decide (self-paced) for each word with which of the three source categories it was previously presented or whether it was not presented at all during the study phase (option new). Participants used the keyboard to indicate their responses ("1" [left], "2" [middle] and "3" [right] for the source categories, assignment randomised for each participant anew at the start of the test, and space key for new items [fixed assignment]).

Following this, participants rated all pictures in terms of valence (first) and arousal (second) after an explanation of both terms. Finally, participants provided demographic information (age, gender, years of education, and highest level of education) and were asked about any problems with legibility of instructions and stimuli and could provide other feedback in an open field. Reimbursement of 2.5 GBP for 30 minutes was awarded via Prolific.

Results

We set $\alpha = .05$ for all analyses. As in Experiment 1, we used the same sub-model of the 2HTSM (Bayen et al., 1996; Keefe et al., 2002) for our main analysis and estimated the parameters based on the aggregated observed response frequencies (see online supplement for performance measures and Bayesian-hierarchical estimates). The model fit the data well in both groups, $G^2(5) = 1.30$, $p = .934$ in the Negative Group, $G^2(5) = 7.36$, $p = .195$ in the Positive Group. Note again that this implies that, similar to Experiment 1, there were no item memory differences across sources. Parameter estimates are listed in Table 1.

Expectedly, source memory for the neutral picture category was comparable between both groups, $\Delta G^2(1) = 0.22$, $p = .636$. To test for arousal effects, we compared source memory for the high-arousing versus low-arousing source (i.e. $d_{high} = d_{low}$) and also tested for the combined emotionality effect of negative/positive high arousal (i.e. $d_{high} = d_{neutral}$) within each valence group. None of these comparisons turned out significant, all ΔG^2s$(1) \leq 0.10$, $ps \geq .755$. In other words, source memory for high- versus low-arousing sources was equal within each valence group. Thus, the effect of reduced source memory for negative high- (versus low-) arousing sources found in Experiment 1 did not replicate.

To test for valence effects, we compared source memory between groups for sources of the same arousal level. More specifically, we equated $d_{negative\_high}$ to $d_{positive\_high}$ for the high-arousing sources, and $d_{negative\_low}$ to $d_{positive\_low}$ for the low-arousing sources. Additionally, to comprehensively test for valence effects, we set $d_{neutral}$ equal to $d_{negative\_low}$ and to $d_{positive\_low}$, respectively, within each valence group. None of these restrictions led to a significant model fit reduction, all ΔG^2s(1) ≤ 0.23, *ps* ≥ .633, suggesting that negative (or positive) valence per se did not have any effects on source memory.

As preregistered, we additionally performed a subgroup analysis by selecting a subgroup of participants who had a particularly high rating-difference between the high- and low-arousing categories in the Negative Group and the Positive Group, respectively. Again, in both valence groups, source memory for the high-arousing source did not differ from source memory for the low-arousing source, $\Delta G^2(1) \leq 1.39$, $p \leq .238$ (see online supplement for parameter estimates). Thus, even when the perceived arousal difference of both sources was particularly high, there was no effect of negative high arousal on source memory, again failing to replicate the observed detrimental effect in Experiment 1.

Discussion

Using emotional pictures as sources and ensuring that they were perceived as intended by the included participants, Experiment 2 did not replicate the effect of reduced source memory for negative high-arousing sources found in Experiment 1 with emotional sounds. There was no evidence for a source memory reduction for negative high-arousing sources, even in a subgroup of participants who perceived a particular high arousal difference between the pictures selected for the high- versus low-arousing negative (or positive) source. This casts further doubt that the effect observed in Experiment 1 was driven by negative high arousal per se rather than being an idiosyncratic effect of the operationalising sound or simply a false positive. To check our study's sensitivity to arousal effects, we again conducted a post hoc sensitivity power analysis using the observed parameter estimates and the total number of observations ($N_{obs} = 15136$ across valence groups) and applying a stricter power criterion ($\alpha = \beta = .05$). This analysis indicated, that our study was sensitive enough to detect smaller-sized differences of .18 between d_{high} and d_{low}. Thus, we can be confident that Experiment 2 was well-powered to find even small effects of source arousal on source memory, further strengthening the conclusion, that there is no effect of arousal per se on source memory.

Clearly, and like Experiment 1, there was no evidence for an emotion-enhancement effect in source memory, in contrast to its typical manifestation in item memory. As in Experiment 1, the current results suggest that there were no effects of arousal or valence per se on source memory.

General discussion

The primary goal of our research was to systematically investigate whether there is a source memory advantage for emotional sources akin to the EEM effect in item memory (Talmi & McGarry, 2012) and suggested in some previous studies on emotional source memory (Bell & Buchner, 2010, 2012; Smith et al., 2005; but see Arnold et al., 2021; Bell et al., 2017). In two experiments, we manipulated source arousal within and source valence between participants to disentangle their potential effects on source memory. Multinomial model-based estimates of source memory revealed no (robust) effects of source emotionality (valence and/or arousal) on source memory. Somewhat surprisingly, in Experiment 1, in which we used emotional sounds as sources and neutral pictures as items, we found a detrimental effect of negative high arousal on source memory. However, this effect was not significant when tested against Bonferroni–Holm adjusted α and did not replicate in Experiment 2, in which we used emotional background pictures as sources with neutral words superimposed as items. There further were no effects of positive high- and low-arousing sources on source memory across both experiments. Altogether, we conclude that there are no robust effects and clearly no memory-enhancing effects of source emotionality (valence and/or arousal) on source memory. That is, emotional valence, emotional arousal or a combination of both do not per se improve source memory.

Potential moderators for source emotionality effects

Why do some studies find an EEM effect for emotional sources and others (including ours) do not? We would like to highlight two methodological aspects that might help clarifying such inconsistencies. First, as discussed in the introduction, many studies finding EEM effects in source memory have used emotional source features that inherently qualified the item (e.g. person [= item] who is cheating [= source]; Bell & Buchner, 2012) or instructions that facilitated integrating item and source (e.g. imagining the item as part of the source; Ventura-Bort et al., 2016). This fosters the storage of the item-source pair as one "emotional unit" which then might profit from similar EEM effects as those typically found for emotional items (see Chiu et al., 2013; Dolcos et al., 2017; Murray & Kensinger, 2013 for a similar reasoning). However such unitised representations also blur the distinction between item and source and their associated memory and neural processes (e.g. hippocampus-reliance; Diana et al., 2008; Murray & Kensinger, 2013). Thus, in our experiments, we deliberately opted for incidental source instructions and a clear distinction between items and sources as we wanted to investigate whether emotional sources per se, independent of the item, are remembered better, which does not seem to be the case.

The second aspect that we deem to be critical for EEM effects in source memory is the number of emotional stimuli used. As we followed the standard source-monitoring paradigm, its typical many-to-few mapping (cf. Chalfonte & Johnson, 1996) involved repeating the emotional sources (23 times in Experiment 1 and 11 times in Experiment 2) across several items. In contrast, other studies on emotional source memory used multiple emotional sources, implementing a less typical one-to-one mapping of items and sources. Based on our experience with emotional stimuli in experiments, we think that emotionality effects are quite dependent on the specific material in use and such material dependencies carry more weight if the used emotional material is limited in number as in the traditional many-to-one mapping in source monitoring. Put simply, in studies with many emotional stimuli it is less important whether a certain stimulus elicits the intended level of valence and arousal because other stimuli may compensate for it. Apart from potentially explaining why other studies on emotional sources with one-to-one mapping found an enhancement effect in source memory (e.g. Bell & Buchner, 2010), this would also explain why valence and arousal effect are more robust in item memory research which necessarily employs multiple emotional items.

Crucially, such material dependencies might also account for the inconsistent results across our two experiments. That is, the negative high-arousing sound in Experiment 1 might have had a specific, idiosyncratic influence on source memory. As we unfortunately did not assess participants' valence and arousal ratings for the sounds in Experiment 1, we cannot directly compare them to the picture stimuli in Experiment 2. However, as we ensured to only include participants who perceived the source pictures as intended in Experiment 2, we are confident that there was no effect of source valence or arousal per se. One may object that the sounds elicited higher (negative) arousal than the pictures but we deem this unlikely. It has been repeatedly shown that emotional pictorial material successfully induces emotional responses in respective face expressions and physiological parameters (Lang et al., 1993). Emotional pictures are considered to be powerful tools to manipulate emotion and have been successfully applied as such in hundreds of studies (Marchewka et al., 2014). Furthermore, participants in our subgroup analysis of Experiment 2 perceived a very strong arousal difference (spanning almost the entire scale), yet showed no source memory difference. One might further emphasise the modality difference between the source manipulations of the two experiments. Perhaps participants were able to ignore the emotional picture sources in Experiment 2 but not the emotional sound sources in Experiment 1. However, we carefully designed the procedure of Experiment 2 to ensure attention to the source pictures (i.e. picture covered most of the screen and was presented alone first; word [item] was superimposed on picture so picture is attended when focusing on the word) and source memory was well above chance and comparable to that in Experiment 1. Therefore, we do not believe that the modality difference of sources across experiments can satisfyingly explain their diverging results or that the effect in Experiment 1 describes a general effect of (a specific level of) negative arousal. Rather, we believe that the observed effect in Experiment 1 idiosyncratically depends on the specific sound selected or was simply a chance effect, as it did not withstand a stricter test against an adjusted α level. Extending our research, future studies could combine our approach of systematically and independently varying source valence and source arousal with a one-to-one mapping of sources and items to study the influence of source valence versus source arousal. From a theoretical perspective, however, a one-to-one mapping of sources to items may challenge the perception of what is source (= context) and what is item (= central information; cf., Glisky et al., 2001), and might thus rather tap into item-to-item binding instead of item-to-source binding. As explained next, emotionality effects might differ between these binding types.

Currently, the evidence on emotionality effects on item-to-item binding is mixed (see Bisby et al., 2016; Pierce & Kensinger, 2011; Touryan et al., 2007 for detrimental effects and Guillet & Arndt, 2009; Nadarevic, 2017 for enhancement effects). To resolve such inconsistencies, more recent accounts have emphasised the importance to consider neural processes and systems involved in forming memory representations of emotional events (Bisby & Burgess, 2017; Chiu et al., 2013). For example, Bisby and Burgess (2017) state in their dual-representation account that negative emotion disrupts only hippocampus-dependent memories (such as associative memory) by down-regulating hippocampal activity. They provide empirical support for their idea by showing impaired memory for item-item associations for pairs containing negative items (i.e. neutral-negative pairs and negative-negative pairs) compared to neutral-only pairs. Given that the hippocampus also supports item-to-source binding (Mitchell & Johnson, 2009), this account would make the same predictions for item-to-source binding. Note however that associative memory (including source memory) becomes less reliant on the hippocampus with unitisation (Diana et al., 2008). Crucially, difficulty of unitisation might systematically vary across item-item versus item-source associations (e.g. emotional sources are often informative about the item, such as its safety, naturally facilitating unitisation), making them more or less hippocampus-dependent. This might also explain why negative emotion sometimes disrupts associative memory even if unitisation instructions are used (e.g. Bisby et al., 2018). Future studies could try to measure unitisation success (e.g. via self-reports, see Murray & Kensinger, 2012) to account for such potential differences.

Notably, we acknowledge that the observed detrimental effect of the negative high-arousing sound on source memory in Experiment 1 might be a true effect in line with the dual-representation account. But it is difficult to reconcile why there was no such detrimental effect in Experiment 2, which ensured perceived high negative source emotionality (high arousal and valence) in included participants. Admittedly, however, as our studies differ from the original studies on the dual-representation account on several aspects (i.e. item-to-source vs. item-to-item binding; incidental vs. intentional source encoding; few vs. many emotional stimuli) and did not originally aim at testing this account, we refrain from drawing conclusions on this account.

Finally, although our studies suggest that there is no effect of source emotionality on source memory assessed shortly after study (or the effect is weak at best), this pattern might change with an extended retention interval (Pierce & Kensinger, 2011). Research on EEM suggests that emotionality effects on memory become more pronounced with a longer retention interval, presumably due to an amygdala-related modulation of consolidation processes (McGaugh, 2000; Talmi, 2013). This is especially true for high arousing material (McGaugh, 2006). Thus, it might be worthwhile to investigate whether the effects of source valence and especially arousal are stronger (or even reverse, see Pierce & Kensinger, 2011) with a longer retention interval.

In sum, we think that future studies on emotionality and associative memory/memory binding need to further refine the concept of binding by differentiating between item-to-item binding versus item-to-source binding (Chiu et al., 2013). Similarly, the effects of emotion on associative memory might depend on how the association is represented in memory (i.e. as bound unit or as associated, but distinct events) and on the length of the retention interval. The use of unitisation instructions (e.g. integrated mental image) and self-reports to measure unitisation success, as well as the systematic manipulation of retention interval length might help to investigate these boundary conditions (Murray & Kensinger, 2012).

Crucially, as suggested by others (Chiu et al., 2013; Dolcos et al., 2017), we think that future research should more thoroughly consider the neural systems and brain regions that are involved in forming and retrieving emotional (bound) representations when investigating emotionality effects on associative memory including source memory.

Strengths and limitations

A major strength that distinguishes our research from previous studies is that we used normed emotional material for the source manipulation and took great effort to match it in terms of valence (within groups) and arousal (between groups), which allows us to disentangle their effects. Also, in Experiment 2, we ensured that the included participants perceived valence and arousal of the sources as intended based on the norms. A further strength of both experiments is that we avoided confounding item and source memory by using external sources not naturally related to the items and keeping source learning incidental. Furthermore, we modelled our data with the 2HTSM model, which dissociates memory processes from guessing bias and thus provides unbiased measures for item and source memory. Finally, we ensured reasonable power to detect emotionality effects on source memory via a priori power analyses.

Having said that, we also acknowledge that our research comes along with certain limitations. Although we conducted an a priori power analysis to ensure a reasonable power ($1 - \beta = .80$) for detecting emotionality effects on source memory, a post hoc sensitivity analysis with a stricter power criterion ($1 - \beta = .95$) in Experiment 1 indicated satisfactory sensitivity only to *d* parameter differences that were considerably larger than the ones assumed in our a priori power analysis (see online supplement). Thus, Experiment 1 was perhaps not reliably sensitive to smaller-sized *d* parameter differences. However, this was different in Experiment 2. There, we conducted a step-wise a priori power analysis, which allowed us to take sensitivity-critical factors (e.g. the level of item memory performance) into account based on the first half of eligible participants. Thus, even when applying a stricter power criterion in the post hoc sensitivity power analysis compared to our a priori power analysis ($1 - \beta = .95$ instead of .80), our study was still sensitive enough to detect *d* parameter differences of .18, which are smaller than the ones typically observed in source emotionality studies (e.g. Bell & Buchner, 2012).

Admittedly, as our second experiment needed to be conducted online there were more factors producing random noise (e.g. variations in screen size, sources of distraction etc.) compared to lab settings. However, there are many comforting indications that participants complied well with our instructions and honestly worked on our online experiment such as memory performance comparable to the laboratory experiment and reasonable task duration (approx. 30 min). Combined with the considerable sample size increase from Experiment 1 to Experiment 2 (i.e. twice as many participants), we do not believe that the online nature of Experiment 2 limits the validity of our results.

Conclusion

To conclude, the goal of our research was to systematically investigate whether there is a source memory advantage for emotional sources (akin to the EEM effects found for item memory) by independently varying source valence and source arousal using normed emotional sounds (Experiment 1) and pictures (Experiment 2) as sources. Both of our studies clearly indicate that there is no beneficial effect of source valence or arousal (or their combination) on source memory, meaning that the EEM effects for item memory do not simply transfer to source memory. If anything, Experiment 1 showed that source memory may be even reduced under conditions of negative high arousal (Experiment 1). However, as we did not replicate the detrimental effect of negative high-arousing sources on source memory in Experiment 2, we cannot straight-forwardly conclude that this effect is clearly ascribable to negative high arousal. Rather we believe that the effect in Experiment 1 was simply a chance effect or emerged due to idiosyncratic properties of the sound used for the operationalisation of this source. Such material-dependent influences might weigh more when only a small number of emotional stimuli is used as typically done in the standard source-monitoring paradigm. Building on our research, future studies should continue to systematically examine source valence and arousal but employ multiple emotional sources in a less typical one-to-one mapping of sources and items and systematically vary item-source unitisation via learning instructions (Chiu et al., 2013; Murray & Kensinger, 2012) as a promising moderator of source emotionality effects.

Acknowledgements

We thank Katja Bitz, Michelle Dörnte, Alina Kias, Franziska Leipold, Paula Schmelzer, and Pia Vogel for help with participant recruitment and data collection.

Disclosure statement

We have no financial or non-financial competing interests that might influence the results reported in this manuscript.

Funding

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) [grant number GRK 2277 "Statistical Modeling in Psychology"].

Data availability statement

Data and analysis code (and/or output files) that support the results reported in this manuscript are publicly available in the Open Science Framework repository at https://osf.io/j2axs/?view_only=9f8283891e2e4136921a862e1fcdc982. Both of our experiments were preregistered and can be accessed via https://osf.io/ey64z/?view_only=0772062849314a07badb9046a4bc159d (Experiment 1) and https://osf.io/kupfs/?view_only=fd08a15e7f504b099fb2b70f1c381abd (Experiment 2).

Details on power analysis, material characteristics and exclusion criteria, as well as additional (explorative) time-course analyses are reported in the online supplement.

ORCID

Nikoletta Symeonidou http://orcid.org/0000-0002-2471-8653
Beatrice G. Kuhlmann http://orcid.org/0000-0002-3235-5717

References

Arnold, N. R., González Cruz, H., Schellhaas, S., & Bublatzky, F. (2021). A multinomial modelling approach to face identity recognition during instructed threat. *Cognition & Emotion*, *35*(7), 1302–1319. https://doi.org/10.1080/02699931.2021.1951175

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037//1089-2680.5.4.323

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 197–215. https://doi.org/10.1037/0278-7393.22.1.197

Bell, R., & Buchner, A. (2010). Valence modulates source memory for faces. *Memory & Cognition*, *38*(1), 29–41. https://doi.org/10.3758/MC.38.1.29

Bell, R., & Buchner, A. (2011). Source memory for faces is determined by their emotional evaluation. *Emotion (Washington, D.C.)*, *11*(2), 249–261. https://doi.org/10.1037/a0022597

Bell, R., & Buchner, A. (2012). How adaptive is memory for cheaters? *Current Directions in Psychological Science*, *21*(6), 403–408. https://doi.org/10.1177/0963721412458525

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

Bisby, J. A., & Burgess, N. (2013). Negative affect impairs associative memory but not item memory. *Learning & Memory*, *21*(1), 21–27. https://doi.org/10.1101/lm.032409.113

Bisby, J. A., & Burgess, N. (2017). Differential effects of negative emotion on memory for items and associations, and their relationship to intrusive imagery. *Current Opinion in Behavioral Sciences*, *17*, 124–132. https://doi.org/10.1016/j.cobeha.2017.07.012

Bisby, J. A., Horner, A. J., Bush, D., & Burgess, N. (2018). Negative emotional content disrupts the coherence of episodic memories. *Journal of Experimental Psychology. General*, *147*(2), 243–256. https://doi.org/10.1037/xge0000356

Bisby, J. A., Horner, A. J., Hørlyck, L. D., & Burgess, N. (2016). Opposing effects of negative emotion on amygdalar and hippocampal memory for items and associations. *Social Cognitive and Affective Neuroscience*, *11*(6), 981–990. https://doi.org/10.1093/scan/nsw028

Bradley, M. M., & Lang, P. J. (2007). *The international affective digitized sounds (2nd edition; IADS-2): Affective ratings of sounds and instruction manual* (Technical report B-3). University of Florida.

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously. A theory of socioemotional selectivity. *The American Psychologist*, *54*(3), 165–181. https://doi.org/10.1037//0003-066X.54.3.165

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*(4), 403–416. https://doi.org/10.3758/BF03200930

Chechile, R. A. (2009). Pooling data versus averaging model fits for some prototypical multinomial processing tree models. *Journal of Mathematical Psychology*, *53*(6), 562–576. https://doi.org/10.1016/j.jmp.2009.06.005

Chiu, Y.-C., Dolcos, F., Gonsalves, B. D., & Cohen, N. J. (2013). On opposing effects of emotion on contextual or relational memory. *Frontiers in Psychology*, *4*, 103. https://doi.org/10.3389/fpsyg.2013.00103

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 730–740. https://doi.org/10.1037/0278-7393.34.4.730

Dolcos, F., Katsumi, Y., Moore, M., Berggren, N., Gelder, B. d., Derakshan, N., Hamm, A. O., Koster, E. H. W., Ladouceur, C. D., Okon-Singer, H., Pegna, A. J., Richter, T., Schweizer, S., van den Stock, J., Ventura-Bort, C., Weymar, M., & Dolcos, S. (2020). Neural correlates of emotion-attention interactions: From perception, learning, and memory to social cognition, individual differences, and training interventions. *Neuroscience and Biobehavioral Reviews*, *108*, 559–601. https://doi.org/10.1016/j.neubiorev.2019.08.017

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

Erk, S., Martin, S., & Walter, H. (2005). Emotional context during encoding of neutral items modulates brain activation not only during encoding but also during recognition. *NeuroImage*, *26*(3), 829–838. https://doi.org/10.1016/j.neuroimage.2005.02.045

Glisky, E. L., Rubin, S. R., & Davidson, P. S. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1131–1146. https://doi.org/10.1037//0278-7393.27.5.1131

Guillet, R., & Arndt, J. (2009). Taboo words: The effect of emotion on memory for peripheral information. *Memory & Cognition*, *37*(6), 866–879. https://doi.org/10.3758/MC.37.6.866

Henninger, F., Shevchenko, Y., Mertens, U., Kieslich, P. J., & Hilbig, B. E. (2021). lab.js: A free, open, online experiment builder. *Behavior Research Methods*. https://doi.org/10.3758/s13428-019-01283-5

Jaeger, A., & Rugg, M. D. (2012). Implicit effects of emotional contexts: An ERP study. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(4), 748–760. https://doi.org/10.3758/s13415-012-0110-1

Janschewitz, K. (2008). Taboo, emotionally valenced, and emotionally neutral word norms. *Behavior Research Methods*, *40*(4), 1065–1074. https://doi.org/10.3758/BRM.40.4.1065

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

Kang, C., Wang, Z., Surina, A., & Lü, W. (2014). Immediate emotion-enhanced memory dependent on arousal and valence: The role of automatic and controlled processing. *Acta Psychologica*, *150*, 153–160. https://doi.org/10.1016/j.actpsy.2014.05.008

Keefe, R. S. E., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source-monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources. *Schizophrenia Research*, *57*(1), 51–67. https://doi.org/10.1016/S0920-9964(01)00306-1

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review: Journal of the International Society for Research on Emotion*, *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., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(2), 110–126. https://doi.org/10.3758/CABN.6.2.110

Klauer, K. C. (2010). Hierarchical multinomial processing tree models: A latent-trait approach. *Psychometrika*, *75*(1), 70–98. https://doi.org/10.1007/s11336-009-9141-0

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the open affective standardized image set (OASIS). *Behavior Research Methods*, *49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

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

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

Marchewka, A., Zurawski, Ł., Jednoróg, K., & Grabowska, A. (2014). The nencki affective picture system (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behavior Research Methods*, *46*(2), 596–610. https://doi.org/10.3758/s13428-013-0379-1

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

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

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

May, C. P., Rahhal, T., Berry, E. M., & Leighton, E. A. (2005). Aging, source memory, and emotion. *Psychology and Aging*, *20*(4), 571–578. https://doi.org/10.1037/0882-7974.20.4.571

McGaugh, J. L. (2000). Memory-a century of consolidation. *Science*, *287*(5451), 248–251. https://doi.org/10.1126/science.287.5451.248

McGaugh, J. L. (2006). Make mild moments memorable: Add a little arousal. *Trends in Cognitive Sciences*, *10*(8), 345–347. https://doi.org/10.1016/j.tics.2006.06.001

Meyer, M. M., Buchner, A., & Bell, R. (2016). Influences of age and emotion on source guessing: Are older adults more likely to show fear-relevant illusory correlations? *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *71*(5), 831–840. https://doi.org/10.1093/geronb/gbv016

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*(4), 638–677. https://doi.org/10.1037/a0015849

Moshagen, M. (2010). Multitree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, *42*(1), 42–54. https://doi.org/10.3758/BRM.42.1.42

Murray, B. D., & Kensinger, E. A. (2012). The effects of emotion and encoding strategy on associative memory. *Memory & Cognition*, *40*(7), 1056–1069. https://doi.org/10.3758/s13421-012-0215-3

Murray, B. D., & Kensinger, E. A. (2013). A review of the neural and behavioral consequences for unitizing emotional and neutral information. *Frontiers in Behavioral Neuroscience*, *7*, 42. https://doi.org/10.3389/fnbeh.2013.00042

Nadarevic, L. (2017). Emotionally enhanced memory for negatively arousing words: Storage or retrieval advantage? *Cognition & Emotion*, *31*(8), 1557–1570. https://doi.org/10.1080/02699931.2016.1242477

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2021). 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

Pierce, B. H., & Kensinger, E. A. (2011). Effects of emotion on associative recognition: Valence and retention interval matter. *Emotion (Washington, D.C.)*, *11*(1), 139–144. https://doi.org/10.1037/a0021287

Rahhal, T. A., May, C. P., & Hasher, L. (2002). Truth and character: Sources that older adults can remember. *Psychological Science*, *13*(2), 101–105. https://doi.org/10.1111/1467-9280.00419

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

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

Schellhaas, S., Arnold, N., Schmahl, C., & Bublatzky, F. (2020). Contextual source information modulates neural face processing in the absence of conscious recognition: A threat-of-shock study. *Neurobiology of Learning and Memory*, *174*, 107280. https://doi.org/10.1016/j.nlm.2020.107280

Smith, A. P. R., Dolan, R. J., & Rugg, M. D. (2004). Event-related potential correlates of the retrieval of emotional and nonemotional context. *Journal of Cognitive Neuroscience*, *16*(5), 760–775. https://doi.org/10.1162/089892904970816

Smith, A. P. R., Henson, R. N. A., Rugg, M. D., & Dolan, R. J. (2005). Modulation of retrieval processing reflects accuracy of emotional source memory. *Learning & Memory*, *12*(5), 472–479. https://doi.org/10.1101/lm.84305

Symeonidou, N., & Kuhlmann, B. G. (2021). A novel paradigm to assess storage of sources in memory: The source recognition test with reinstatement. *Memory*, *29*(4), 507–523. https://doi.org/10.1080/09658211.2021.1910310

Talmi, D. (2013). Enhanced emotional memory. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Talmi, D., & McGarry, L. M. (2012). Accounting for immediate emotional memory enhancement. *Journal of Memory and Language*, *66*(1), 93–108. https://doi.org/10.1016/j.jml.2011.07.009

Touryan, S. R., Marian, D. E., & Shimamura, A. P. (2007). Effect of negative emotional pictures on associative memory for peripheral information. *Memory*, *15*(2), 154–166. https://doi.org/10.1080/09658210601151310

Ventura-Bort, C., Löw, A., Wendt, J., Moltó, J., Poy, R., Dolcos, F., Hamm, A. O., & Weymar, M. (2016). Binding neutral information to emotional contexts: Brain dynamics of long-term recognition memory. *Cognitive, Affective & Behavioral Neuroscience*, *16*(2), 234–247. https://doi.org/10.3758/s13415-015-0385-0

Ventura-Bort, C., Wendt, J., Wirkner, J., König, J., Lotze, M., Hamm, A. O., Dolcos, F., & Weymar, M. (2020). Neural substrates of long-term item and source memory for emotional associates: An fMRI study. *Neuropsychologia*, *147*, 107561. https://doi.org/10.1016/j.neuropsychologia.2020.107561

Is There an Emotionality Effect in Older Adults' Source Memory?

Nikoletta Symeonidou*, Abdolaziz Hassan, Isabel Porstein, & Beatrice G. Kuhlmann

Department of Psychology, School of Social Sciences, University of Mannheim, Germany

Author Note

Nikoletta Symeonidou https://orcid.org/0000-0002-2471-8653

Beatrice G. Kuhlmann https://orcid.org/0000-0002-3235-5717

*Correspondence concerning this article should be addressed to Nikoletta Symeonidou, Department of Psychology, University of Mannheim, D-68161 Mannheim, Germany. +49(0)621-181-3390, E-mail: nsymeoni@mail.uni-mannheim.de.

Abstract

The goal of our research was to investigate whether older adults show a source memory enhancement for emotionally valenced sources. Additionally, building on research on the socioemotional selectivity theory and the age-related positivity effect (Carstensen et al., 1999), we tested whether older adults show a larger enhancement for positive compared to negative (and neutral) sources than younger adults. In Experiment 1 (n_{old} = 25, n_{young} = 27), we used one positive, one negative, and one neutral picture to manipulate source valence (many-to-one mapping of items to sources), whereas, in Experiment 2 (n_{old} = 62, n_{young} = 62), we used multiple pictures per source valence category (one-to-one mapping of items to sources) to counteract potential habituation effects. In both experiments, sources had medium and matching arousal levels. Items were neutral words superimposed on the source pictures. To support an implicit, natural information processing, participants rated the words in terms of pleasantness. We analyzed memory data with a multinomial processing tree model to disentangle memory processes from guessing bias. Across both experiments, an age-related positivity effect occurred in participants' pleasantness ratings. This effect, however, did not carry over to older adults' source memory. That is, in source memory, we found a general emotionality effect for younger but not for older adults and no age-related positivity effect. We propose that due to older adults' pronounced difficulties in remembering the item-to-source *link* (i.e., associative deficit), even a greater focus on an inherently emotional source might be insufficient to boost source memory.

Keywords: source memory, aging, emotion, positivity effect, multinomial modeling

Is There an Emotionality Effect in Older Adults' Source Memory?

Memory gets worse as we age – a widespread assumption among laypeople that found steady empirical support over the last decades: Research has confirmed that, indeed, aging comes along with several cognitive deficits such as impaired attention control, reduced working memory capacity, and worse episodic memory (Carstensen & Mikels, 2005; Milham et al., 2002; Nilsson, 2003; Salthouse, 1994). One of the most profound age-related declines concerns source memory, that is, memory for the (spatial, temporal, social, or emotional) context of information (Johnson et al., 1993). Many studies have shown that older adults compared to younger adults have more difficulties in remembering the speaker, spatial context, or peripheral features (e.g., color or font) of an information (Chalfonte & Johnson, 1996; Kuhlmann & Boywitt, 2016) due to a reduced ability to bind the item information to its source features (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008). These behavioral observations are further corroborated by neurological findings that show a substantial age-related volume shrinkage in brain regions that are crucially involved in the formation of associations and bound memory representations (i.e., hippocampus and prefrontal cortex; Driscoll et al., 2003; Shing et al., 2011), such as source memory (see Mitchell & Johnson, 2009).

Having said that, there is also evidence suggesting that age-related deficits in memory (or other cognitive functions) are partially driven by differences in processing priorities (i.e., motivational differences) between older and younger adults, making these deficits malleable to a certain degree. This idea traces back to Carstensen's *socioemotional selectivity theory* (SST), which suggests that older compared to younger adults prioritize present-related (as opposed to future-related) goals due to their greater awareness of the finite nature of life (Carstensen et al., 1999; Carstensen, 2006). Consequently, older adults are more motivated to elaborate on emotionally meaningful and especially on positive information, which in turn improves their memory for such information. This emotionality effect (and positivity effect in particular) has been well established for item memory (i.e., memory for central information; see Reed et al., 2014 for a meta-analysis). Two studies suggest that older adults also remember presumably emotional sources, but these studies did not use normed emotional stimuli and did not distinguish between positive and negative valence (May et al., 2005; Rahhal et al., 2002).

Thus, the goal of our research was to investigate whether there is an emotionality effect, and a positivity effect in particular, in older adults' source memory. In both of our experiments, we systematically manipulated source valence by using background pictures of either positive, negative, or neutral valence as sources, keeping arousal constant at a medium level. This allowed us to specifically test whether older adults show prioritized processing and, consequently, better memory for positively valenced compared to negatively (and neutrally) valenced sources, as predicted by the socioemotional selectivity theory reviewed next.

The Socioemotional Selectivity Theory and the Positivity Effect in Memory

The *socioemotional selectivity theory* (SST; Carstensen et al., 1999) assumes that age-related differences in cognition are partially driven by differences in motivational priorities. More specifically, SST suggests that motivational priorities change over time as a function of perceived time until death. That is, "when time is perceived as open-ended, knowledge-related goals are prioritized. In contrast, when time is perceived as limited, emotional goals assume primacy" (Carstensen et al., 1999, p. 165). As age and (perceived) time until death are interrelated, the theory predicts that (healthy) younger adults rather focus on optimizing future outcomes (e.g., by acquiring knowledge), whereas older adults rather pursue present-oriented goals (e.g., maximizing emotional meaning and satisfaction). Importantly, this age-related shift in the goal hierarchy goes along with a shift in processing preferences: Younger adults put more emphasis on information seeking and knowledge gain. In contrast, older adults rather elaborate on emotionally meaningful information. Since its proposition, the SST has motivated many studies and gained steady empirical support from different psychological areas, including social, differential, and cognitive psychology (Carstensen et al., 1999). For instance, research on emotional item memory has shown that younger and older adults both preferably process emotional over neutral items (e.g., pictures or words), but differ in the *type* of emotional material on which they focus (Carstensen, 2006). More specifically, studies have shown that younger adults typically show a negativity bias in item memory (i.e., better memory for negative items), whereas older adults seem to preferably process and memorize positive items (i.e., positivity bias) or put less emphasis on negative items (i.e., reduced negativity bias; Reed et al., 2014). This phenomenon, in both manifestations, is termed *positivity effect* (Mather & Carstensen, 2005) and, in short, describes older compared to younger adults' relative preference for positive over negative information.

The positivity effect has been well investigated for item memory (see meta-analysis by Reed et al., 2014) but not for source memory. This is surprising given that the age-related source memory deficit might be reduced if older adults' processing preference for emotional (and especially positive) material (see SST) extends to emotional sources. There is some initial evidence showing that older adults' source memory is boosted for emotional (i.e., goal-relevant) instead of perceptual (i.e., goal-irrelevant) source information. May and colleagues (2005) presented food items at different screen positions (i.e., left vs. right) to younger and older adults. Participants were either told that the item's screen position is indicative of its serving temperature (i.e., hot vs. cold) or of its safety (i.e., spoiled vs. not spoiled). In the subsequent source memory test, dependent on their assigned test group, participants had to indicate the original screen position (perceptual source) versus serving temperature (conceptual, non-emotional source) versus safety (conceptual, emotional source) of the food item, or to indicate that the item was new. Results yielded that, while younger adults' source memory was comparably high across all types of sources, older adults' source memory was substantially enhanced for the emotionally meaningful source cues (i.e., safety of the food item) and even reached the level of younger adults' source memory. Similarly, Rahhal et al. (2002) report that older adults' source memory benefited if the source was indicative of the item's truth status or source's character and thus qualified with emotional meaning instead of simply perceptual (i.e., source's voice). In contrast, younger adults did not show such a benefit. Again, the increase in older adults' source memory due to the emotional sources leveled their performance to those of younger adults.

These two studies impressively illustrate the crucial role of motivational factors in source memory and challenge the deterministic idea that the pronounced age differences in source memory are solely the result of age-related structural and neuronal changes in critical brain areas such as the hippocampus (Mitchell & Johnson, 2009). Note, however, that May et al. (2005) and Rahhal et al. (2002) did not differentiate between positive and negative sources when reporting source memory and speak of general source memory enhancements for emotional sources in their studies. Considering the previously outlined research on the age-related positivity effect in item memory, a differentiation between positive and negative sources is advisable and would generally promote a more fine-grained insight into the influence of source valance on older adults' source memory. Further crucial, May et al. (2005) and Rahhal et al. (2002) both used conceptually emotional source cues (e.g., source signals item's safety vs. danger), thus manipulating the emotionality of the item-source *link*, not the inherent emotionality of the source per se. Notably, in an experiment by Davidson et al. (2006), using inherently emotional sources (i.e., voices with emotional tone), older adults' source memory did not benefit as much from source emotionality as younger adults'.However, as Davidson et al. only used negative emotional sources, this may reflect a reduced negativity bias, and for inherently positive sources, older adults may show a particular boost in source memory, in line with a positivity effect.

The Current Research

We aimed at testing whether there is a general emotionality effect in source memory for older adults, that is, whether older adults remember inherently emotional (i.e., negative and positive) sources better compared to neutral sources. Furthermore, we wanted to investigate whether this emotionality effect in older adults was more pronounced for inherently positive compared to inherently negative sources, which would indicate a positivity bias in source memory. For a more comprehensive test of the positivity effect in older adults, we also recruited a comparison group of younger adults. This allowed us to test whether older adults' source memory profits relatively more from positive (vs. negative) sources compared to younger adults' source memory.

In both experiments, we carefully selected our source material in terms of valence and arousal (see below for more detailed information). We systematically differentiated between positive versus negative sources to test for a positivity effect, and we kept arousal level constant across sources to exclude it as a potential confound in any occurring emotionality effects. Moreover, we created experimental conditions that favored the unfolding of motivational effects. We kept arousal as low as possible to ensure that top-down motivational processes were not overshadowed by bottom-up attentional processes (cf., Kensinger, 2008). This is well in line with the broader finding that an experimentally directed information processing counteracts the motivation-based positivity effect because it works against the inherently motivated way of processing (cf., SST; Reed et al., 2014). Relatedly, we chose incidental instead of intentional learning instructions in both experiments to ensure a more natural way of information processing. More specifically, we asked participants to rate the pleasantness of (neutral) words that were presented as central items together with an emotional versus neutral source in the study phase.

Considering the sparse evidence, we were cautious in formulating strict hypotheses regarding source memory. In general, we expected to find better source memory for emotional (i.e., negative and positive) sources compared to neutral sources in both older and younger adults. We further deemed it plausible to find a positivity effect in source memory for older adults, that is, an age-related relative preference to focus on (and better memorize) positive over negative source information. As in item memory (Reed et al., 2014), this positivity effect might manifest in two possible ways: Older adults might show enhanced source memory for positive sources compared to negative sources (i.e., positivity bias). Or, if there is a negativity bias, such that source memory for positive sources is poorer than that for the negative sources, this difference might be less pronounced in older compared to younger adults (i.e., reduced negativity bias). Either way, to establish a positivity effect, it is necessary to contrast memory for positive versus memory for negative source information (Reed et al., 2014). Note, however, that it is important to first test whether there are general emotionality effects, that is, whether memory for the positive or negative source differs from memory for the neutral source, as the latter provides a baseline memory level. Put differently, the reduction or absence of differences between positive and negative sources only then indicates a positivity effect if memory for the neutral source is still lower than memory for either emotional (positive or negative) source. Otherwise, the pattern (negative = positive = neutral) would suggest a null effect of emotionality.

Furthermore, we reasoned that the effects of source valence on participants' pleasantness ratings would parallel those on source memory. More specifically, we expected higher versus lower scores for words presented with a positive versus negative source. In other words, the influence of source valence on participants' pleasantness ratings provides a measure for their attention to and processing of sources.[1]

Experiment 1

Experiment 1 followed the standard source-monitoring paradigm, using three pictures (one positive, one negative, one neutral) as sources in the study phase, with each source presenting several items (i.e., *many-to-one* mapping, cf. Glisky et al., 2001; Schacter et al., 1994). Instructions, layout (words superimposed on images), the number of study items, and the encoding time were motivated by a pilot study, which we had conducted with older adults only. We uploaded a detailed description of this study and its results on the Open Science Framework (OSF) repository at https://osf.io/9suqj/?view_only=fe02f43d968e4ccc9f2e8938682cbcc5.

Method

Design

We used three different pictures (negative, positive, neutral) to manipulate source valence (see Appendix, Table A1). The experiment followed a 2 (age group: older adults vs. younger adults) × 3 (source valence: positive, negative, neutral) mixed design with age group manipulated between and source valence manipulated within participants. Our main dependent variable was source memory as measured by parameter *d* of the 2HTSM. Additionally, we examined participants' mean pleasantness ratings (per source picture) in the study phase.

[1] Originally, we also hypothesized that source valence will influence participants' response times for providing the pleasantness ratings. However, we had to implement a rather long initial fixed presentation time to ensure above-chance memory. Therefore, response time after this long processing was little informative of processing preferences that likely unfolded earlier and did not vary with valence in either age group.

Participants

Using the software G*Power (Faul et al., 2007), we conducted an a priori power analysis with $\alpha = .05$ and $1\text{-}\beta = .80$ and an expected effect size of $d = 0.55$ based on Kensinger's (2008) experiment, which investigated the age-related positivity effect in item memory (with a repeated measures design). More specifically, we drew on the therein reported difference between older adults' corrected hit rates for positive versus neutral items.[2] Assuming a repeated measures correlation of .40 (based on Reed et al., 2014), the resulting required sample size for our study was $n = 22$ (per age group). To ensure that our four counterbalancing conditions were equally sized, we aimed at recruiting $n = 24$ participants per age group (i.e., $n = 6$ per counterbalancing condition; actual distribution negligibly ranged from 6 to 8 participants) thus $N = 48$ participants in total.

Ultimately, 33 older adults (>= 50 years old) and 28 younger adults (18-30 years old) participated in our online study recruited via snowballing and university courses for senior citizens. Eight older participants were excluded from data analysis because they did not meet pre-defined eligibility requirements (German as native language [i.e., learned before the age of six]; age: 50+ years old for older adults, 18-30 years old for younger adults; no diagnosed depression disorder within the past 6 months; no history of heart attack, stroke, pneumonia or COPD, severe head/brain injury, or addiction to alcohol or drugs; no Parkinson's disease; no untreated [i.e., drug-controlled] hypertension; no dementia; no previous or current treatment with chemotherapy; no recent [i.e., past month] intake of benzodiazepines). One younger participant was excluded because they never responded "new" in the source memory test. Thus, the final data analysis was based on 52 participants (25 older adults aged $M = 63.76$ years, $SD = 5.99$ years, and 27 younger adults aged $M = 21.96$ years, $SD = 2.39$ years). At the

[2] As the positive-negative difference in Kensinger's (2008) study was more strongly pronounced than the positive-neutral difference, we drew on the latter to derive a more conservative effect size estimate.

time of testing, older adults had completed slightly more total years of education (*M* = 19.38, *SD* = 5.07) than the younger adults (*M* = 15.37, *SD* = 2.63), $t(50) = 3.62$, $p < .001$, $d = 1.00$.

To characterize our online sample more comprehensively, we additionally measured performance on a pattern comparison task (Salthouse, 1996) and a vocabulary task (Riegel, 1967, see below for more details) in both age groups. Older adults classified fewer patterns correctly (*M* = 30.00 out of 60 patterns, *SD* = 8.46) compared to younger adults (*M* = 45.00 out of 60, *SD* = 8.26), showing that, as expected, younger adults had faster processing speed compared to older adults, $t(49) = 6.42$, $p < .001$, $d = 1.80$[3]. Both younger and older adults completed slightly more patterns correctly in this online computerized assessment of the pattern comparison task than our previous lab samples on the paper-based task (e.g., Kuhlmann & Touron, 2016) but the age difference was comparably pronounced. In reverse, older adults outperformed younger adults on the vocabulary task, *M* = 75% correct answers (*SD* = 11%) versus *M* = 65% correct answers (*SD* = 10%) for older and younger adults, respectively, $t(50) = 3.57$, $p < .001$, $d = 0.99$, again as expected. Vocabulary performance and the age difference therein was comparable to that of lab samples of younger and older adults, which we previously assessed with the same computerized task (Kuhlmann & Undorf, 2018).

Materials

Three emotional pictures were taken from the standardized picture database *Open Affective Standardized Image Set* (OASIS; Kurdi et al., 2017) and were intended to serve as emotional source information. All pictures depicted sceneries to support their background character. We ensured that the positive and negative pictures matched in terms of absolute

[3] Responses on the pattern comparison task were missing for one participant, presumably due to some idiosyncratic browser-task incompatibilities or because the person took a short break after the study phase. However, their retention interval was approximately three minutes long and thus comparable to the pre-defined retention interval (of three minutes) for the remaining participants.

valence, and all pictures were matched in terms of arousal such that they all had low arousal levels (see Appendix, Table A1).

Neutral words, superimposed on the pictures, served as items. There were drawn from the *Berlin Affective Word List Reloaded* (BAWL-R; Võ et al., 2009), a database with German words that are normed for valence, arousal, and imageability (among other criteria). One hundred twenty words of neutral valence (]1.5; 1.5[on a rating scale rating from -3 [negative] to +3 [positive]), low arousal (≤ 2.5 on a 5-point rating scale, with higher values indicating higher arousal levels), and moderate imageability (> 3 on a 7-point rating scale, with higher values indicating higher imageability) were chosen for the experiment. From these, we randomly selected 60 words and distributed them on four lists (à 15 words each), matched on mean valence, arousal, and imageability. The assignment of the item lists to the target sets versus the distractor set was counterbalanced between participants so that across participants, each list was (approximately) equally often presented with the negative versus positive versus neutral source picture or as a distractor in the test phase.

Procedure

The study was approved by the ethics board of the University of Mannheim (the same applies to the second study). The experiment was built in lab.js (Henninger et al., 2021) and hosted on the server application OpenLab (https://open-lab.online/). Personal information for the reimbursement was collected on SoSci Survey (https://www.soscisurvey.de/) (Leiner, 2019) after the experiment. Participants were randomly assigned to one of the four counterbalancing conditions using OpenLab's urn function. After providing informed consent, participants performed a scaling task to adapt the size of the background source pictures to participants' screen size. With this, we ensured that the source pictures in the study phase covered a large part of participants' screen, emphasizing their contextual character. To ensure good visibility of the context pictures, we required participants to have a minimum screen size of 13 inches, which the scaling task also checked.

In the subsequent study phase, participants were presented with 45 neutral words (see Material section for counterbalancing of word-set assignment to sources). Each word was shown in a black-framed, white box and superimposed on the (negative, positive, or neutral) source picture. The 45 words were equally split between the three pictures (i.e., 15 words per picture) and presented in random order with the constraint of maximum four successive same-picture repetitions. For each word, participants were asked to provide a self-paced pleasantness rating (on a 5-point scale, ranging from 1: very unpleasant to 5: very pleasant). There was no explicit reference to an upcoming memory test or the source pictures. To set a lower bound for presentation (i.e., processing) time, we showed the picture on its own for 750 ms before pairing it with a superimposed word for another 3000 ms. After the 3000 ms, the rating scale for the self-paced pleasantness ratings appeared underneath the word-picture pair, and participants were asked to judge how pleasant/unpleasant they perceived the word at that present moment ("Wie unangenehm oder angenehm finden Sie das Wort in diesem Moment?", roughly translates to "How pleasant or unpleasant do you perceive the word at the present moment?"). This was intended to provide participants with the option to also consider the contextual pictures when judging the pleasantness of the words. Immediately after the response, the next trial was initiated by a 500 ms fixation cross.

After the study phase, participants completed a pattern-comparison task for three minutes. This did not only function as a distractor task to eliminate the recency effect but also served as an assessment of processing speed (Salthouse, 1996) for sample characterization. In this task, participants were presented with two patterns of lines side-by-side and had to decide whether the patterns were the same (by pressing key 1) or different (by pressing key 0) as fast as possible. Participants completed two blocks of 30 seconds, which corresponded to the two pages of the original paper-based pattern comparison task. Each block ended automatically after 30 seconds. If participants completed both blocks, including instructions, in less than 3 minutes, the remaining items of the previous two blocks were presented and if there was still time remaining, already completed patterns may have repeated. This was only for filling the retention interval and not scored. Directly afterward, participants were given instructions on the test phase, which consisted of a standard source-monitoring test: All 45 words from the study phase plus 15 new distractor words were presented individually in random order at the top center of the screen. Below, the three source pictures (left, center, right; screen position was counterbalanced across participants) were presented side-by-side on the screen. The option "new" was printed at the center bottom. Participants had to decide for each word with which of the three pictures (negative, positive, neutral) it was previously presented or whether it was not presented at all during the study phase (new). Participants used the keyboard to indicate their responses ("1", "2," and "3" for left, center, and right picture, space key for new) in the self-paced memory test.

Following the test phase, participants rated the valence and arousal levels of the three used source pictures, which served as a manipulation check for the emotionality manipulation. Each picture was first presented for 3000 ms (order of presentation was random). After the picture disappeared, participants provided self-paced valence ratings (first) and arousal ratings (second) on a 7-point rating scale (ranging from 1: very negative to 7: very positive for valence, and 1: very low to 7: very high for arousal). We used the instructions provided by Kurdi et al. (2017) for the OASIS (translated into German) for a detailed explanation of both emotionality dimensions. Then, participants received instructions on a computerized version of Riegel's (1967) vocabulary task, which – similar to the pattern comparison task – served to characterize our online sample. Participants were presented with 20 words consecutively. Each word was printed at the top of the screen with five response options (other words or short phrases) beneath (labeled 1-5). Participants had to decide which of the five options matched best to the meaning of the target word by pressing the corresponding number. They were instructed to guess if they did not know (or were unsure about) the correct answer. The vocabulary task was self-paced.

Finally, participants provided demographic and health information. They were additionally asked whether they could work on the study focused and without disruptions and indicated whether all instructions and stimuli were legible (and, if not, they were asked to specify the exact problem in an open text field). Further crucially, they were asked whether they used any tools to artificially boost their memory performance (e.g., taking screenshots or the like). None of the responses here indicated any serious issues or cheating. Lastly, they had the chance to give optional feedback/notes in an open text field, which again did not reveal any problems.

They were then debriefed about our research intention and redirected to the SoSci survey website to provide their e-mail address (which was saved separately from their responses in the experiment) for monetary compensation and for information on our study results if desired.

Results and Discussion

Alpha level was fixed to $\alpha = .05$ for all analyses.

Pleasantness Ratings

Mean pleasantness ratings are displayed in Figure 1 (left-hand plot) and were submitted to a 2 (age group) × 3 (source valence) mixed ANOVA. There was a main effect of source valence, $F(1.52, 75.82) = 72.64$, $p < .001$, $\eta^2_p = .59$, a main effect of age, $F(1, 50) = 6.04$, $p = .018$, $\eta^2_p = .11$, and a source valence × age interaction effect, $F(1.52, 75.82) = 6.81$,

$p = .004$, $\eta^2_p = .12$.[4] Follow-up simple main effect analyses indicated significant valence effects in both age groups, $F(2, 49) = 11.66$, $p < .001$, $\eta^2_p = .32$ for the older adults, $F(2, 49) = 40.82$, $p < .001$, $\eta^2_p = .63$ for the younger adults. Bonferroni-Holm adjusted pairwise comparisons revealed that, in both age groups, words paired with the positive source were rated more pleasant compared to words paired with the neutral source, $t(50) = 4.50$, $p < .001$, $d = 1.14$, for older adults, and $t(50) = 7.59$, $p < .001$, $d = 1.26$, for younger adults, and also more pleasant compared to words paired with the negative source, $t(50) = 4.60$, $p < .001$, $d = 1.12$, for older adults, and $t(50) = 9.03$, $p < .001$, $d = 1.52$, for younger adults (see also Figure 1, left-hand plot). Also, words paired with the neutral source were rated more pleasant than words paired with the negative source, $t(50) = 5.74$, $p < .001$, $d = 1.10$, in the younger group, and $t(50) = 2.29$, $p = .026$, $d = 0.46$, in the older group. Comparing both age groups on each level of source valence further revealed that older adults rated words paired with the negative source less unpleasant than younger adults, $t(50) = 3.36$, $p = .002$, $d = 0.93$, whereas the age groups did not differ in their neutral and positive pleasantness ratings of words paired with the other respective source pictures, $t(50) = 1.65$, $p = .105$, and $t(50) = -1.02$, $p = .312$, for words paired with the neutral and positive source, respectively. This suggests that older participants showed a less pronounced negativity bias in their pleasantness ratings compared to younger adults, which is in line with previous research on the positivity effect sometimes manifesting as a reduced negativity bias (Reed et al., 2014).

[Figure 1 near here]

Taken together, these results suggest that both younger and older adults attended to the source pictures and incorporated their valence in their pleasantness ratings. Thus, source valence had an impact on participants' processing. Notably, older adults' ratings were less

[4] Due to violation of the sphericity assumption for the three-level within-subjects factor source valence in both experiments, the Greenhouse-Geisser correction was applied for all ANOVA tests involving this factor.

influenced by the negative pictures compared to younger adults, which points towards a reduced negativity bias (i.e., positivity effect) in older adults' pleasantness ratings.

Source Memory

We used the 2HTSM (Bayen et al., 1996), extended to three sources (Keefe et al., 2002) to obtain measures for item memory and, crucially, source memory. The 2HTSM rests on the assumption that memory performance in a source-monitoring task is jointly determined by item memory (parameter D), source memory (parameter d), and guessing processes (item old/new guessing, parameter b; source guessing, parameter g; see Figure 2). Thereby, the memory parameters of the model provide separate measures for item and source memory that are further corrected for response biases. Considering our research question as well as previous research on the statistical identifiability of sub-models of the 2HTSM (Bayen et al., 1996), we freely estimated the following model-specific parameters from the current source-monitoring test responses: The probability of item memory (i.e., detecting a target or distractor word), represented by parameter D (and assumed to be equal across the source pictures); the probability of source memory, separately for the positive, negative and neutral sources pictures, which was measured by parameters $d^{positive}$, $d^{negative}$, and $d^{neutral}$, respectively; in the case of a source memory failure (1-d), the probability of guessing the positive source ($g^{positive}$) or, when the positive source was not guessed (1-$g^{positive}$), the probability of guessing the negative source ($g^{negative}$) versus the neutral source (1- $g^{negative}$); and, finally, the probability of guessing that an item is old (b) versus new (1-b), if item memory fails.

[Figure 2 near here]

We estimated the parameters of this model version based on the aggregated observed response frequencies and evaluated model fit via maximum likelihood (ML) estimation methods using the software *multiTree* (Moshagen, 2010). To test for the presence of emotionality effects, we estimated the 2HTSM model separately in each age group and performed pairwise comparisons of the parameters (e.g., $d^{positive}$ versus $d^{negative}$) within each age group.

The model fit the data well, $G^2(5) = 7.53$, $p = .184$ for younger adults, and $G^2(5) = 8.72$, $p = .121$ for older adults. The resulting parameter estimates for each age group are listed in Table 1; the observed source memory pattern is additionally visualized in Figure 3 (left-hand plot). Note again that the memory parameters of the 2HTSM (i.e., *D* and all *d*s) are already corrected by guessing bias. Therefore, 0 denotes chance performance. As apparent from Table 1, all memory parameter estimates were substantially above 0 (i.e., did not include 0 in their confidence intervals), indicating above-chance item and source memory performance.

We first tested for general emotionality effects by comparing source memory for emotional sources (i.e., positive and negative) to source memory for the neutral source in both age groups. More specifically, we set the respective source memory parameters equal (i.e., $d^{positive} = d^{neutral}$; and $d^{negative} = d^{neutral}$) and compared the fit of these models to the fit of our baseline model by means of χ^2 distributed difference tests.

[Table 1 near here]

[Figure 3 near here]

For the younger group, these difference tests revealed that source memory for the positive source was better compared to source memory for the neutral source, $\Delta G^2(1) = 7.75$, $p = .005$. Although the descriptive pattern was the same for the negative-neutral comparison (negative > neutral; see Figure 3, left-hand plot), this difference was not statistically reliable, $\Delta G^2(1) = 2.18$, $p = .140$. Furthermore, source memory for negative and positive sources did not differ, $\Delta G^2(1) = 1.94$, $p = .163$. For a more powerful test of emotionality effects, we set $d^{positive}$ and $d^{negative}$ equal (i.e., $d^{valanced}$), providing a more reliable estimate of source memory for emotionally valenced sources because twice as many data points fed the estimation of $d^{valanced}$. Contrasting the joint estimate $d^{valanced}$ against $d^{neutral}$ led to a significant model fit reduction, $\Delta G^2(1) = 6.00$, $p = .014$, speaking for a general emotionality effect in younger adults' source memory (i.e., better source memory for emotional over neutral sources).

In contrast, for the older group, the imposed parameter restrictions did not worsen model fit substantially, $\Delta G^2(1) = 0.01$, $p = .938$, for $d^{positive} = d^{neutral}$, and $\Delta G^2(1) = 0.02$, $p = .879$, for $d^{negative} = d^{neutral}$, indicating that older adults' source memory for the emotional sources did not differ significantly from their source memory for the neutral source (see Figure 3, left-hand plot). Thus, there was no evidence for a general emotionality effect in older adults' source memory. To test for a potential positivity bias more specifically, we also equalized $d^{positive}$ and $d^{negative}$ and tested this restriction against the baseline model. Again, this did not lead to a significant model fit reduction, contradicting a positivity bias, $\Delta G^2(1) = 0.01$, $p = .929$. Thus, there was no evidence for a positivity bias in older adults' source memory. For a more powerful test of emotionality effects, we set $d^{positive}$ and $d^{negative}$ equal (i.e., $d^{valanced}$) and contrasted this joint estimate against $d^{neutral}$. Note that this additionally served as a replication attempt for the findings of May et al. (2005) and Rahhal et al. (2002), who similarly compared source memory for emotional versus neutral sources. Again, this comparison of $d^{valanced}$ with $d^{neutral}$ did not lead to a significant model fit reduction, suggesting that there were no general effects of source valence on source memory in the older group, $\Delta G^2(1) = 0.02$, $p = .902$.

Finally, to test for age-group differences in source memory, we equalized d parameters of the same valence across age groups (e.g., $d_{negative\_YA} = d_{negative\_YA}$). Replicating previous studies on the age-related source memory deficit (Old & Naveh-Benjamin, 2008), we found that source memory was overall lower in older compared to younger adults, $\Delta G^2(3) = 13.05$, $p = .005$. Somewhat surprisingly, however (and contrary to the results of May et al., 2005; Rahhal et al., 2002), this deficit was driven by older (compared to younger) adults' poorer source memory for positive sources, $\Delta G^2(1) = 9.87$, $p = .002$. Source memory for the negative source also tended to be lower in older compared to younger adults, but not significantly, $\Delta G^2(1) = 2.56$, $p = .110$. Source memory for the neutral source did not differ across age groups, $\Delta G^2(1) = 0.09$, $p = .770$.

Overall, contrary to our expectations, we did not find a positivity effect nor a general emotionality effect in source memory for older adults. In contrast, younger adults' source memory pattern suggested a general emotionality effect. Note, however, that although younger adults' source memory for the negative source was descriptively better than for the neutral source, this pairwise comparison was not significant. This might have been a low-power issue, as a post-hoc sensitivity power analysis (with $\alpha = .05$ and $1\text{-}\beta = .80$) indicated that our study was sensitive to detect source-memory differences of $\Delta d = .23$, thus detecting the positive-neutral difference ($\Delta d = .21$) but not the smaller positive-negative difference ($\Delta d = .13$). We considered this issue when designing our second experiment. Crucially note that while specific tests comparing the effects of positive versus negative valence were underpowered here, the test for finding a positivity effect in older adults' source memory (i.e., $d^{positive} > d^{neutral}$) was sufficiently powered. More specifically, we computed the post-hoc power for finding a positivity effect in older adults' source memory if the effect was of comparable size as the one found in younger adults' source memory (i.e., $\Delta d = .21$) with the given sample size (i.e., $n = 25$) and $\alpha = .05$. The analysis yielded a satisfactory power of $1\text{-}\beta = .83$ to find a comparable emotionality effect ($\Delta d = .21$) in older adults' source memory, rendering low power a rather unlikely explanation for the observed null effect.

Valence and Arousal Ratings

To check whether our source valence manipulation was successful, we analyzed participants' valence and arousal ratings. As the source pictures were selected such that they varied in terms of valence but not in terms of arousal, we expected a main effect of source valence on participants' valence ratings and no effect on their arousal ratings in both age groups.

Regarding participants' valence ratings, a repeated-measures ANOVA revealed a main effect of source valence, $F(1.46, 72.98) = 360.26$, $p < .001$, $\eta^2_p = .88$, a main effect of age, $F(1, 50) = 4.76$, $p = .034$, $\eta^2_p = .09$, and no source valence × age interaction, $F(1.46, 72.98) = 1.88$, $p = .169$. Bonferroni-Holm adjusted pairwise comparisons of the source valence level revealed that, as expected, the positive picture was rated more positively compared to the neutral picture, $t(50) = 13.92$, $p < .001$, $d = 1.90$, and compared to the negative picture, $t(50) = 41.08$, $p < .001$, $d = 5.57$; and the neutral picture was rated more positively than the negative one, $t(50) = 10.02$, $p < .001$, $d = 1.40$. Also, the means show that, across age groups, the positive, neutral, and negative pictures were perceived as positive, neutral, and negative, respectively, corroborating our manipulation. As to the main effect of age, the mean valence ratings across all emotionality levels indicated that older adults' ratings were overall more positive than younger adults' ratings, suggesting an age-related positivity effect in the perception of the pictures' valence (see also Grühn & Smith, 2008; Ready et al., 2017). Descriptively, this was particularly the case for the negative and neutral pictures, though the source valence × age interaction was not significant.

With regard to the arousal ratings, the pattern was more complicated. There was a main effect of source valence, $F(2, 100) = 5.69$, $p = .005$, $\eta^2_p = .10$, and a source valence × age interaction, $F(2, 100) = 3.39$, $p = .038$, $\eta^2_p = .06$, but no main effect of age, $F(1, 50) = 1.21$, $p = .277$. Follow-up simple main effect analysis indicated significant valence effects only for younger adults, $F(2, 49) = 10.95$, $p < .001$, $\eta^2_p = .31$, but not for older adults, $F(2, 49) = 1.18$, $p = .315$. Bonferroni-Holm adjusted pairwise comparisons further revealed that younger adults judged the negative picture to be more arousing than the neutral picture, $t(50)$

= 4.36, p < .001, d = 0.89 and the positive picture, $t(50) = 3.05$, p = .007, d = 0.56. There was however no difference between the positive and neutral picture, t<1.

Overall, these results suggest that we manipulated source valence as intended in both age groups. However, we were less successful in keeping source arousal constant across all three sources in the younger sample because younger adults tended to perceive the negative source as more arousing relative to the positive source. Note, however, that this variation in arousal cannot sufficiently account for the observed source valence effect on younger adults' source memory: If arousal (rather than valence) was the driving factor behind the effect, younger adults' source memory should have been better for the (higher-arousing) negative compared to the (lower-arousing) positive source. Yet our results yielded the opposite memory pattern (descriptively $d^{positive} > d^{negative}$ and significantly $d^{positive} > d^{neutral}$ despite comparable arousal). In contrast, older adults perceived the positive and negative (and neutral) sources as comparably low arousing, as intended. As such, these ratings suggest optimal preconditions for emotionality effects on older adults' memory, rendering the absence of such effects on source memory particularly noteworthy.

Taken together, the results of Experiment 1 show that both younger and older adults incorporated source valence in their pleasantness ratings. That is, participants' pleasantness ratings varied symmetrically with source valence and even showed typical age-related patterns (i.e., positivity effect). Importantly, this impact on the pleasantness ratings translated into a source memory benefit for emotionally valanced sources only in younger, but not in older adults. Put differently, although older adults' ratings were affected by the emotionally valanced sources in an expected direction, this impact did not translate into source memory benefits.

Notably, the results of Experiment 1 are constrained by the drawback that we used only one picture per source category, which possibly compromised the effectiveness of our emotionality manipulation. Although this many-to-one mapping procedure is the standard method to investigate source memory, this procedure seems suboptimal for emotional sources. For one, it bears the risk that participants habituate to the emotional material, thus reducing its potential effects. Yet, habituation effects should have been similarly pronounced across both age groups in our experiment and thus cannot fully explain our age- and source-valence-sensitive result pattern for both pleasantness ratings and source memory. Alternatively, one might argue that the three pictures were simply unsuitable for inducing emotionality effects in older adults' source memory. This is, however, unlikely because picture valence influenced older adults' pleasantness ratings as expected. Also, their valence ratings at the end of our study suggest that our valence manipulation was successful, even after the repeated exposure during the study. But we admit that these ratings might have been biased by demand characteristics as each picture was rated in the context of only two other pictures, making the demanded valence rather obvious. Finally, another disadvantage of using one picture per source is that this procedure confounds picture content with its valence. That is, the absence of emotionality effects in older adults could have been due to source valence or due to idiosyncratic features of the used pictures. Likewise, the presence of emotionality effects in younger adults is potentially confounded by idiosyncrasies of the chosen pictures. To counter these constraints, we conducted a second experiment, in which we used several pictures per source valence category.

Experiment 2

The main goal of our second experiment was to corroborate the results of Experiment 1 using many pictures (instead of only one) per source valence category. More specifically, each item was paired with a unique (negative, positive, or neutral) picture, resulting in a one-to-one mapping (instead of many-to-one mapping) of items to sources (Glisky et al., 2001). We selected pictures based on their OASIS norm ratings. As we used several pictures for each valence category, idiosyncratic deviations of a picture's perceived emotionality from the norm ratings have relatively little influence because other pictures of that category can compensate for it. Thus, we refrained from collecting valence and arousal ratings for each picture at the end of the experiment, which would have been rather lengthy given the increased number of source pictures, and as discussed in Experiment 1, such ratings at the end of the study may be biased by demand characteristics and habituation effects (although not repeated at study, source pictures were repeatedly presented in the source-monitoring test).

Furthermore, based on a tailored a priori power analysis, we considerably increased our sample size (n = 62 per age group) to ensure sensitivity even for small differences in source memory parameters (i.e., Δd =.15; see description below). Finally, we applied a stricter age criterion for both younger and older adults to maximize the age difference between both groups, which has been shown to boost the positivity effect (Reed et al., 2014).

Method

Design

Our design was again a 2 (age group: older adults vs. younger adults) × 3 (source valence: positive, negative, neutral) mixed design with age group manipulated between and source valence manipulated within participants. Pleasantness ratings and source memory were our main dependent variables.

Participants

Based on an a priori power analysis, we aimed for n = 62 participants per age group (N = 124 in total, which corresponds to 124*60 trials = 7440 total observations) to detect .15 differences between source memory parameters with power 1-β = .80 and α = .05 (power calculations were conducted with multiTree, Moshagen, 2010). This a priori power analysis was based on the parameter estimates of Experiment 1, in which differences between emotional and neutral source memory ranged from .13 (negative-neutral comparison) to .21 (positive-neutral comparison). Population parameter values, which are required for the power analysis, were fixed to the empirically observed parameter values of the older adults in Experiment 1 as we especially wanted to maximize the chance to find emotionality effects for the older group (i.e., $D = .73$; $b = .15$; $g^{positive} = .27$; $g^{negative} = .50$; $d^{neutral} = .39$; $d^{positive}$ and $d^{negative} = d^{neutral} + .15 = .54$). In other words, this tailored power-analysis ensured that Experiment 2 was sensitive to even small effects of source valence on source memory. Ultimately, 62 eligible older adults (aged *M* = 63.76 years, *SD* = 5.99 years) and 62 eligible younger adults (aged *M* = 21.96 years, *SD* = 2.39 years) participated in our online study, recruited via the platform Prolific (https://www.prolific.co/)[5]. All participants were English native speakers with US residence and met our pre-defined eligibility requirements (same as listed for Experiment 1, except English [not German] as native language [i.e., learned before the age of six]; age: 60+ years old for older adults, 18-25 years old for younger adults).

At the time of testing, older adults had completed slightly more total years of education (*M* = 15.71, *SD* = 3.32) than the younger adults (*M* = 14.74, *SD* = 2.79), but the difference was not significant, $t(122) = 1.74$, $p = .084$. We again measured performance on the pattern comparison task (see Experiment 1) and a computerized version of the English vocabulary task (part B) by Ekstrom et al. (1979) in both age groups. Similar to Experiment 1, younger adults performed better in the pattern comparison task (*M* = 46.26 correct out of 60 patterns, *SD* = 7.30) than older adults (*M* = 34.23 correct out of 60 patterns, *SD* = 7.74), $t(122) = 8.91$, $p < .001$, $d = 1.60$, however, worse on the vocabulary task, *M* = 41% (*SD* = 15%) for younger adults, versus *M* = 65% (*SD* = 20%), $t(122) = -6.64$, $p < .001$, $d = 1.58$. The observed age differences are comparable to previous lab research with US samples of

[5] We first piloted 30 participants to make sure that the experiment is properly working and participants' memory level was not at floor. We did not perform any hypothesis tests with these data and did not include it into the final sample.

younger and older adults employing these tasks (Kuhlmann & Touron, 2016). We additionally measured participants' perceived nearness to death, which was low in general (M = 2.35, SD = 1.17 for older adults; M = 2.15, SD = 1.14 for younger adults on a 5-point ratings scale) and surprisingly did not differ across age groups, $t(122) = 1.01$, $p = .316$.

Materials

Forty-five scenery pictures (15 per valence category) were taken from the standardized picture-database OASIS (Kurdi et al., 2017). As previously, we controlled for arousal and absolute valence level: We made sure that all three source valence categories were, on average, matched on a low arousal level and that the positive and negative categories additionally matched on average absolute valence (see Appendix, Table A1). Words to serve as items were drawn from the *Affective Norms for English Words* (ANEW; Bradley & Lang, 2017), which contains English words that are normed for valence and arousal. Sixty nouns of neutral valence (]4.5; 5.5[on a rating scale rating from 1 [negative] to 9 [positive]), and low arousal (≤ 3.7 on a 9-point rating scale, with higher values indicating higher arousal levels), were chosen for this study. These words were distributed on four lists (à 15 words each), matched in mean valence, arousal, and imageability[6] using the R package "anticlust" (Papenberg & Klau, 2020). The lists were randomly assigned to serve as target sets versus the distractor set in the experiment.

Procedure

The procedure was the same as in Experiment 1 with the following exceptions: We used Prolific's built-in screening filters to approach our eligibility criteria. However, we additionally checked our specific exclusion criteria in a self-report demographic survey at the beginning of the experiment. If a participant was not eligible, the program terminated, and

[6] As the ANEW does not contain imageability ratings, we obtained these from Brysbaert et al. (2014).

participants received partial reimbursement for their time spent on the screening questions. In the study phase, participants were presented with 45 neutral words and 45 pictures (15 negative, 15 positive, 15 neutral). To better control for encoding time, the word-picture pair disappeared after 3000 ms, and the pleasantness-rating scale was presented on its own. The rating question was the same as in Experiment 1 but worded in the past tense, as the to-be-rated word was no longer visible on the screen.

The test looked the same as in Experiment 1. That is, one positive, one neutral, and one negative picture were presented as response options (alongside the new option). To realize this here, despite the many source pictures employed, we proceeded as follows: For target trials (i.e., old items), one of the three pictures was the one originally paired with the item, whereas the other two were pictures (from the remaining two valence categories) originally paired with other items. For distractor trials (i.e., new items), all three pictures were originally paired with other items in the study phase. Thus, each of the 45 pictures from the study phase appeared four times in the test phase: once as the correct source option, twice as the incorrect source option, and once with a distractor item.

As in Experiment 1, participants completed a vocabulary test after the source-monitoring test. Given the English-speaking sample, this test was changed from the 20-item German SASKA to the 18-item English vocabulary test (part B) by Ekstrom et al. (1979). The display of target items and the five response options was as in Experiment 1 but in addition, a sixth option "skip the item" was given as in the original task. Further, following the original task instructions, participants' time on this test was limited to 4 minutes in total; a clock showing the elapsed time was displayed at the top of the vocabulary test screen. In addition to answering the questions on display/legibility problems and use of memory aids (the other two questions from Experiment 1 were dropped), participants indicated their agreement with the statement "I have the feeling that my time is coming to an end" (Lang, 2000, p. 162) on a 5-point rating scale ranging from strongly disagree (1) to strongly agree (5). With this, we aimed at measuring participants' time horizon, which according to SST, underlies age-related changes in socio-emotional processing. Participants were then debriefed and reimbursed via Prolific.

Results and Discussion

Alpha level was fixed to $\alpha = .05$ for all analyses.

Pleasantness Ratings

Mean pleasantness ratings are displayed in Figure 1 (right-hand plot). Pleasantness ratings were submitted to a 2 (age group) × 3 (source valence) mixed ANOVA. There was a main effect of age, $F(1, 122) = 19.22$, $p < .001$, $\eta^2_p = .14$, and a main effect of source valence, $F(1.19, 145.50) = 239.79$, $p < .001$, $\eta^2_p = .66$, but no age × source valence interaction effect, $F(1.19, 145.50) = 1.17$, $p = .291$. As evident in Figure 1 (right-hand plot), older adults rated word pleasantness overall higher compared to younger adults, independent of source valence, in line with an age-related positivity effect. Bonferroni-Holm adjusted pairwise comparisons for the source valence levels revealed that words paired with the positive source were rated more pleasant compared to words paired with the neutral source, $t(122) = 13.81$, $p < .001$, $d = 1.24$, and also more pleasant compared to words paired with the negative source, $t(122) = 16.24$, $p < .001$, $d = 1.46$, and words paired with the neutral source were rated more pleasant than words paired with the negative source, $t(122) = 14.80$, $p < .001$, $d = 1.33$, again suggesting that both, younger and older adults incorporated source valence in their pleasantness ratings.

Overall, these results mostly replicate the findings of Experiment 1, with the only difference that in Experiment 1, we found an age-related positivity effect only when comparing pleasantness ratings for the negative source (source valence × age interaction), whereas, in Experiment 2, we found a more general positivity effect (though descriptively, the effect seemed to be more pronounced for neutral and negative sources, see Figure 1). This stronger effect might be due to the use of multiple pictures per source valence, which presumably reduced habituation effects that might have occurred in Experiment 1.

Source Memory

We used the same sub-model of the 2HTSM for three sources (Keefe et al., 2002) as in Experiment 1 to obtain measures for item and source memory. Parameter estimation, evaluation of fit, and parameter difference tests were performed with the software multiTree (Moshagen, 2010). We again estimated the 2HTSM model separately in each age group and performed pairwise comparisons of the parameters (e.g., $d^{positive}$ versus $d^{negative}$) within each age group.

The model fit the data well, $G^2(5) = 5.22$, $p = .390$ for younger adults, and $G^2(5) = 5.02$, $p = .414$ for older adults. The resulting parameter estimates for each age group are listed in Table 1; the observed source memory pattern is additionally visualized in Figure 3 (right-hand plot). As evident from the 95% CIs presented in Table 1, all memory parameters were well above chance (i.e., CI not including 0) in both age groups. We first tested for the presence of general emotionality effects within each age group. For the younger group, source memory for the positive sources was again better compared to source memory for the neutral sources, $\Delta G^2(1) = 5.85$, $p = .016$. The descriptive pattern was the same for the negative-neutral comparison (negative > neutral; see Figure 3, right-hand plot), however not significant, $\Delta G^2(1) = 1.70$, $p = .193$. Furthermore, source memory for negative and positive sources did not differ, $\Delta G^2(1) = 1.19$, $p = .275$. For a more powerful test of emotionality effects, we again set $d^{positive}$ and $d^{negative}$ equal (i.e., $d^{valanced}$) and compared this joint estimate against $d^{neutral}$. This revealed that source memory was better for emotional compared to neutral sources, $\Delta G^2(1) = 4.67$, $p = .031$.

In contrast, for the older group, there was no difference between source memory for the positive versus neutral sources, $\Delta G^2(1) = 0.02$, $p = .881$, and no difference between source memory for the negative versus neutral sources, $\Delta G^2(1) = 0.20$, $p = .657$. This again speaks against a general emotionality effect in older adults' source memory (i.e., better source memory for emotional over neutral sources). Similarly, there was no difference between the positive and negative sources, $\Delta G^2(1) = 0.09$, $p = .762$, ruling out a positivity bias. For a more powerful test of emotionality effects, we set $d^{positive}$ and $d^{negative}$ equal (i.e., $d^{valanced}$) and compared this joint estimate against $d^{neutral}$. Again, there was no difference between parameters, $\Delta G^2(1) = 0.11$, $p = .739$. In total, we neither found any evidence for emotionality effects nor a positivity bias in source memory for older adults, fully replicating the results of Experiment 1.

To test for the typical age-related source memory deficit, we compared source memory across age groups by equalizing *d* parameters of the same valence across groups (e.g., $d_{negative\_YA} = d_{negative\_YA}$). As expected, source memory was significantly lower in the older compared to the younger group, $\Delta G^2(3) = 69.28$, $p < .001$. Different from Experiment 1, the deficit manifested for all three types of sources in this experiment, all ΔG^2s$(1) \geq 13.09$, *p*s $< .001$.

Overall, the results for source memory fully replicate the findings of Experiment 1 for both age groups: We again did not find a positivity effect nor a general emotionality effect in older adults' source memory, this time ruling out habituation effects. In contrast, for younger adults, we again found enhanced source memory for positive compared to neutral sources for the pairwise comparison and a general emotionality effect when estimating source memory jointly for the negative and positive sources. Interestingly, similar to Experiment 1, the pairwise comparison between source memory for negative versus neutral sources was not significant, though descriptively present. The observed numerical difference (Δd =.09) was, however, smaller than in Experiment 1 (Δd =.13), on which we based our power analysis. Our study was thus not a priori tailored to find such a small effect. Albeit not significant, the persistent (descriptive) pattern (positive > negative > neutral) across experiments speaks for acertain consistency and raises the question of why the emotional source memory benefit is higher for positive than for negative sources in younger adults. Notably, previous research has shown that positive emotion (in contrast to negative emotion) generally benefits associative memory in an associative cued-recall paradigm (i.e., item-item binding; Madan et al., 2019). This effect might be due to the broader attentional and cognitive scope associated with a positive emotional state (Fredrickson & Branigan, 2005). However, more research is needed to corroborate that positive emotion not only benefits item-item-binding but also item-source-binding, at least in younger adults (but see Symeonidou & Kuhlmann, 2022 for a null-effect of positive sources on source memory).

General Discussion

The general aim of our research was twofold. First, we wanted to investigate whether older adults' source memory is better for (inherently) emotional sources compared to neutral sources. Furthermore, we wanted to examine whether older adults' processing preference for positive over negative stimuli, which translates into a positivity effect in item memory, generalizes to source memory. In two experiments, we used emotional pictures drawn from the OASIS database to manipulate source valence (while matching arousal levels). In Experiment 1, we applied the standard many-to-one-mapping procedure of source monitoring, using one source picture per valence category. In Experiment 2, to counteract potential habituation effects and strengthen the effectiveness of our emotionality manipulation, we applied a one-to-one-mapping procedure, using several source pictures per valence category (thus pairing each item with a unique source picture). Across both experiments, we did *not* find evidence for any emotionality effect in older adults' source memory. Also, there was no evidence of a specific age-related positivity bias or positivity effect in source memory. This was true despite evidence for positivity effects in older adults' pleasantness ratings and the same source material eliciting an emotionality effect in younger adults' source memory, especially for the positive source, in both experiments.

Before discussing potential explanations for the absence of an emotionality effect, and specifically a positivity effect, on older adults' source memory, we first must consider whether there indeed is no such effect or whether limitations in our study design confounded the results. One potential objection might be that the item-related orienting task during encoding was too directive and thus counteracted participants' encoding preferences. In fact, the meta-analysis of Reed et al. (2014) suggests that experimentally imposed constraints on information processing, even incidental learning instructions, might reduce the positivity effect. In our experiments, the pleasantness ratings put primary focus on the neutral items. To ensure a certain focus on the sources as well, we worded the pleasantness questions such that it directed participants to consider the whole screen (including the source) when providing their rating, not only the item. This was motivated by our pilot study (see OSF link), in which we had phrased the pleasantness question neutrally and found that participants focused only on the items and did not attend to the sources resulting in floor/chance-level source memory. Indeed, this change was successful as results on the pleasantness ratings across both experiments suggest that older and younger adults not only attended to the source pictures and considered their emotionality for their ratings, but they did this in a way that seemed to match their processing priorities (i.e., greater influence of the negative sources on younger relative to older adults in Experiment 1; generally higher pleasantness ratings given by older relative to younger adults in Experiment 2). This shows that, although our instructions were to a certain degree directive, participants still had enough freedom to follow their encoding preferences. Despite this (directed) processing of sources according to participants' preferences, older adults did not show corresponding source memory benefits.

Another potential objection might be that our source valence manipulation was not effective. We relied on previous norms in selecting the source material (Kurdi et al., 2017), but the selected pictures may have been perceived differently in the present paradigm and/or in the recruited samples. This problem particularly pertains to Experiment 1, where only one picture was used per source: Source emotionality was fully dependent on the perceived emotionality of the picture constituting the source. Thus, individual deviations from the norms had a greater impact on the success of the emotionality manipulation. Although we checked (and confirmed) manipulation success by collecting valence and arousal ratings in Experiment 1, we admit that these ratings could have been biased by demand characteristics and habituation effects due to the small number of used pictures. Thus, to make the success of the emotionality manipulation less dependent on one specific picture, we used several pictures per source valence in Experiment 2. This made sure that individual norm deviations for one specific picture of a source category are negligible because the other pictures of this category can compensate for it. Further, the use of several emotional pictures considerably reduces the risk of habituation effects, making the emotionality induction stronger. Yet, even under these improved conditions, we did not find emotionality effects in older adults' source memory, replicating the results of Experiment 1.

One might further argue that our older adults were relatively young (especially in Experiment 1) and healthy (all older participants in Experiment 1 and 82% in Experiment 2 indicated their health to be good or excellent), making their time horizon still quite broad. This was corroborated by participants' subjective nearness-to-death ratings in Experiment 2, which were relatively low and did not differ across age groups. This could have undermined a positivity effect because, according to SST, the effect is based on a reduced time horizon rather than age per se (though both are naturally correlated; Carstensen, 2006). Considering, however, the subjective nature of these ratings and the fact that a positivity effect was present in participants' pleasantness ratings, these nearness-to-death estimates should not be overemphasized. Further note that our recruited samples are comparable in terms of mean age and healthiness to the majority of previous studies on the age-related positivity effect (Reed et al., 2014). As such, despite comparable sample characteristics, an age-related positivity effect seems to arise in item memory (Kensinger, 2008; Reed et al., 2014) but not in source memory, suggesting that the effect is at best weaker (if at all present) for source compared to item memory.

Taken together, the employed experimental material proved effective in inducing a positivity bias (Experiment 1) and effect (Experiment 2) in older adults' pleasantness ratings. Further, the material consistently affected younger adults' ratings and source memory. Placed in this context, our experiments seem to provide specific evidence against a positivity (or a general emotionality) effect in older adults' source memory.

Explanations for the Absent Positivity Effect in Older Adults' Source Memory

Why didn't the age-related positivity effect for the pleasantness ratings translate to improved source memory? We think that due to older adults' pronounced difficulty in binding (Naveh-Benjamin, 2000; Old & Naveh-Benjamin, 2008), even a greater elaboration on the positive source might be insufficient to boost source memory. Put differently, even if older adults focused more on the positive source picture(s), they still might have had difficulties encoding the item-to-source *association* (i.e., binding problem), which ultimately is crucial for intact source memory. In line with this explanation, Nashiro and Mather (2011) found emotional arousal to improve only younger adults' but not older adults' memory binding for picture pairs. Likewise, although Davidson et al. (2006) found that negative emotionality improved older adults' source memory, this improvement was reduced compared to that seen in younger adults, again suggesting that emotion is less beneficial in improving older adults' binding. Notably, this explanation can also explain the seemingly stark contrast of our results with the results of May et al. (2005) and Rahhal et al. (2002), who found emotional source information to more strongly boost older than younger adults' source memory: In both studies, the authors used neutral sources (e.g., location of a food item in May et al., voice of a spoken statement in Rahhal et al.) but manipulated the emotionality of the *link* between the (neutral) source and (neutral) item via instructions (i.e., source indicated safety of the presented food item in May et al., or trustworthiness of the presented statement in Rahhal et al.). Put differently, the source feature carried affective information about the item (i.e., safety in May et al.; trustworthiness in Rahhal et al.), which potentially made it easier for older adults to remember the item-to-source link. Indeed, it has been shown that older adults source memory can benefit from encoding strategies providing a mediator to link the item to the source (Kuhlmann & Touron, 2012); more generally, such mediators can improve older adults' associative memory (Dunlosky & Hertzog, 2001). In a future study, one could test whether combining an emotional mediator for the item-to-source link with an additionally inherently emotional source leads to even further enhancements than the improvements observed for the inherently neutral source in May et al. (2005) and Rahhal et al. (2002). Thus, there may be some advantage in older adults' source memory for inherently emotional sources, but only if they receive sufficient aid in encoding these source-item associations. It should, however, be considered that such a conceptual emotional source-to-item mediator, as implied by May et al. (2005) and Rahhal et al. (2002), may induce participants to encode the source as an intrinsic feature of the item (i.e., item-source-unitization; Bastin et al., 2013). If so, the observed emotionality effect may rather be on familiarity-based item memory as opposed to recollection-based source memory (cf. Diana et al., 2008).

The current results thus show that although prior studies found improved memory for emotional sources in older adults (May et al., 2005; Rahhal et al., 2002), this effect seems to be bound by specifics about the emotional source manipulations employed in these studies. Neither of these previous studies used standardized emotional source stimuli drawn from a picture database normed for valence and arousal. Employing such standardized emotional source stimuli, no systematic influence of inherent source valence on older adults' source memory could be established. Further crucial, there is to date no evidence specifically for a positivity effect in older adults' source memory. Taken together, it seems that older adults are not well able to benefit from emotionality for enhancing recollection-based source memory.

Conclusion

This research aimed at investigating age-related emotionality effects, and in particular the positivity effect, in source memory by using inherently emotional sources and applying more advanced statistical tools to measure source memory separately from item memory and guessing biases. Although an age-related positivity effect occurred in participants' processing (i.e., pleasantness ratings) in both experiments, this effect did not transfer to source memory for older adults. That is, across both experiments, our results suggest better source memory for emotional compared to neutral sources only in younger but *not* in older adults. Although the absence of emotionality effects and especially a positivity effect in older adults was somewhat surprising, we believe that our experiments, using standardized emotional stimuli as sources, point out the need to re-evaluate the specific source valence manipulations in previous studies (May et al., 2005; Rahhal et al., 2002), which found enhanced emotional source memory in older adults. It seems likely that the specific emotional source manipulations of those studies facilitated encoding of the item-to-source association and that source emotionality benefits can only arise with such encoding facilitation, given older adults' pronounced difficulties with building associations in memory.

Acknowledgments and Funding

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant KU3329/1-1.

Data Availability Statement

The data that support the results reported in this manuscript are publicly available in the Open Science Framework repository (OSF) at https://osf.io/9suqj/?view_only=fe02f43d968e4ccc9f2e8938682cbcc5.

Disclosure of Interest

We have no financial or non-financial competing interests that might influence the results reported in this manuscript.

References

Bastin, C., Diana, R. A., Simon, J., Collette, F., Yonelinas, A. P., & Salmon, E. (2013). Associative memory in aging: The effect of unitization on source memory. *Psychology and Aging*, *28*(1), 275–283. https://doi.org/10.1037/a0031566

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 197–215. https://doi.org/10.1037/0278-7393.22.1.197

Bradley, M.M. & Lang, P.J. (2017). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings. Technical Report C-3.* Gainesville, FL:UF Center for the Study of Emotion and Attention.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science (New York, N.Y.)*, *312*(5782), 1913–1915. https://doi.org/10.1126/science.1127488

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously. A theory of socioemotional selectivity. *The American Psychologist*, *54*(3), 165–181. https://doi.org/10.1037//0003-066X.54.3.165

Carstensen, L. L., & Mikels, J. A. (2005). At the Intersection of Emotion and Cognition. *Current Directions in Psychological Science*, *14*(3), 117–121. https://doi.org/10.1111/j.0963-7214.2005.00348.x

Chalfonte, B. L., & Johnson, M. K [M. K.] (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*(4), 403–416. https://doi.org/10.3758/BF03200930

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

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 730–740. https://doi.org/10.1037/0278-7393.34.4.730

Driscoll, I., Hamilton, D. A., Petropoulos, H., Yeo, R. A., Brooks, W. M., Baumgartner, R. N., & Sutherland, R. J. (2003). The aging hippocampus: Cognitive, biochemical and structural findings. *Cerebral Cortex (New York, N.Y.: 1991)*, *13*(12), 1344–1351. https://doi.org/10.1093/cercor/bhg081

Dunlosky, J., & Hertzog, C. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition*, *29*(2), 247–253. https://doi.org/10.3758/BF03194918

Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monographs*. *79*(2), 3–84.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion*, *19*(3), 313–332. https://doi.org/10.1080/02699930441000238

Glisky, E. L, Rubin, S. R., & Davidson, P. S. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning,* *Memory, and Cognition*, *27*(5), 1131–1146. https://doi.org/10.1037//0278-7393.27.5.1131

Grühn, D., & Smith, J. (2008). Characteristics for 200 words rated by young and older adults: Age-dependent evaluations of German adjectives (AGE). *Behavior Research Methods*, *40*(4), 1088–1097. https://doi.org/10.3758/BRM.40.4.1088

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2021). Lab.Js: A free, open, online study builder. *Behavior Research Methods, 54*(2), 556-573. https://doi.org/10.3758/s13428-019-01283-5

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

Keefe, R. S., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source-monitoring deficits for self-generated stimuli in schizophrenia: multinomial modeling of data from three sources. *Schizophrenia Research*, *57*(1), 51–67. https://doi.org/10.1016/S0920-9964(01)00306-1

Kensinger, E. A. (2008). Age differences in memory for arousing and nonarousing emotional words. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *63*(1), P13-8. https://doi.org/10.1093/geronb/63.1.P13

Kuhlmann, B. G., & Boywitt, C. D. (2016). Aging, source memory, and the experience of "remembering". *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *23*(4), 477–498. https://doi.org/10.1080/13825585.2015.1120270

Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(5), 1352–1364. https://doi.org/10.1037/a0027863

Kuhlmann, B. G., & Touron, D. R. (2016). Aging and memory improvement through semantic clustering: The role of list-presentation format. *Psychology and Aging*, *31*(7), 771–785. https://doi.org/10.1037/pag0000117

Kuhlmann, B. G., & Undorf, M. (2018). Is all metamemory monitoring spared from aging? A dual-process examination. *Psychology and Aging*, *33*(8), 1152–1167. https://doi.org/10.1037/pag0000318

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods*, *49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Lang, F. R. (2000). Endings and Continuity of Social Relationships: Maximizing Intrinsic Benefits within Personal Networks when Feeling Near to Death. *Journal of Social and Personal Relationships*, *17*(2), 155–182. https://doi.org/10.1177/0265407500172001

Leiner, D. J. (2019). SoSci Survey (Version 3.1.06) [Computer software]. Available at https://www.soscisurvey.de

Madan, C. R., Scott, S. M. E., & Kensinger, E. A. (2019). Positive emotion enhances association-memory. *Emotion (Washington, D.C.)*, *19*(4), 733–740. https://doi.org/10.1037/emo0000465

Mather, M., & Carstensen, L. L [Laura L.] (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences*, *9*(10), 496–502. https://doi.org/10.1016/j.tics.2005.08.005

May, C. P., Rahhal, T., Berry, E. M., & Leighton, E. A. (2005). Aging, source memory, and emotion. *Psychology and Aging*, *20*(4), 571–578. https://doi.org/10.1037/0882-7974.20.4.571

Milham, M. P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., & Cohen, N. J. (2002). Attentional control in the aging brain: Insights from an fMRI study of the stroop task. *Brain and Cognition*, *49*(3), 277–296. https://doi.org/10.1006/brcg.2001.1501

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin*, *135*(4), 638–677. https://doi.org/10.1037/a0015849

Moshagen, M. (2010). Multitree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, *42*(1), 42–54. https://doi.org/10.3758/BRM.42.1.42

Nashiro, K., & Mather, M. (2011). How arousal affects younger and older adults' memory binding. *Experimental Aging Research*, *37*(1), 108–128. https://doi.org/10.1080/0361073X.2011.536746

Naveh-Benjamin, M [M.] (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1170–1187. https://doi.org/10.1037//0278-7393.26.5.1170

Nilsson, L.-G. (2003). Memory function in normal aging. *Acta Neurologica Scandinavica. Supplementum*, *179*, 7–13. https://doi.org/10.1034/j.1600-0404.107.s179.5.x

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

Papenberg, M., & Klau, G. W. (2020). Using anticlustering to partition data sets into equivalent parts. *Psychological Methods, 26*(2), 161–174. https://doi.org/10.1037/met0000301

Rahhal, T. A., May, C. P., & Hasher, L. (2002). Truth and character: Sources that older adults can remember. *Psychological Science*, *13*(2), 101–105. https://doi.org/10.1111/1467-9280.00419

Ready, R. E., Santorelli, G. D., & Mather, M. A. (2017). Judgment and classification of emotion terms by older and younger adults. *Aging & Mental Health*, *21*(7), 684–692. https://doi.org/10.1080/13607863.2016.1150415

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging*, *29*(1), 1–15. https://doi.org/10.1037/a0035194

Riegel, K. F. (1967). *Der sprachliche Leistungstest SASKA*. Hogrefe.

Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology*, *8*(4), 535–543. https://doi.org/10.1037/0894-4105.8.4.535

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Schacter, D. L., Osowiecki, D., Kaszniak, A. W., Kihlstrom, J. F., & Valdiserri, M. (1994). Source memory: Extending the boundaries of age-related deficits. *Psychology and Aging*, *9*(1), 81–89. https://doi.org/10.1037/0882-7974.9.1.81

Shing, Y. L., Rodrigue, K. M., Kennedy, K. M., Fandakova, Y., Bodammer, N., Werkle-Bergner, M., Lindenberger, U., & Raz, N. (2011). Hippocampal subfield volumes: Age, vascular risk, and correlation with associative memory. *Frontiers in Aging Neuroscience*, *3*, 1-8. https://doi.org/10.3389/fnagi.2011.00002

Symeonidou, N., & Kuhlmann, B. G. (2022). Better memory for emotional sources? A systematic evaluation of source valence and arousal in source memory. *Cognition & Emotion*, *36*(2), 300–316. https://doi.org/10.1080/02699931.2021.2008323

Ventura-Bort, C., Löw, A., Wendt, J., Dolcos, F., Hamm, A. O., & Weymar, M. (2016). When neutral turns significant: Brain dynamics of rapidly formed associations between neutral stimuli and emotional contexts. *The European Journal of Neuroscience*, *44*(5), 2176–2183. https://doi.org/10.1111/ejn.13319

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods*, *41*(2), 534–538. https://doi.org/10.3758/BRM.41.2.534

Table 1

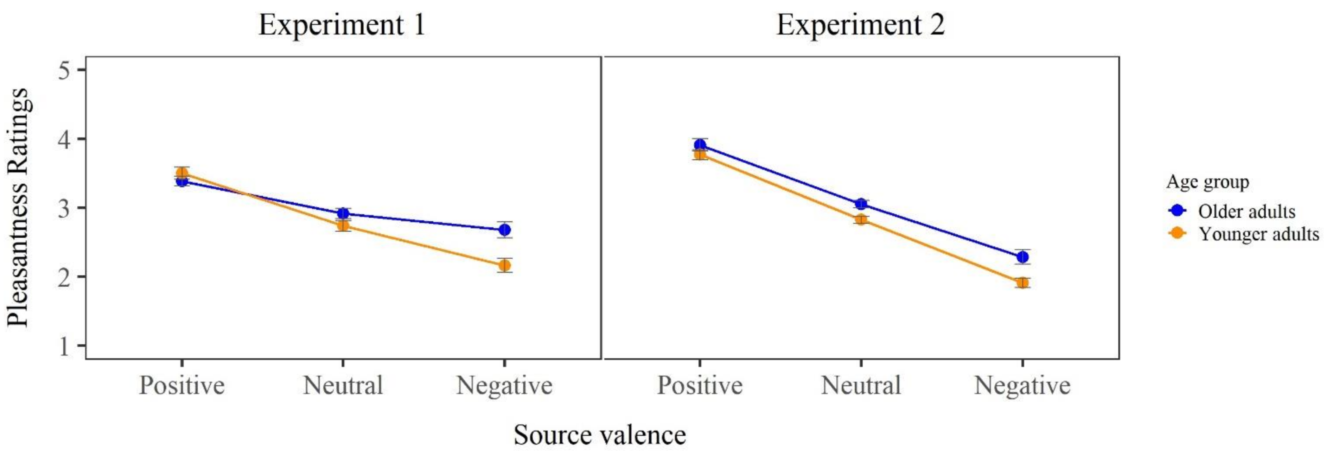
Parameter Estimates and Model Fit of the Two-High-Threshold Multinomial Model of Source Monitoring (2HTSM) for Experiment 1 and Experiment 2

	Model Fit	Parameter estimates						
Age Group	$G^2(5)$	D	$d^{positive}$	$d^{negative}$	$d^{neutral}$	b	$g^{positive}$	$g^{negative}$
				Experiment 1				
OA	8.72,	.73	.38	.37	.39	.15	.27	.50
	p = .121	[.70; .77]	[.28; .48]	[.26; .48]	[.28; .50]	[.09; .22]	[.22; .32]	[.44; .56]
YA	7.53,	.81	.57	.49	.36	.38	.22	.39
	p = .184	[.78; .84]	[.50; .65]	[.40; .58]	[.23; .49]	[.29; .47]	[.18; .26]	[.32; .45]
				Experiment 2				
OA	5.02,	.53	.40	.43	.39	.50	.32	.52
	p = .414	[.50; .56]	[.31; .50]	[.33; .52]	[.30; .49]	[.47; .54]	[.29; .35]	[.48; .55]
YA	5.22,	.49	.82	.75	.66	.56	.32	.49
	p = .390	[.46; .52]	[.72; .92]	[.65; .85]	[.55; .76]	[.53; .59]	[.28; .35]	[.45; .53]

Note. Brackets indicate 95% confidence intervals. D = probability of detecting a word as previously presented (equated across the positive, negative, and neutral source) or not presented; d^i = probability of correctly recalling the i = positive, negative, or neutral source of a recognized word; b = probability of guessing that a word was previously presented; $g^{positive}$ = probability of guessing the positive source for a detected or undetected word (.33 if unbiased); $g^{negative}$ = probability of guessing the negative (vs. neutral source) for a detected or undetected word if the positive source was not guessed (.50 if unbiased). OA = older adults; YA = younger adults.

Figure 1

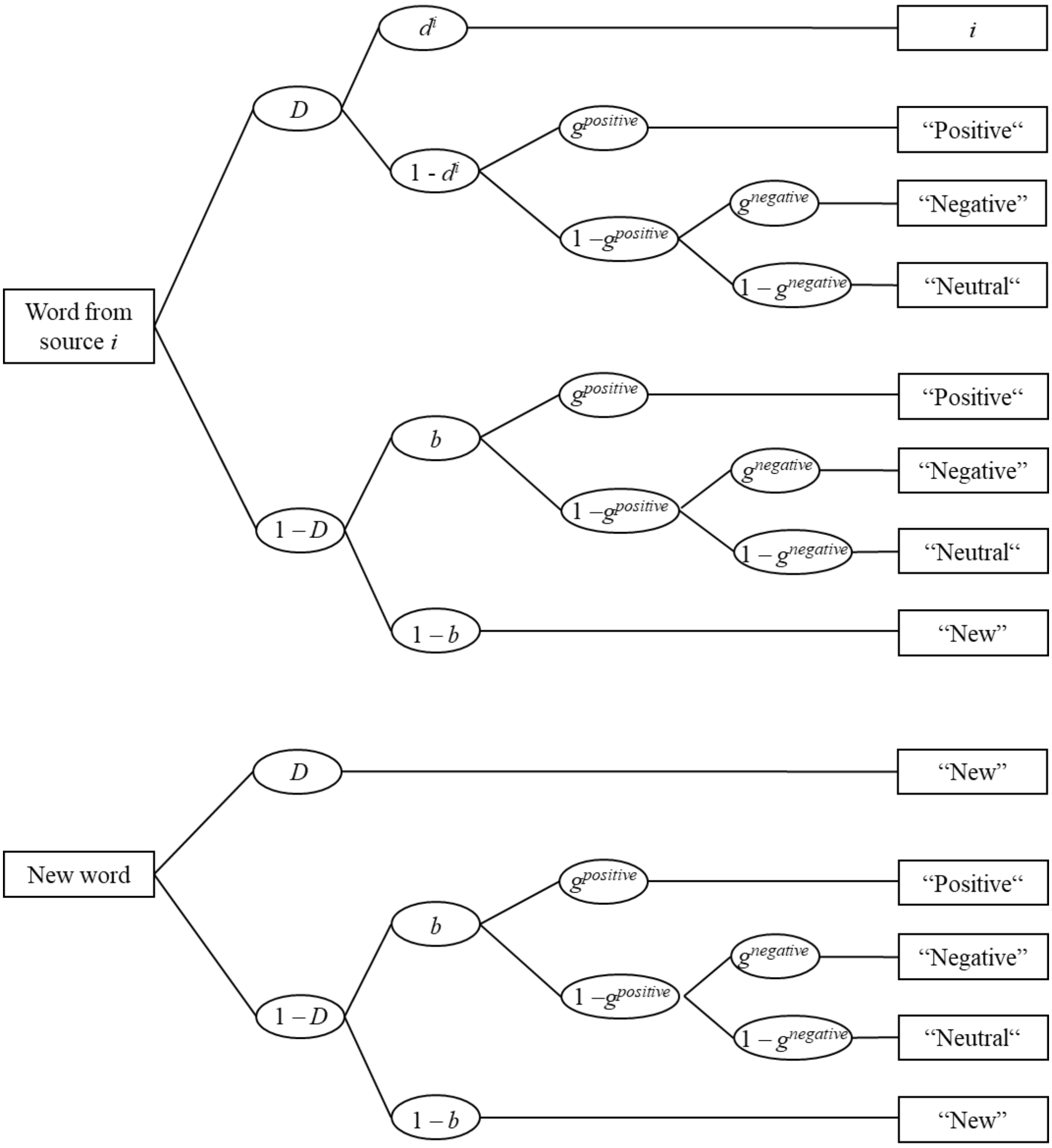
Pleasantness Ratings of Older and Younger Adults for Words Presented with the Positive, Neutral and Negative Source in Both Experiments



Note. Depicted are the mean pleasantness ratings of older and younger adults in Experiment 1 (left-hand plot) and Experiment 2 (right-hand plot) for the neutral words shown with the positive, neutral, and negative sources, respectively. Error bars indicate one standard error of the mean. Pleasantness scale ranged from 1 = very unpleasant to 5 = very pleasant. In Experiment 1, a significant main effect of age group (older adults > younger adults), a main effect of source valence (positive > neutral > negative) and an age group × source valence interaction (for negative sources: older adults > younger adults; for the other sources see main effect) occurred in pleasantness ratings. In Experiment 2, a significant main effect of age group (older adults > younger adults) and a main effect of source valence (positive > neutral > negative) occurred in pleasantness ratings.

Figure 2

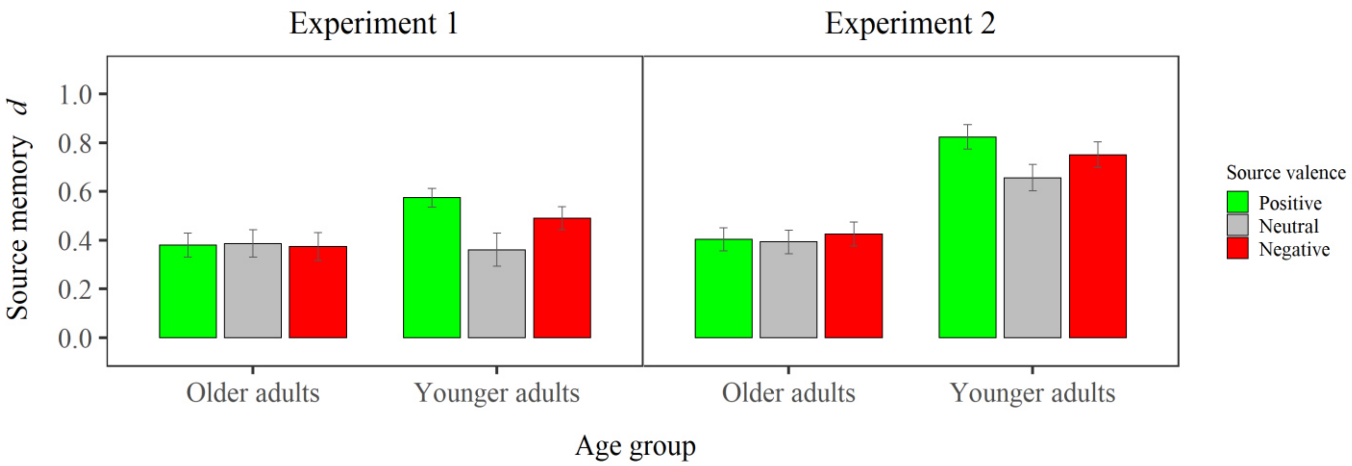
Graphical Representation of the Two-High-Threshold Multinomial Model of Source Monitoring (2HTSM) for Three Sources



Note. The figure shows sub-model 5d of the 2HTSM (Bayen et al., 1996) for target words (upper tree) and for new words (lower tree), extended to three sources. *i* denotes valence of the source the item was originally paired with, $i \in$ {positive, negative, neutral}. Boxes on the right represent participants' answers in the source memory test. *D* = probability of detecting a word as previously presented or not presented; d^i = probability of correctly recalling the source of a recognized word; *b* = probability of guessing that a word was previously presented; $g^{positive}$ = probability of guessing the positive source for a detected or undetected word $g^{negative}$ = probability of guessing the negative (vs. neutral) source for a detected or undetected word if the positive source was not guessed. Adapted from "Source monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources", by Keefe et al. (2002, p. 63).

Figure 3

Older and Younger Adults' Source Memory for the Positive, Negative, and Neutral Sources in Both Experiments



Note. The figure shows source memory performance of older and younger adults in Experiment 1 (left-hand plot) and Experiment 2 (right-hand plot), separately for the positive, negative, and neutral source valence categories. Source memory was measured by parameter *d* of the 2HTSM (Bayen et al., 1996). Error bars indicate one standard error of estimate. Across both experiments, older adults' source memory did not differ dependent on source valence, whereas younger adults' source memory was enhanced for emotional (especially positive) compared to neutral sources.

Appendix: Pictures for the Source Manipulation

Table A1

OASIS Norm Ratings for Valence and Arousal for the Source Pictures of Each Experiment

Emotionality	Picture Label	Valence	Arousal
	Experiment 1		
Positive	"Lake 3"	5.96 (0.93)	3.22 (1.97)
Negative	"Garbage dump 8"	2.06 (1.12)	3.15 (1.84)
Neutral	"Car race 1"	4.39 (1.17)	3.21 (1.83)
	Experiment 2		
Positive	"Wedding 4","Flowers 5","Bridge 1","Wedding ring 1","Food 6","Sunset 3","Sunset 1","Sunflower 1","Wedding 8","Beach 8","Lake 8","Parasailing 1","Lake 1","Rainbow 1","Flowers 9"	5.90 (1.03)	3.77 (1.77)
Negative	"Garbage dump 1", "Destruction 2","Feces 1", "Destruction 10","Jail 2","Garbage dump 4","Destruction 7","Flood 2","Car accident 4","Plane crash 3","Car crash 2","Destruction 8","Fire 3","Car crash 1","Fire 7"	2.06 (0.99)	3.86 (1.89)
Neutral	"Sun 1","Lightning 7","Research 1","Cemetery 2","Fire 4","Thunderstorm 7","Soldiers 1","Gargoyle 1","Graveyard 3","Cold 1","Desert 1","Bar 3","Woods 1","Alcohol 7","Alcohol 2"	4.18 (1.27)	3.77 (1.71)

Note. Standard deviation in brackets. Pictures were drawn from the Open Affective Standardized Image Set (OASIS; Kurdi et al., 2017). Valence and arousal scales ranged from 1 = very negative/very low to 7 = very positive/very high (with 4 = neutral/neither low nor high). All three picture-sets within an experiment were matched on arousal. The positive and negative picture-sets were additionally matched on absolute valence.

Enhanced Source Memory for Emotional Sources:

Does an Affective Orienting Task Make the Difference?

Nikoletta Symeonidou* & Beatrice G. Kuhlmann

Department of Psychology, School of Social Sciences, University of Mannheim, Germany

Author Note

Nikoletta Symeonidou https://orcid.org/0000-0002-2471-8653

Beatrice G. Kuhlmann https://orcid.org/0000-0002-3235-5717

*Correspondence concerning this article should be addressed to Nikoletta Symeonidou, Department of Psychology, University of Mannheim, D-68161 Mannheim, Germany. +49(0)621-181-3390, E-mail: nsymeoni@mail.uni-mannheim.de.

Abstract

Previous research on whether source memory is enhanced for emotional sources yielded inconclusive results. To identify potential boundary conditions, we tested whether encoding instructions that promote integrative versus segregated item-source-processing foster versus hamper source-emotionality effects. In both experiments, we used neutral words as items superimposed on emotional (negative or positive) or neutral pictures as sources. Source pictures were selected based on valence and arousal ratings collected in a pre-study, and source memory was measured with a multinomial model. In Experiment 1, we applied an affective, item-focused orienting task (OT; i.e., word-pleasantness ratings) during item-source encoding and found enhanced source memory for emotional (positive and negative) compared to neutral sources. In Experiment 2, we systematically manipulated encoding instructions and again found enhanced source memory for emotional sources with an affective OT, but no such effects with an integrative OT (item-source-fit judgments). Similarly, no effects appeared when participants were oriented towards the items only (living-non-living judgments) or instructed to learn the items only (no-OT). Source memory was surprisingly better for positive than negative sources with intentional item encoding. We conclude that source-emotionality effects might unfold only if affective item processing takes place.

Keywords: source memory, emotion-enhanced memory, orienting task, multinomial modelling

Enhanced Source Memory for Emotional Sources: Does an Affective Orienting Task Make the Difference?

A great bulk of research shows that our memory is more accurate for emotional compared to non-emotional information (see Kensinger, 2009; Mather, 2007; Talmi, 2013 for reviews). This emotion-enhanced memory (EEM) effect has been typically shown for emotional item (i.e., central) information. Results are, however, mixed with regard to EEM effects in source memory, that is, memory for the context of central information: When neutral items (e.g., words or objects) are paired with emotional versus neutral sources (e.g., background pictures), participants' memory sometimes shows a benefit for the emotional compared to neutral sources (e.g., Bell & Buchner, 2012; Ventura-Bort, Löw, Wendt, Moltó et al., 2016) but sometimes does not show such a benefit (e.g., Arnold et al., 2021; Symeonidou & Kuhlmann, 2021). Upon closer examination, it appears that most of the studies reporting EEM effects in source memory used (affective) encoding instructions/tasks that fostered integrative item-source processing, whereas those not finding an EEM effect in source memory did not use such instructions/tasks. Building on this, the goal of the herein-reported research was to investigate the role of encoding instructions for EEM effects in source memory.

When Is Source Memory Enhanced for Emotional Sources?

In contrast to the robust emotion-enhanced memory (EEM) effects for emotional items, findings regarding EEM for emotional sources are rather mixed. Buchner et al. (2009) report enhanced source memory for socially threatening compared to neutral (or trustworthy) sources across several experiments (see Bell & Buchner, 2012 for a review) when pairing neutral faces (=items) with socially threatening versus trustworthy versus neutral behavioural descriptions (=sources). Similarly, using sentences (=items) spoken by voices (=sources) with an emotional versus neutral tone, Davidson et al. (2006; Experiment 3) also established an

EEM effect in source memory. Furthermore, in Ventura-Bort, Löw, Wendt, Moltó et al. (2016), participants viewed neutral item objects superimposed on positive versus negative versus neutral scenery pictures as inherently emotional sources and also showed better source memory for the emotional (positive/negative) compared to the neutral sources (see also Smith et al., 2005 for a similar procedure and results). In contrast, Arnold et al. (2021) paired neutral faces as items with threat of shock (vs. safety vs. neutral) colour contexts as emotional sources and did not find EEM effects in source memory. Also, other studies that applied conceptual manipulations of source emotionality (e.g., safety/danger [= emotional source feature] of a neutral food item; May et al., 2005) show an EEM effect in source memory only for older, but not for younger adults (see also Rahhal et al., 2002).

Examined more closely, it appears that most of the studies that report EEM effects in source memory used material or instructions that fostered integrative item-source processing. For example, Bell et al. typically applied an affective orienting task in their studies (e.g., attractiveness or likability judgments; Bell & Buchner, 2011; Buchner et al., 2009) and additionally used sources whose interpretation could be easily integrated with the items (e.g., faces [=items] with descriptions of cheating behaviour [=sources] can be processed as cheaters). Similarly, Davidson et al. (2006) applied affective encoding instructions (emotionality judgments during encoding) and easy-to-integrate item-source material (sentences [=items] spoken by voices [=sources] in different emotional tones) and equally observed an EEM effect in source memory. In these studies, EEM effects might rely on integrative processing (elicited via the highly relatable item-source-material) or on affective processing specifically (or on both). Considering other studies, however, the integrative component appears to be already sufficient to establish EEM effects in source memory: For example, Ventura-Bort, Löw, Wendt, Moltó et al. (2016) and Smith et al. (2005) did not apply an affective encoding task but rather explicitly instructed participants to engage in integrative item-source processing (e.g., asking participants to imagine the neutral item as part of the source), finding an EEM effect in source memory. However, there are also results contradicting the idea that an integrative item-source processing (without an affective component) suffices to induce EEM effects: Similar to Bell and Buchner (2011), May et al. (2005), and Rahhal et al. (2002) also used highly relatable item-source material (dangerous vs. safe food items in May et al.; dishonest vs. honest persons in Rahhal et al.), thus fostering integrative processing, but did not find any evidence for an EEM effect in younger adults' source memory. Unlike Bell and Buchner (2011), however, these authors did not use affective encoding instructions, which might explain their null results. The role of affective processing is further emphasized in two studies from our own lab (Symeonidou et al., in press): Using unrelated item and source material (neutral words [= items] superimposed on emotional vs. neutral pictures [= sources]) and an affective orienting task during encoding (i.e., asking participants to judge the pleasantness of the neutral words), we observed better source memory for the emotional compared to the neutral sources in younger adults. This might suggest that the effect is specifically tied to affective encoding but note that an integrative explanation still cannot be ruled out: The affective orienting task in our study encouraged participants to use picture's emotionality to inform their pleasantness judgment of the neutral item. This in turn again fostered an integrative item-source processing. Thus, a more systematic manipulation of instructions (within one experiment) is necessary to dissociate the affective and integrative component and derive a clear conclusion on this matter.

In contrast, null effects of source emotionality seem to appear when items and sources are unrelated and instructions do not encourage integrative processing. For example, in two other experiments from our research group (Symeonidou & Kuhlmann, 2021), we instructed participants to learn the items only (intentional item & incidental source learning) instead of using an affective orienting task (or integrative instructions) and did not find EEM effects in source memory (Symeonidou & Kuhlmann, 2021). Similarly, Arnold et al. (2021) instructed participants to passively view face items along with their (shock vs. safety) source contexts and equally report null effects of source emotionality.

Overall, there seems to be good evidence for an EEM effect in source memory with an affective orienting task but also some evidence for an EEM effect with integrative encoding instructions, whereas unguided encoding does not seem to result in EEM. In conceiving the current experiments, we thus assumed that any (affective or non-affective) orienting task that prompts participants to relate the source to the item and thus facilitates an integrative item-source encoding should induce EEM effects. In other words, we predicted that EEM effects in source memory occur whenever people consider and integrate the emotional source while processing its item.

The Current Research

The main goal of our research was to systematically investigate whether the EEM effect in source memory depends on the type of orienting task or encoding instructions used for item-source processing. In a first step (Experiment 1), we wanted to test whether the use of an affective orienting task robustly leads to EEM effects in source memory, replicating our previous research. Then, after establishing the EEM effect with an affective orienting task, we conducted a second (preregistered) experiment to test our assumption that the EEM effect in source memory is not tied to an affective orienting task but rather occurs for any orienting task that fosters an integrative item-source encoding. More specifically, we implemented four conditions in Experiment 2: three conditions with differing orienting tasks (affective, integrative but non-affective, or non-integrative), and additionally a control condition without an orienting task (but intentional item learning).

Both experiments were conducted online via the recruitment platform Prolific (https://www.prolific.co/). In both experiments, we used neutral words as items and superimposed them on one of three background (scenery) pictures of varying valence (i.e., negative, positive, neutral) as sources. We carefully selected the three source pictures based on valence and arousal ratings, which were collected in a pre-study (with the same participants). Further, to measure source memory independent of guessing bias, we applied the two-high-threshold multinomial model of source monitoring (2HTSM; Bayen et al., 1996).

Experiment 1

We used pleasantness ratings as affective orienting task in this experiment, that is, participants judged the pleasantness of neutral words (=items) presented with emotional or neutral background pictures (=sources). We hypothesized that participants would incorporate the emotionality of the source into their pleasantness judgment of the neutral items, thus facilitating an (affective) integrative item-source processing. More specifically, we predicted that words paired with negative (positive) source pictures should receive lower (higher) pleasantness ratings compared to words paired with neutral sources. Thus, the pleasantness ratings provided a proxy for the integrative item-source processing. This affective integrative processing, in turn, should lead to EEM effects in source memory, that is, better source memory for negative/positive compared to neutral sources.

Method

Design

We used one negative, one positive, and one neutral background picture of medium (matched) arousal as sources, resulting in a simple one-factorial design (source emotionality manipulated within participants). Our dependent variables were participants' pleasantness ratings during encoding and participants' source memory, which we measured by parameter *d* of the 2HTSM (Bayen et al., 1996).

Participants

Based on an a priori power analysis conducted with the software *multiTree* (Moshagen, 2010), we aimed at $N = 70$ to find source memory differences of .15 between emotional (positive, negative) and neutral source memory parameters of the 2HTSM with a power of $1\text{-}\beta = .80$ and $\alpha = .05$. This was a conservative estimate of the EEM effect in the source memory parameters based on our previous study, which observed differences of .13 to .21 (Symeonidou et al., in press, Experiment 1). To conduct the power analysis, we estimated the emotionality effect (i.e., $\Delta d = .15$) and population parameter values based on this previous study, in which we similarly had used pleasantness ratings as an affective orienting task (i.e., $D = .81$; $b = .38$; $g_{positive} = .22$; $g_{negative} = .39$; $d_{neutral} = .36$; $d_{positive}$ and $d_{negative} = d_{neutral} + .15 = .51$). Specifically, the power analysis yielded that 4145 observations were necessary to detect a difference of .15 between source memory for the negative or positive source versus neutral source with a power of .80 and $\alpha = .05$. This corresponded to $N = 70$ participants (i.e., 70 x 60 Trials = 4200 observations). Eighty-four participants were invited to the main experiment (see section Procedure), of whom 76 actually participated. Of these, 8 were excluded because they did not fulfil pre-defined demographic eligibility criteria (i.e., German as native language [i.e., learned before the age of six]; age: 18-30 years old; student status [i.e., enrolled at a university]; no diagnosed depression and/or anxiety disorder within the past 6 months), resulting in 68 eligible participants (44 female, 1 non-binary; aged $M = 22.71$, $SD = 3.21$).

Material

We used pictures from the standardized picture-database *Open Affective Standardized Image Set* (OASIS; Kurdi et al., 2017) to manipulate source emotionality. To ensure the effectiveness of our emotionality manipulation, we first conducted a pre-study on Prolific, in which we collected valence and arousal ratings for a reasonable pre-selection of potentially suitable pictures for our source manipulation (for more details on the selection of these 45 pictures, see preregistration of Experiment 2). More specifically, we asked participants to rate 45 pictures (15 negative, 15 positive, 15 neutral) in terms of valence and arousal on a 7-point rating scale, using the OASIS rating instructions provided by Kurdi et al. (2017), translated into German. Based on these ratings, we selected one negative, one positive, and one neutral picture as sources for the main study, such that, as evident in Table 1, the negative and positive pictures were matched on absolute valence (controlling for valence strength) and all three pictures had comparable medium arousal levels (to avoid potential detrimental effects of high arousal on source memory, cf. Mather, 2007). We then invited these pre-study participants to our main experiment, thus ensuring that the source material used in the main study was tailored to the recruited sample in terms of emotionality.

For the item material in the main experiment, we used neutral words drawn from the *Berlin Affective Word List Reloaded* (BAWL-R; Võ et al., 2009). We chose 60 words of neutral valence (]1.5; 1.5[on a rating scale ranging from -3 [negative] to +3 [positive]), low arousal (≤ 2.5 on a 5-point rating scale, with higher values indicating higher arousal levels), and moderate imageability (> 3 on a 7-point rating scale, with higher values indicating higher imageability). Using the *R* package *anticlust* (Papenberg & Klau, 2020), we divided the words into four lists (à 15 words) such that they matched on mean valence, mean arousal, and mean imageability. The lists were randomly assigned to be presented with the negative, positive, or neutral source in the learning phase or to serve as distractor set in the test phase of the experiment.

Procedure

The experiment was approved by the ethics board of the University of Mannheim (the same applies to Experiment 2). The study was advertised as a two-part study (first part: pre-study with ratings, second part: main study with the source-memory task) on Prolific. We used Prolific pre-screening filters (age: 18-30; German as native language; student status; no diagnosed, ongoing mental health/ illness/ condition) to approach our eligibility criteria but additionally checked these at the beginning of the pre-study.

Both the pre-study and the main study were built in *lab.js* (Henninger et al., 2021) and hosted on the server application *OpenLab* (https://open-lab.online/). In the pre-study (first part), participants provided informed consent and answered our specific demographic and health questions to ensure full eligibility. If a participant was not eligible, they were excluded from further participation and received partial reimbursement for their time spent on the screening survey. Participants then performed a scaling task to adapt the size of the later presented OASIS pictures to their screen. More specifically, the scaling task allowed us to estimate participants’ physical screen size (in cm) and, in turn, to adjust pictures’ size such that they covered a large part of the screen, thus emphasizing their background nature (see preregistration for details). After the scaling task, participants rated 45 pictures (15 negative, 15 positive, and 15 neutral) in terms of valence and arousal. Each picture was first presented for 3750 ms and then replaced by a valence rating scale (ranging from 1: very negative to 7: very positive) which remained on the screen until the participant provided an answer. After participant’s valence rating, the arousal rating scale appeared (ranging from 1: very low to 7: very high) and remained until the participant answered. Finally, participants were asked whether there were any problems regarding the presentation of pictures and instructions (and if yes, to specify the exact problem in an open text field) and gave general feedback in an open text field if they wished.

All eligible participants were then invited to the main study (i.e., the second session). Before starting with the memory task, participants first performed the above-described scaling task again. Then, they received instructions to judge 45 words in terms of their pleasantness (i.e., affective orienting task; incidental learning). Each trial started with a 500 ms fixation cross. Afterward, one of the three source pictures (negative, positive, neutral) was presented on its own for 750 ms before a neutral word in a black-framed, white-background box was superimposed on the picture for another 3000 ms. Then, a 5-point rating scale appeared below the word-picture pair, and participants were asked to judge (self-paced) how pleasant/unpleasant they perceived the word at the present moment (1: very unpleasant to 5: very pleasant). The 45 words were equally divided between the three pictures (i.e., 15 words per picture) and presented in random order with the constraint of maximally four successive same-picture repetitions. After the study phase, participants of all four conditions completed a pattern-comparison task (Salthouse, 1996) for three minutes. In the subsequent test phase, participants were presented with all 45 words from the study phase plus 15 new distractor words. Below, the three source pictures were presented next to each other on the screen (left, centre, right; screen position was assigned randomly for each participant anew), and the option "new" was shown at the centre bottom. For each word, participants decided self-paced with which of the three pictures (negative, positive, neutral) the word was previously paired or whether it was not presented at all during the study phase (option new) by using their keyboard ("1", "2" and "3" for left, centre and right picture, space key for new items).

Finally, participants were asked whether there were any problems regarding the presentation of pictures and instructions (and if yes, to specify the exact problem in an open text field) and whether they had used any aids during the task (and if yes, to specify the used tools). They gave general feedback in an open text field, if they wished, and were debriefed about the research aim of the study.

Results and Discussion

The alpha level was fixed to $\alpha = .05$ for all analyses.

Pleasantness Ratings

A within-participants analysis of variance (ANOVA) indicated a main effect of source emotionality on participants pleasantness ratings, $F(1.69, 113.52) = 79.12$, $p < .001$, $\eta^2_p = .54$ (Greenhouse-Geisser corrected). Bonferroni-Holm adjusted pairwise comparisons revealed that, as expected, participants rated words paired with the positive source as more pleasant compared to words paired with the neutral source, $t(67) = 6.47$, $p < .001$, $d = 0.78$, and more pleasant compared to words paired with the negative source, $t(67) = 10.65$, $p < .001$, $d = 1.29$ (see also Figure 1, left-hand plot). Similarly, words paired with the neutral source were rated more pleasant than words paired with the negative source, $t(67) = 7.73$, $p < .001$, $d = 0.94$. This pattern suggests that participants used the source's emotionality to inform their pleasantness rating, indicating that they incorporated the source when processing the item.

Source Memory

For analysing source memory, we used the 2HTSM (Bayen et al., 1996), extended to three sources (Keefe et al., 2002; see Figure 2). The 2HTSM assumes that memory performance in source-monitoring tasks relies upon item recognition, source memory, and guessing processes. The model can disentangle these processes by estimating different probabilities for each process which are expressed via parameters. The probability of item memory (i.e., memory for the words) is measured in parameter D (i.e., $D_{negative}$, $D_{positive}$, and $D_{neutral}$, for items presented with the negative, positive, and neutral source, respectively). In case of a successful item recognition, the original source can also be remembered with probability $d_{negative}$ for the negative source, $d_{positive}$ for the positive source, or $d_{neutral}$ for the neutral source. If the source cannot be remembered (e.g., 1-$d_{negative}$), guessing processes take place, which are captured by the a parameters, that is, $a_{neutral}$ for the probability to guess the neutral source. If the neutral source was not guessed (with probability 1-$a_{neutral}$), guessing between the two emotional sources is captured by $a_{positive}$ for the probability to guess the positive source and, complementary, 1-$a_{positive}$ for the probability to guess the negative source. In case of failed item recognition (1-*D*), participants' answers are based on guessing processes only: If they guess that a word was previously presented in the study phase (expressed via probability *b*), they also need to guess whether it was presented with either the neutral ($g_{neutral}$) or, conditional on not guessing the neutral source (1-$g_{neutral}$), the positive ($g_{positive}$) or the negative source (1-$g_{positive}$). With the complementary probability 1-*b*, participants guess that the word is new. Relying on our previous research (Symeonidou & Kuhlmann, 2021), we assumed that item memory would be equal across sources and that source guessing does not differ for recognized versus unrecognized items. We estimated model fit via maximum likelihood estimation methods (with the software multiTree; Moshagen, 2010). Somewhat surprisingly, this submodel did not fit the data, $G^2(5) = 18.74$, $p = .002$. This was because participants showed differing source guessing biases for recognized versus unrecognized items ($a \neq g$), such that for recognized items, participants preferred to guess the neutral (compared to the negative and positive) source, whereas there was no such source guessing bias for unrecognized items (i.e., the probability for guessing one of the sources was even). Thus, we used the submodel with separate *a* and *g* parameters as baseline model, $G^2(3) = 6.61$, $p = .09$. Note that this does not compromise our analysis of source memory, as parameter *d* is estimated independently of guessing in the 2HTSM. Parameter estimates are listed in Table 2.

To test for EEM effects in source memory, we set source memory parameters equal and tested whether these equality restrictions worsen model fit significantly, which would indicate substantial source memory differences. As expected, source memory was higher for positive and negative sources compared to the neutral source, $\Delta G^2(1) = 24.76$, $p < .001$[1] for the positive-neutral comparison, $\Delta G^2(1) = 16.26$, $p < .001$ for the negative-neutral comparison, whereas the difference between the negative and positive source was not significant (see also Figure 1, right-hand plot). This pattern indicates an EEM effect in source memory, as expected.

All in all, Experiment 1 showed that, when using an affective orienting task during item-source encoding, source memory was enhanced for emotional compared to neutral sources. This replicates our previous research (Symeonidou et al., in press) and conceptually replicates Buchner et al. (2009; see also Bell & Buchner, 2011). However, as noted above, the affective orienting task not only engages affective processing but also engages integrating the item and the source because the source can be used to inform the affective item rating. We deemed it plausible that the latter type of integrative processing without the affective component already suffices to foster an EEM in source memory (see Ventura-Bort, Löw, Wendt, Moltó et al., 2016). To specifically test what type of item-source processing fosters the EEM effect in source memory, a systematic variation of encoding instructions is needed. Thus, in Experiment 2, we manipulated item-source encoding by applying different learning instructions and orienting tasks.

Experiment 2

To test more systematically which type of item-source processing results in an EEM effect in source memory, we implemented four experimental conditions (varied between participants), which differed in their learning instructions and item-source encoding: The *affective orienting task (OT) condition* replicated Experiment 1. In the *integrative OT condition,* participants judged how well neutral words fit to the background pictures, thus

[1] All reported *p*-values for the pairwise comparisons of MPT-model parameters in both experiments are Bonferroni-Holm adjusted.

explicitly fostering integrative item-source processing. In the *non-integrative OT condition,* participants judged whether the presented word rather described something living or something non-living, thus promoting a rather segregated item-source learning. Finally, in the *no-orienting-task condition* (no-OT condition), we told participants to study the words only, without explicitly referring to the emotional sources. Similar to the non-integrative OT condition, the no-OT condition fostered segregated item-source processing, but it additionally tested whether the presence of an orienting task per se might already induce EEM effects in source memory.

As to our hypotheses regarding participants' ratings in the OT conditions, we expected to replicate the effect of source emotionality on participants' pleasantness ratings in the affective OT condition (see Experiment 1). But we did not expect any impact of source emotionality on participants' item-source-fit judgments in the integrative OT condition or on their living-non-living judgments in the non-integrative OT condition. Put differently, item-source-fit and items' living-non-living status should be rated equally for negative, positive, and neutral sources. Concerning source memory, we expected to replicate the EEM effect in the affective OT condition (see Experiment 1). Assuming that this EEM effect is driven by relational item-source encoding, we also predicted the effect in the integrative OT condition. We did not expect to find any effect of source emotionality on source memory in the non-integrative OT and no-OT conditions, which promoted a segregated item-source encoding. More specifically, we predicted better source memory for negative and positive sources compared to neutral sources in the affective OT and the integrative OT conditions and no source memory differences in the non-integrative OT and no-OT conditions.

Method

Design

Participants were randomly assigned to either the affective OT, the integrative OT, the non-integrative OT, or the no-OT condition, which in turn determined the item-source learning instructions. In each condition, we used one negative, one positive, and one neutral background picture of medium (matched) arousal as sources (pictures were the same across conditions). This resulted in a four × three-mixed design with type of instructions (affective OT vs. integrative OT vs. non-integrative OT vs. no-OT) manipulated between participants and source emotionality (negative vs. positive vs. neutral) manipulated within participants.

Participants

We aimed at $n = 54$ per condition (i.e., $N = 216$ in total) to detect .25 differences between the source memory parameters with a power of $1-\beta = .80$ and $\alpha = .05$ (power analysis was conducted with multiTree; Moshagen, 2010). Our estimate for the emotionality effect (i.e., $\Delta d = .25$) was based on Experiment 1. To derive reasonable estimates for population parameter values for the remaining three conditions (i.e., integrative OT, non-integrative OT, no-OT), we piloted 10 participants per condition (these data were not used in our main analysis). The following parameter values entered the power analysis: $D = .62$; $b = .44$; $a_{neutral} = .54$; $a_{positive} = .49$; $g_{neutral} = .29$; $g_{positive} = .44$; $d_{neutral} = .13$ (see preregistration for more details).The power analysis yielded a maximum of 3230 observations to detect a difference of .25 between source memory for the negative or positive source versus neutral source with a power of .80 and $\alpha = .05$, corresponding to $n = 54$ participants (i.e., 54 x 60 Trials = 3240 observations) per condition. Two hundred sixty-five participants were invited to the main experiment (see section Procedure), of whom 222 actually participated. From these, 6 were excluded because they did not fulfil two of our preregistered criteria (i.e., had either more than 15% missing trials in the study phase [i.e., ratings] or did not use all four of the test response options at least once in the test phase), resulting in 216 eligible participants (54 female, 61 male, 1 non-binary, 1 unspecified; aged $M = 24.46$, $SD = 3.64$), that is 54 per condition, as preregistered. All participants were German native speakers and met our pre-defined demographic and health requirements (German as native language [i.e., learned before the age of six]; age: 18-30 years old; no diagnosed depression and/or anxiety disorder within the past 6 months[2]).

Material

To select our source pictures, we again conducted a pre-study using the same pre-selected 45 OASIS pictures as in Experiment 1. Based on participants' valence and arousal ratings, we selected one negative, one positive, and one neutral picture, with the negative and positive pictures matching on absolute valence and all three pictures matching on (medium) arousal level (see Table 1). All (eligible) pre-study participants were invited to our main experiment.[3] For the item material in the main experiment, the same neutral words (drawn from the BAWL-R; Võ et al., 2009) were used as in Experiment 1.

Procedure

The procedure was similar to Experiment 1 with the following exceptions. In the pre-study (first session), we made the presentation of pictures self-paced to approach the original

[2] We dropped student status (enrolled at university) as eligibility criterion because it considerably reduced the already sparse number of German native speakers on Prolific. With the student status criterion, we originally aimed at maximizing demographic similarity to our previous (lab) studies, but do not perceive this criterion as crucial for the research question at hand.

[3] As the time interval for two-part studies on Prolific is limited to three weeks and recruitment for part 1 was rather slow in our case (due to the German sample), we had to undertake two recruitment waves to achieve our desired N. In the first wave, 117 (eligible) participants completed both parts, and in the second wave 99 participants. Notably, we selected the source pictures based on the first wave of participants and used these pictures for the second wave as well. However, we first carefully inspected whether the second wave of participants rated the three source pictures comparable to the first wave in terms of valence and arousal, which was the case. Also, as apparent from Table 1, the selected sources continue to fully meet our valence and arousal criteria in the final sample.

procedure of Kurdi et al. (2017) for the OASIS ratings. More specifically, each picture was first presented for 750 ms to ensure a minimum presentation time. Directly afterward, the valence rating scale (same 7-point scale as in Experiment 1) appeared below the picture and, along with the picture, remained on the screen until the participant provided an answer. After participant's valence rating, the 7-point arousal rating scale (again as in Experiment 1) appeared and remained, along with the picture, until the participant answered.

In the main study (second session), participants received differing learning instructions dependent on the assigned condition. In the study phase of the affective OT condition, participants were asked to judge how pleasant/unpleasant they perceived the word at the present moment, as in Experiment 1. In the integrative OT condition, participants judged how well the word fit the background picture on a 5-point rating scale, ranging from 1 (doesn't fit at all) to 5 (fits very well). In the non-integrative OT condition, participants rated on a 5-point scale whether the word described something non-living (rating 1) or something living (rating 5). To match the presentation time of the word-picture pairing in the OT conditions to the study time in the no-OT condition (described below), we had to pace participants' ratings. That is, after the source picture was presented on its own for 750 ms (as in Experiment 1), the word was superimposed on the picture, and, unlike Experiment 1, the rating scale appeared simultaneously below the word-picture pair. Participants had 5 seconds to make their judgment by pressing a number from 1 to 5. If they needed less than 5 seconds for their response, the selected number turned blue to indicate that their answer was logged. The next trial then began after the 5 seconds had elapsed. In the no-OT condition, participants were instructed to memorize the neutral words for a later memory test. No explicit instructions on memorizing the word-picture pairing were given (i.e., incidental source learning). The procedure and study screen were the same as in the other three conditions, except that there was no rating scale. Thus, the presentation time for the picture-word pairing was fixed to 5 seconds in all conditions. The remainder of the procedure was identical to Experiment 1.

Results and Discussion

Ratings in the Orienting-Task Conditions

The mean OT responses by source valence and OT task condition are displayed in Figure 3. Given the different meanings of these ratings, we analysed them separately by OT condition in a simple repeated-measures analysis with source valence (positive, negative, neutral) as the only factor.

In the affective OT condition, a main effect of source emotionality on participants' pleasantness ratings occurred, $F(1.53, 81.06) = 36.62$, $p < .001$, $\eta^2_p = .41$ (Greenhouse-Geisser corrected). Replicating the results of Experiment 1, Bonferroni-Holm adjusted pairwise comparisons showed that words paired with the positive source were rated more pleasant compared to words paired with the neutral source, $t(53) = 6.19$, $p < .001$, $d = 0.84$, and more pleasant compared to words paired with the negative source, $t(53) = 6.82$, $p < .001$, $d = 0.93$, whereas words paired with the neutral source were rated more pleasant than words paired with the negative source, $t(53) = 3.87$, $p < .001$, $d = 0.53$. Again, this pattern suggests that participants incorporated the emotionality of the source into their pleasantness ratings.

In the integrative OT condition, an unexpected main effect of source emotionality on participants' item-source fit occurred, $F(1.68, 88.89) = 9.00$, $p < .001$, $\eta^2_p = .15$ (Greenhouse-Geisser corrected). Bonferroni-Holm adjusted pairwise comparisons revealed that item-source fit was rated higher for the positive source compared to the neutral source, $t(53) = 3.17$, $p = .005$, $d = 0.43$, and compared to the negative source, $t(53) = 4.44$, $p < .001$, $d = 0.60$, whereas there was no difference between negative and neutral, $t<1$. This suggests that, contrary to our expectation, it was easier for participants to see an item-source connection with the positive source.

An unexpected main effect of source emotionality also occurred in the non-integrative OT condition, $F(2, 106) = 3.59$, $p = .031$, $\eta^2_p = .06$. However, Bonferroni-Holm adjusted pairwise comparisons did not reveal any differences in participants' living-non-living judgments between the positive and negative source, $t(53) = 1.27$, $p = .238$, and between the negative and neutral source, $t(53) = 1.58$, $p = .238$, but there was a small descriptive (non-significant) difference between the positive and neutral source, $t(53) = 2.37$, $p = .065$. However, numerically, average living-non-living ratings were comparable across sources ($M = 2.49$, $SD = 0.39$ for positive, $M = 2.43$, $SD = 0.38$ for negative, $M = 2.35$, $SD = 0.36$ for neutral), suggesting that the repeated-measures ANOVA was potentially overpowered due to the large sample and thus sensitive to already small (but substantively meaningless) rating differences. To foreshadow, source memory (reported next) was at floor-level in this condition, indicating that participants did not attend to the sources in this condition, further suggesting that the effect on the pleasantness rating is probably meaningless. Thus, we do not further interpret these differences.

Memory Performance

We again used the 2HTSM to measure item memory and, crucially, source memory. This time, the most parsimonious submodel for our research question with both, the item memory restriction ($D_{negative} = D_{positive} = D_{neutral} = D_{new}$) and the source guessing restriction ($a = g$) showed good fit in all conditions, $G^2(5) = 1.67$, $p = .892$ for the affective OT condition, $G^2(5) = 6.34$, $p = .274$ for the integrative OT condition, $G^2(5) = 6.07$, $p = .299$ for the non-integrative OT condition, and $G^2(5) = 6.71$, $p = .243$ for the no-OT condition, respectively. Thus, we used this model version as the baseline model for all difference tests. Parameter estimates are listed in Table 2. We first report results on item memory before turning to source memory.

Item memory. Item memory in the affective OT and in the non-integrative OT condition was higher compared to item memory in the integrative condition, which in turn was higher than in the no-OT condition, all ΔG^2s(1) ≥ 12.85, *p*s < .001. The affective and integrative OT condition, however, did not differ, $\Delta G^2(1) = 1.11$, $p = .292$. Although we did not have specific predictions on item memory, the observed pattern across conditions seems plausible: Item memory benefited from orienting tasks that specifically focused on the item (i.e., affective OT and non-integrative OT) rather than the item-source relation (i.e., integrative OT condition). Also, incidental orienting-task instructions that promote deep processing (as used in our experiment; see also Craik & Tulving, 1975) seem to better support item recognition compared to intentional learning instructions (no-OT condition), which is in line with previous research (e.g., Postman & Kruesi, 1977).

Source memory. Mean source memory parameter estimates by source valence are plotted separately for each encoding condition in Figure 4. We first compared overall source memory levels between conditions by simultaneously equalizing the three *d* parameters of the same source emotionality across conditions (e.g., $d_{negative\_affective} = d_{negative\_integrative}$ & $d_{positive\_affective} = d_{positive\_integrative}$ & $d_{neutral\_affective} = d_{neutral\_integrative}$). This yielded that source memory was overall highest in the integrative OT condition, followed by the no-OT condition, then the affective OT condition, and lastly the non-integrative OT condition, all ΔG^2s(3) ≥ 14.92, *p*s $\leq$ .002. Again, despite no *a priori* predictions, this pattern seems plausible, as the condition with the highest item focus (i.e., non-integrative OT condition) had the lowest source memory, whereas the condition with the highest source focus (i.e., integrative OT) had the highest source memory. Interestingly, the affective and the no-OT condition both showed intermediate levels of source memory, which is in line with Experiment 1 (for the affective OT condition) and a previous study reported in Symeonidou and Kuhlmann (2021; for the no-OT condition). In the affective OT condition, this again suggests that, although the pleasantness ratings referred to the items, participants were inclined to additionally consider the sources, presumably due to their informative value for the pleasantness judgment. In the no-OT condition, the medium source-memory levels might suggest that the strategies participants used to encode the item left enough time and resources to attend to (and encode) the sources as well.

To test our hypotheses whether EEM effects in source memory depend on the type of item-source processing, we compared source memory parameters within each condition. In the affective OT condition, source memory was higher for the positive source compared to the neutral source, $\Delta G^2(1) = 12.97$, $p = .001$. Similarly, source memory for the negative source was descriptively higher than for the neutral source. However, this difference was not significant, $\Delta G^2(1) = 2.64$, $p = .104$. Thus, as expected, source memory was lower for the neutral source compared to the positive source (significantly) and also descriptively lower compared to the negative source, mostly replicating the pattern observed in Experiment 1. Different from Experiment 1, source memory for the positive source was also descriptively higher compared to the negative source, but this was not significant against a .05 α level, $\Delta G^2(1) = 4.90$, $p = .054$. Thus, an EEM effect in source memory again occurred with an affective orienting task and seemed most pronounced for the positive source.

Turning to the integrative OT condition, our analysis revealed no differences in source memory across source emotionality levels, $\Delta G^2(1) = 1.22$, $p = .538$ for the positive-neutral comparison, $\Delta G^2(1) = 0.23$, $p = .628$ for the negative-neutral comparison, and $\Delta G^2(1) = 2.89$, $p = .267$ for the positive-negative comparison. That is, contrary to our expectation, source memory for the positive and negative source was not higher compared to source memory for the neutral source, suggesting that an integrative (but non-affective) item-source encoding does not foster EEM effects in source memory.

In the non-integrative OT condition, source memory was at floor as all *d* parameters included 0 in their confidence interval (see Table 2). Accordingly, there were no differences in source memory across source emotionality levels, all ΔG^2s(1) ≤ 0.32, ps $\geq .574$. This suggests that the item-focus induced by the living-non-living judgments was too high, leading to general neglect of the sources. We originally intended this condition as a control condition to ensure that EEM effects in source memory do not generally emerge with any OT. This, however, can already be ruled out given that no EEM effects occurred in the integrative OT condition. All in all, even if the item-focus in this intended control OT was too strong, the pattern of results clearly shows that the EEM effect in source memory does not occur with any OT but rather seems specifically tied to an affective OT.

Finally, in the no-OT condition, as expected, there were no EEM effects in source memory. That is, source memory for the negative and positive source was not higher compared to the neutral source, $\Delta G^2(1) = 3.28$, $p = .140$ for the positive-neutral comparison, $\Delta G^2(1) = 2.23$, $p = .140$ for the negative-neutral comparison. Source memory again tended to be higher for the positive source, and this difference was significant when tested against source memory for the negative source, $\Delta G^2(1) = 13.29$, $p = .001$.

Taken together, we observed EEM effects in source memory only in the affective OT condition but not in the integrative OT condition, contradicting the idea that EEM effects are fostered by integrative item-source processing. Interestingly, across all conditions (except the failed non-integrative OT condition), source memory was descriptively enhanced for positive compared to the negative sources, and this difference was significant in the no-OT condition. In fact, this pattern matches the item-source-fit judgments from our integrative OT conditions: Apparently, it was easier for participants to see a connection between the item and the positive source picture. This might have facilitated the binding of the items to the positive source, leading to a specifically enhanced source memory for this source in the affective and no-OT conditions. Note, however, that it is unclear whether this effect is due to the positive valence of the source picture or due to the specific depicted picture theme, as both were confounded in our study (see also General Discussion).

General Discussion

The goal of our research was to investigate whether emotionality effects in source memory depend on the type of orienting task and instructions used during encoding. In two experiments, we used neutral words as items, which were superimposed on a negative, positive, or neutral picture as sources. In Experiment 1, we implemented an affective orienting task during item-source encoding, that is, we asked participants to rate the pleasantness of neutral items which were presented with emotional or neutral sources. Results indicated that participants incorporated the source's emotionality into their item-pleasantness rating and showed enhanced source memory for the positive and negative compared to the neutral source, suggesting an EEM effect in source memory. In Experiment 2, we wanted to test whether EEM effects in source memory more generally occur whenever integrative item-source processing is encouraged and are thus not tied to an affective orienting task. For this, we implemented an affective OT condition (see Experiment 1), an integrative OT condition (item-source-fit ratings), a non-integrative OT condition (living-non-living ratings for the item), and a no-OT condition (no orienting task, intentional item and incidental source learning). Replicating Experiment 1, participants in the affective OT condition incorporated the source's emotionality into their pleasantness ratings and showed EEM effects in source memory (significant for the positive source and descriptive for the negative source). Surprisingly, however, participants in the integrative OT condition rated item-source fit higher for the positive compared to the negative and neutral source and did not show EEM effects in source memory. In the non-integrative OT condition, source memory was at floor. In the no-OT condition, no EEM effects occurred in source memory, as predicted, but, somewhat surprisingly, source memory was enhanced for the positive compared to the negative source. Overall, the pattern of results suggests that an affective OT fosters EEM effects in source memory. Non-affective OTs, on the other hand, hinder EEM effects, even when they foster the integration of item and source. These results prompt the question of which potential mechanisms drive the effect in the affective OT condition, if not integrative item-source processing.

Explanations for the EEM Effect in the Affective OT Condition

Across both experiments and in our previous research (Symeonidou et al., in press), the EEM effect in source memory consistently occurred with an affective OT (also in line with Bell & Buchner, 2011). Whereas pleasantness (or likeability in Bell & Buchner) judgments allow integration of the source into the OT response, the comparisons to other OTs in Experiment 2 suggest that the affective nature of the OT is what drives the EEM effect in source memory. We think that the affective encoding instructions stand out from the other instructions in two main ways:

1) The affective task did not force participants to process and integrate the sources because the judgment referred to the items only. Participants had thus the freedom to follow their own preferences when processing the sources.
2) The pleasantness judgments made the emotionality (i.e., valence) of the source more salient because pleasantness directly maps onto valence. This is underpinned by participants' pleasantness ratings, which show that the source's valence informed their rating of the item. Put simply, the affective orienting question led participants to process the items in an emotional way.

Note that this can also explain the observed result patterns in the other conditions. In the integrative OT condition, the explicit integrative instructions strongly encouraged participants to engage in an integrative item-source-processing for all three sources (also the neutral source). This, in turn, led to an overall boost in source memory for all three sources. Put simply, the integrative instructions were potentially too directive and may have rather distracted attention from the emotionality of the source, thus counteracting source emotionality effects. In contrast, in the non-integrative OT condition, participants were strongly directed towards processing the items only (as the sources were irrelevant for the judgment), which resulted in poor (chance-level) source memory performance. Interestingly, source memory was at a medium level in the no-OT condition. This suggests that, although the intentional item learning instructions focused on the items, participants had enough freedom to volitionally direct their attention to their preferred sources.

Overall, our results indicate that an emotional item processing specifically might promote the unfolding of source emotionality effects. Admittedly, however, not all previous studies that found EEM effects in source memory used such affective encoding instructions (see introduction for a review). For example, source emotionality effects were also established in studies with unitization instructions (e.g., imagining the item as part of the source; Ventura-Bort, Löw, Wendt, Moltó et al., 2016), which are not per se affective. Note, however, that item-source-unitization blurs the distinction between item and source, thus substantially altering the cognitive processes that underlie a source judgment (Diana et al., 2008). As item and source become a unit, the item might adopt the emotionality of the source, making it difficult to disentangle emotionality effects on source versus item memory (see also Symeonidou & Kuhlmann, 2021 for a discussion). Thus, the emotionality effects reported in these studies most likely rely on different processes than the effects observed here. Having said that, we still deem it important that future studies investigate whether unitization instructions foster EEM effects in source (and item) memory, as, to the best of our knowledge, this has not been systematically tested to date.

Why Is Source Memory Higher for Positive Compared to Negative Sources?

What merits further discussion is the observed higher source memory for positive compared to negative sources in the affective OT condition of Experiment 2. In fact, in that condition, the EEM effect in source memory (emotional > neutral) was only significant for the positive source, and only descriptively present for the negative source. Note that the observed negative-neutral difference of $\Delta d = .10$ was considerably smaller than our (based on Experiment 1) estimated EEM effect of $\Delta d = .25$, which we had used for power analysis. Thus, Experiment 2 was underpowered to detect this smaller effect (i.e., post-hoc power to detect a difference of $\Delta d = .10$ was $1\text{-}\beta = .37$). But even considering that, this implies that the EEM effect might be reliably stronger for the positive compared to the negative source. Of note, this pattern of higher source memory for positive compared to negative sources did not only show in the affective OT (descriptively) but also in the integrative OT (descriptively) and the no-OT condition (significantly). Interestingly, the fit judgments in the integrative OT condition give a hint towards an explanation for this seemingly better memorability of the positive source in Experiment 2: Item-fit ratings were higher for the positive compared to the negative and neutral sources. This suggests that participants could more readily create an item-source connection for the positive source, boosting binding processes and thus source memory. This relatedness explanation would also account for the finding that the bias towards the positive source was less pronounced (i.e., non-significant) in the integrative OT condition than in the no-OT (and affective OT) condition: In the integrative condition, we generally encouraged participants to create an item-source-connection for all sources, thus counteracting the a priori relatedness-benefit for the positive source, whereas in the no-OT condition this benefit could unfold more fully. However, it is unclear whether positive material is generally easier relatable than negative and neutral material or whether the higher relatedness observed in Experiment 2 arose from some idiosyncratic features of the picture used as the positive source. In fact, we did not find such a bias towards the positive source in Experiment 1, in which source memory for positive versus negative sources did not differ. Note, however, using several (instead of one) positive, negative, and neutral sources in Symeonidou et al. (in press; Experiment 2) and thus ruling out idiosyncratic effects of a specific picture, we equally observed descriptively higher source memory for the positive compared to the negative source. Furthermore, Ventura-Bort, Löw, Wendt, Dolcos et al. (2016) similarly report higher relatedness for positive material: When asking participants to imagine neutral objects as part of emotional or neutral scenery pictures and querying whether or not they were successful in doing so, the authors found higher reported success for positive compared to negative pictures (but no difference to the neutral pictures). Taken together, the evidence is rather in favour of the idea that positive material is generally easier relatable, at least to neutral items as used in our research (and in Ventura-Bort, Löw, Wendt, Dolcos et al., 2016).

Of note, this higher relatedness of the positive sources is potentially not a general characteristic of positive material but rather results from participants' emotional state induced by the positive material. More specifically, previous studies have shown that a positive emotional state broadens the attentional and cognitive scope, thus fostering a more global (vs. local) processing and facilitating creative and integrative thinking (see Fredrickson & Branigan, 2005). With regard to our study, the positive sources presumably induced a positive emotional state, leading to a more global and integrative processing. This, in turn, resulted in higher item-source-fit judgments for positive sources and, more importantly, in higher source memory for positive sources in the affective and the no-OT condition, in which participants' attention was not strongly constrained by the applied instructions. In fact, there is some evidence suggesting that positive emotion fosters associative memory. Madan et al. (2019) showed that participants' associative memory for word pairs (measured with a cued recall test) was enhanced if words were both positive rather than purely neutral or mixed. Note, however, that in two other studies from our own lab (Symeonidou & Kuhlmann, 2021), we did not find enhanced source memory for positive sources (using intentional item and incidental source instructions as in the no-OT condition). Similarly, a study by von Hecker and Meiser (2005) suggests that people with a low positive trait effect (i.e., high depressiveness) do not show worse source memory than their healthy counterparts (and even better source memory for irrelevant source information). Thus, the empirical evidence on whether associative memory and source memory, in particular, is enhanced by positive emotions is rather mixed. Future studies could more directly manipulate participants' emotional state by applying mood-induction procedures, as typically done in this research area (Fredrickson & Branigan, 2005).

Limitations

As noted above, in both experiments, we followed the standard approach of a many-to-few mapping of items to sources (i.e., many items are paired with the same source; Glisky et al., 2001), thus using only one source per emotionality type (three sources in total). This confounds source emotionality with the specific picture content, meaning that the observed EEM effects might be driven by the specific picture content, not its valence (see also Symeonidou & Kuhlmann, 2021 for a discussion of this issue). However, such idiosyncratic effects are rather unlikely because we used different source pictures in Experiment 1 and Experiment 2 yet observed similar EEM effects in source memory when applying affective encoding instructions. This systematic dependence of EEM effects on the type of instructions cannot be sufficiently explained by idiosyncratic (picture-specific) effects. Furthermore, in a previous study (Symeonidou et al., in press), we used several pictures per emotionality type (one-to-one mapping of items to sources) and equally found EEM effects in source memory with affective encoding instructions. As the many-to-few-mapping (vs. one-to-one-mapping) procedure more readily taps into item-source binding (vs. item-to-item binding) and did not compromise the EEM effect in our previous research, we deliberately opted for this standard procedure in the herein reported research.

One might further criticize that we did not measure and control for participants' pre-experimental mood. Considering that participants' emotional state can influence their cognitive-attentional scope (Fredrickson & Branigan, 2005), a priori differences in participants' mood across conditions might have biased our observed effects. Although we cannot entirely rule out this possibility, we deem that such systematic biases are rather unlikely: As the assignment of participants to conditions was random, participants' pre-experimental mood should have been (on average) comparable across conditions. We agree, however, that future studies should measure and control for mood or, as noted above, even manipulate participants' emotional state to study its effects on associative memory, including source memory. If a positive emotional state indeed fosters source memory, then a direct manipulation could potentially amplify the effect of positive emotion and thus make it easier to find memory differences between positive and negative emotion. In our studies, the negative-positive difference in source memory was small and not significant (except in the no-OT condition), as our studies had only limited power to detect such small differences in source memory parameters, despite their decent sample size. Future studies could thus aim for a stronger manipulation of emotion.

Another point of criticism might pertain to the online nature of our studies because the experimental setting cannot be fully controlled in online studies (e.g., quiet/loud environment, used device, used aids, etc.), thus introducing the risk of systematic biases. Note, however, that we implemented different checks throughout the entire experiment (e.g., checks for browser window changes, check of each task screen's duration), which at least ensured that participants were continuously working on the experiment without interruptions.

The above-chance (source and item) memory performance further indicates that participants complied with our instructions. Also note that, using similar online control measures in a previous study (Symeonidou & Kuhlmann, 2021), we obtained comparable memory patterns in a lab and in an online setting.

Conclusion

Across two experiments, we showed that emotionality effects in source memory depend on encoding instructions: EEM effects emerged reliably when participants were oriented towards an affective item-source processing but did not occur when participants engaged in non-affective relational item-source encoding or were instructed to focus attention on the (neutral) items only. This, in part, explains the mixed findings regarding EEM effects in source memory in previous research and once more corroborates that the mere presence of emotional sources is not sufficient to induce EEM effects in source memory. Crucially, our studies stand out against previous research in that we carefully selected the source pictures in terms of valence and arousal by conducting a pre-study, thus ensuring the effectiveness of the source emotionality manipulation. Furthermore, we used statistical modelling to measure source memory free from guessing bias. Overall, our research considerably contributes to identifying important boundary conditions under which source emotionality has versus does not have beneficial effects on source memory. But it also highlights open questions and potential future directions. For example, it is yet unclear whether unitization instructions (which lead to item-source unitization) can similarly moderate EEM effects in source memory and whether these effects are then rather driven by enhanced familiarity versus enhanced recollection of emotional compared to neutral sources. Future studies could investigate this question by varying unitization instructions and applying a Remember/Know procedure to tap into familiarity versus recollection.

Acknowledgments and Funding

We thank Arndt Bröder and Monika Undorf for their helpful comments on the interpretation and discussion of the obtained results.

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grants GRK 2277 "Statistical Modeling in Psychology" and KU 3329/1-1.

Data Availability Statement

The data supporting the results reported in this manuscript are publicly available in the Open Science Framework (OSF) repository at https://osf.io/8dhmj/?view_only=de8bb8c733a74b549ec5dcc436b9af1d. The preregistration of Experiment 2 can be accessed via https://osf.io/g56ec/?view_only=98bae34f4bfa47bfb614a49f5dcb921f.

Disclosure of Interest

We have no financial or non-financial competing interests that might influence the results reported in this manuscript.

References

Arnold, N. R., González Cruz, H., Schellhaas, S., & Bublatzky, F. (2021). A multinomial modelling approach to face identity recognition during instructed threat. *Cognition & Emotion*, *35*(7), 1–18. https://doi.org/10.1080/02699931.2021.1951175

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 197–215. https://doi.org/10.1037/0278-7393.22.1.197

Bell, R., & Buchner, A. (2011). Source memory for faces is determined by their emotional evaluation. *Emotion (Washington, D.C.)*, *11*(2), 249–261. https://doi.org/10.1037/a0022597

Bell, R., & Buchner, A. (2012). How adaptive is memory for cheaters? *Current Directions in Psychological Science*, *21*(6), 403–408. https://doi.org/10.1177/0963721412458525

Buchner, A., Bell, R., Mehl, B., & Musch, J. (2009). No enhanced recognition memory, but better source memory for faces of cheaters. *Evolution and Human Behavior*, *30*(3), 212–224. https://doi.org/10.1016/j.evolhumbehav.2009.01.004

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

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

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 730–740. https://doi.org/10.1037/0278-7393.34.4.730

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion*, *19*(3), 313–332. https://doi.org/10.1080/02699930441000238

Glisky, E. L., Rubin, S. R., & Davidson, P. S. (2001). Source memory in older adults: An encoding or retrieval problem? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1131–1146. https://doi.org/10.1037//0278-7393.27.5.1131

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2021). Lab.Js: A free, open, online study builder. *Behavior Research Methods, 54*(2), 556-573. https://doi.org/10.3758/s13428-019-01283-5

Keefe, R. S.E., Arnold, M. C., Bayen, U. J., McEvoy, J. P., & Wilson, W. H. (2002). Source-monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources. *Schizophrenia Research*, *57*(1), 51–67. https://doi.org/10.1016/S0920-9964(01)00306-1

Kensinger, E. A. (2009). Remembering the Details: Effects of Emotion. *Emotion Review: Journal of the International Society for Research on Emotion*, *1*(2), 99–113. https://doi.org/10.1177/1754073908100432

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods*, *49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Madan, C. R., Scott, S. M. E., & Kensinger, E. A. (2019). Positive emotion enhances association-memory. *Emotion (Washington, D.C.)*, *19*(4), 733–740. https://doi.org/10.1037/emo0000465

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

May, C. P., Rahhal, T., Berry, E. M., & Leighton, E. A. (2005). Aging, source memory, and emotion. *Psychology and Aging*, *20*(4), 571–578. https://doi.org/10.1037/0882-7974.20.4.571

Moshagen, M. (2010). Multitree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, *42*(1), 42–54. https://doi.org/10.3758/BRM.42.1.42

Papenberg, M., & Klau, G. W. (2020). Using anticlustering to partition data sets into equivalent parts. *Psychological Methods, 26*(2), 161–174. https://doi.org/10.1037/met0000301

Postman, L., & Kruesi, E. (1977). The influence of orienting tasks on the encoding and recall of words. *Journal of Verbal Learning and Verbal Behavior*, *16*(3), 353–369. https://doi.org/10.1016/S0022-5371(77)80056-X

Rahhal, T. A., May, C. P., & Hasher, L. (2002). Truth and character: Sources that older adults can remember. *Psychological Science*, *13*(2), 101–105. https://doi.org/10.1111/1467-9280.00419

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Smith, A. P. R., Henson, R. N. A., Rugg, M. D., & Dolan, R. J. (2005). Modulation of retrieval processing reflects accuracy of emotional source memory. *Learning &*

Memory (Cold Spring Harbor, N.Y.), *12*(5), 472–479. https://doi.org/10.1101/lm.84305.

Symeonidou, N., & Kuhlmann, B. G. (2021). Better memory for emotional sources? A systematic evaluation of source valence and arousal in source memory. *Cognition & Emotion*, *36*(2), 300-316. https://doi.org/10.1080/02699931.2021.2008323

Symeonidou, N., Hassan, A., Porstein, I. & Kuhlmann, B. G. (in press). Is there an emotionality effect in older adults' source memory? *Aging, Neuropsychology and Cognition*.

Talmi, D. (2013). Enhanced emotional memory. *Current Directions in Psychological Science*, *22*(6), 430–436. https://doi.org/10.1177/0963721413498893

Ventura-Bort, C., Löw, A., Wendt, J., Dolcos, F., Hamm, A. O., & Weymar, M. (2016). When neutral turns significant: Brain dynamics of rapidly formed associations between neutral stimuli and emotional contexts. *The European Journal of Neuroscience*, *44*(5), 2176–2183. https://doi.org/10.1111/ejn.13319

Ventura-Bort, C., Löw, A., Wendt, J., Moltó, J., Poy, R., Dolcos, F., Hamm, A. O., & Weymar, M. (2016). Binding neutral information to emotional contexts: Brain dynamics of long-term recognition memory. *Cognitive, Affective & Behavioral Neuroscience*, *16*(2), 234–247. https://doi.org/10.3758/s13415-015-0385-0

Võ, M. L.-H., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin Affective Word List Reloaded (BAWL-R). *Behavior Research Methods*, *41*(2), 534–538. https://doi.org/10.3758/BRM.41.2.534

von Hecker, U., & Meiser, T. (2005). Defocused attention in depressed mood: Evidence from source monitoring. *Emotion (Washington, D.C.)*, *5*(4), 456–463. https://doi.org/10.1037/1528-3542.5.4.456

Table 1

Pre-Study Ratings for Valence and Arousal of the Source Pictures Chosen for the Main Study of Experiment 1 and Experiment 2

Experiment	Emotionality	Picture Label	Valence	Arousal
Experiment 1	Positive	"Lake 4"	5.93 (0.94)	4.31 (1.69)
	Negative	"Destruction 10"	2.31 (0.80)	4.22 (1.43)
	Neutral	"Volcano 1"	4.15 (1.12)	4.35 (1.16)
Experiment 2	Positive	"Bridge 1"	5.65 (0.91)	3.89 (1.56)
	Negative	"Garbage dump 1"	2.28 (0.94)	3.72 (1.45)
	Neutral	"Car race 4"	3.89 (1.17)	3.74 (1.56)

Note. Standard deviation in brackets. Pictures were originally drawn from the *Open Affective Standardized Image Set* (OASIS; Kurdi et al., 2017). Pictures were selected such that the negative and positive pictures matched on absolute valence, and all three pictures matched on arousal level. The reported mean ratings refer to the final sample (i.e., to the eligible participants who participated in the pre-study and the main study of the respective experiment).

Table 2

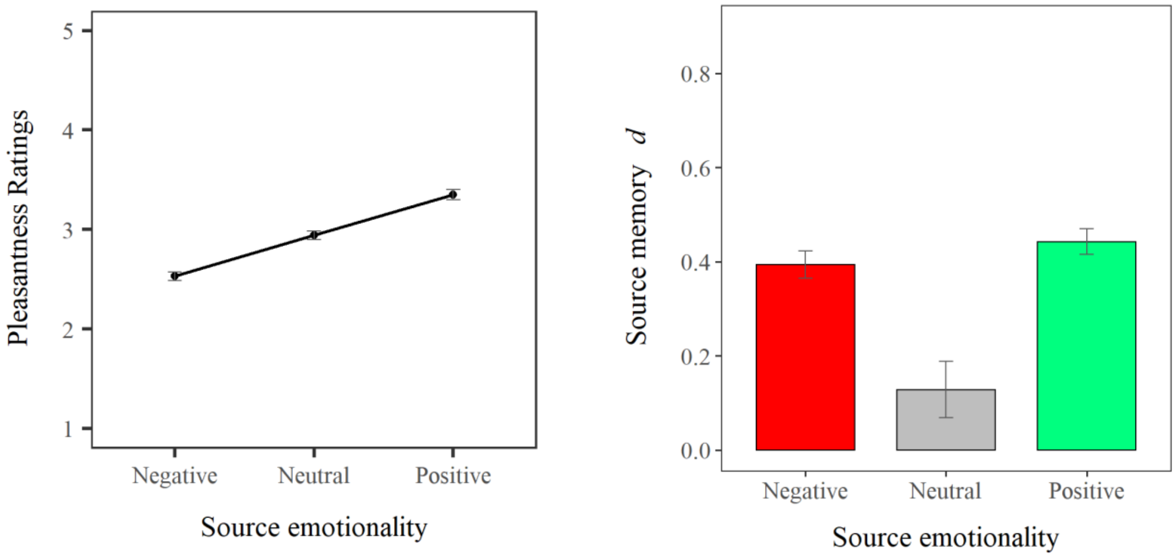
Model Fit and Parameter Estimates of the Two-High-Threshold Multinomial Model of Source Monitoring (2HTSM) for Experiment 1 and Each Condition of Experiment 2

Orienting Task/ Condition	Model Fit	Parameter estimates								
		D	$d_{positive}$	$d_{negative}$	$d_{neutral}$	b	$a_{neutral}$	$a_{positive}$	$g_{neutral}$	$g_{positive}$
	$G^2(3)$	Experiment 1								
Affective	6.61, p = .086	.78 [.76; .80]	.44 [.39; .50]	.39 [.34; .45]	.13 [.01; .25]	.40 [.35; .45]	.54 [.49; .59]	.49[.43; .55]	.29 [.19; .38]	.44 [.32; .56]
	$G^2(5)$	Experiment 2								
Affective	1.67, p = .892	.77 [.75; .80]	.44 [.38;.50]	.33 [.26;.40]	.23 [.13;.32]	.43 [.37; .49]	-	-	.44 [.40; .47]	.47 [.42; .51]
Integrative	6.07, p = .299	.79 [.76; .81]	.04 [-.04;.11]	.02 [-.05;.09]	.05 [-.03;.13]	.58 [.53; .64]	-	-	.37 [.34; .40]	.50 [.47; .54]
Non-Integrative	6.34, p = .274	.71 [.68; .73]	.75 [.70;.81]	.68 [.61; .75]	.70 [.63;.78]	.33 [.28; .38]	-	-	.41 [.36; .45]	.38 [.32; .44]
None	6.71, p = .243	.59 [.56; .62]	.55 [.47;.63]	.33 [.23; .42]	.43 [.34;.53]	.38 [.34; .43]	-	-	.40 [.36; .43]	.46 [.41; .51]

Note. Brackets indicate 95% confidence intervals. D = probability of detecting a word as previously presented (equated across the positive, negative, and neutral source) or not presented; d_i = probability of correctly recalling the i = positive, negative, or neutral source of a recognized word; b = probability of guessing that a word was previously presented; $a/g_{neutral}$ = probability of guessing the neutral source for a detected/undetected word (equated and estimated as one parameter g_{neutral} in Experiment 2); $a/g_{positive}$ = probability of guessing the positive (vs. negative) source for a detected/undetected word if the neutral source was not guessed (equated and estimated as one parameter g_{positive} in Experiment 2).

Figure 1

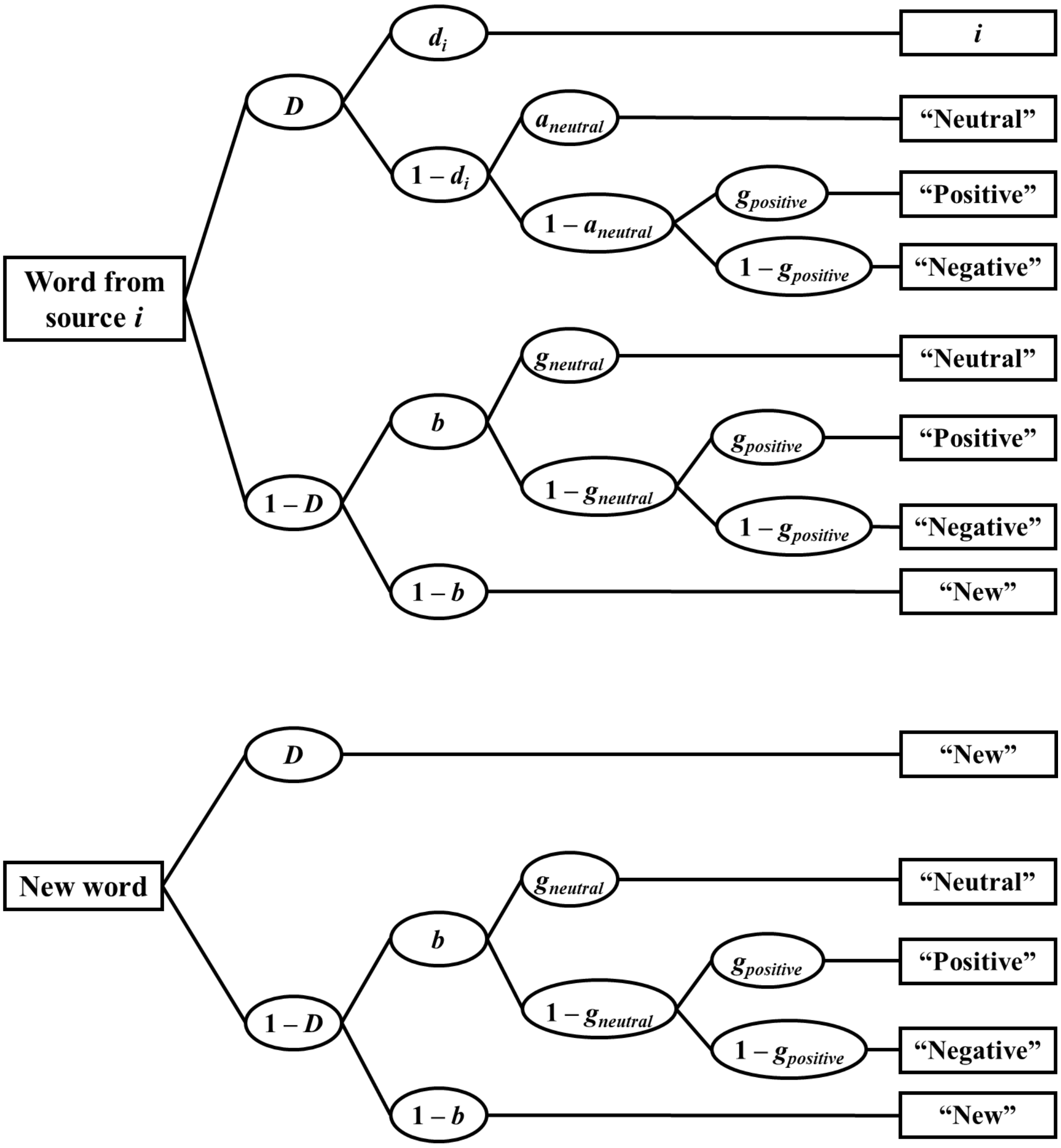
Pleasantness Ratings and Source Memory for Negative, Neutral, and Positive Sources in Experiment 1



Note. Left-hand plot: Depicted are mean pleasantness ratings as a function of source emotionality. Error bars indicate one standard error of the mean. Pleasantness scale ranged from 1 = very unpleasant to 5 = very pleasant. Right-hand plot: Depicted is source memory for the negative, neutral, and positive source. Source memory was measured by parameter *d* of the 2HTSM (Bayen et al., 1996). Error bars indicate one standard error of estimate.

Figure 2

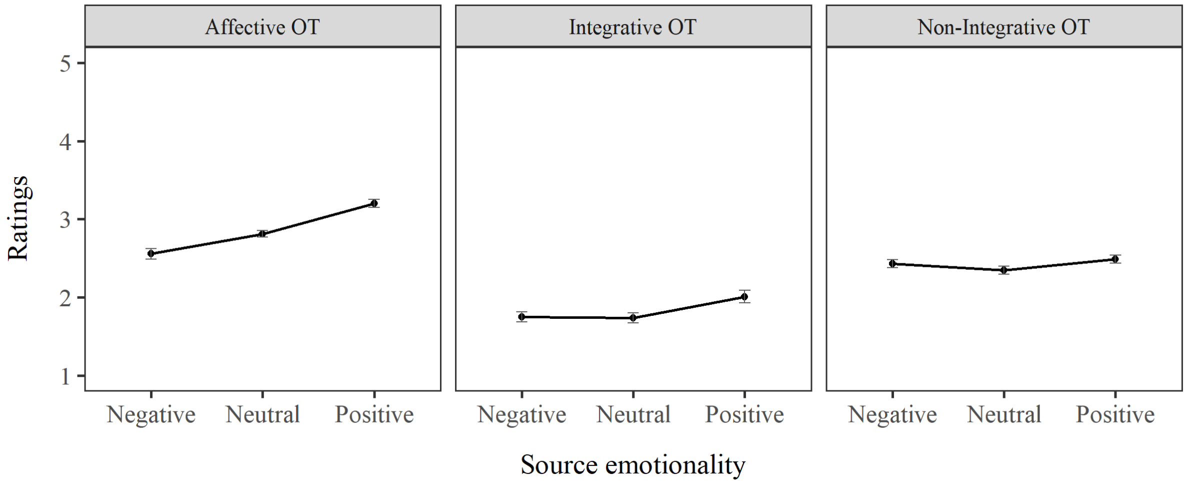
Graphical Representation of the Two-High-Threshold Multinomial Model of Source Monitoring (2HTSM) for Three Sources



Note. The figure shows submodel 5d of the 2HTSM (Bayen et al., 1996) for target words (upper tree) and for new words (lower tree), extended to three sources. *i* denotes the emotionality of the source with which the item was originally paired, $i \in$ {negative, neutral, positive}. Boxes on the right represent participants' answers in the source memory test. *D* = probability of detecting a word as previously presented or not presented; d_i = probability of remembering the source of a recognized word; *b* = probability of guessing that a word was previously presented; $a/g_{neutral}$ = probability of guessing the neutral source for a detected/undetected word (equated and estimated as one parameter $g_{neutral}$ in Experiment 2); $a/g_{positive}$ = probability of guessing the positive (vs. negative) source a for a detected/undetected word if the neutral source was not guessed (equated and estimated as one parameter $g_{positive}$ in Experiment 2). Adapted from "Source monitoring deficits for self-generated stimuli in schizophrenia: Multinomial modeling of data from three sources", by Keefe et al. (2002, p. 63).

Figure 3

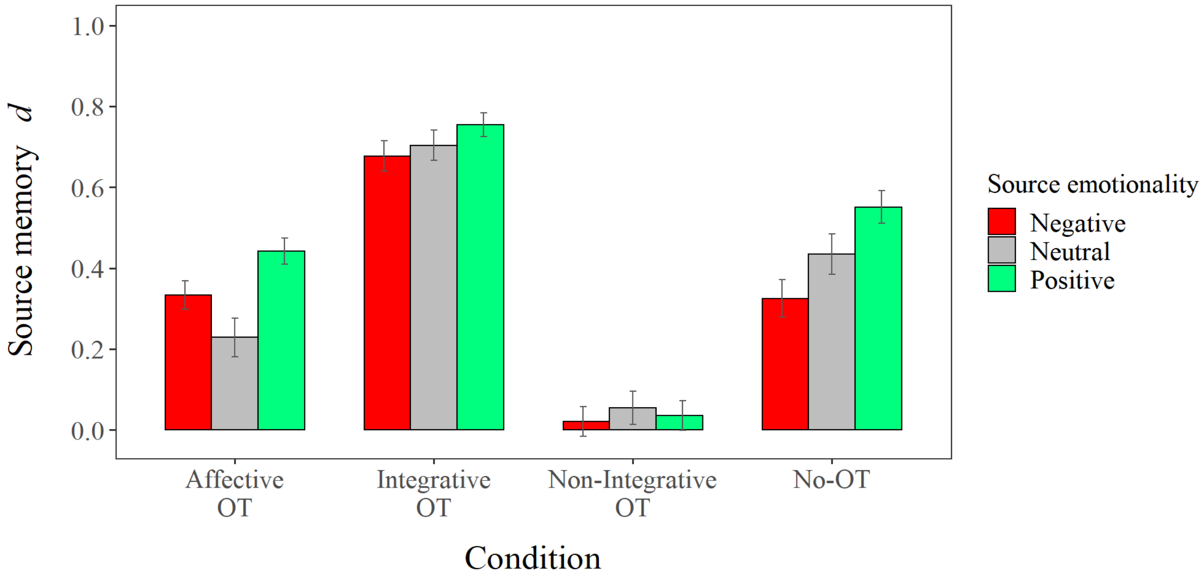
Orienting Task Ratings as a Function of Source Emotionality in Experiment 2



Note. The figure shows mean pleasantness ratings (affective orienting task [OT]; higher = more pleasant), mean item-source-fit ratings (integrative OT; higher = higher fit), and mean living-non-living ratings (non-integrative OT; higher = living) as a function of source emotionality. Error bars indicate one standard error of the mean. The rating scale ranged from 1 to 5 in all three OTs.

Figure 4

Source Memory for Negative, Neutral, and Positive Sources in Each Condition of Experiment 2



Note. Depicted is source memory for negative, neutral, and positive sources in the affective orienting task (OT), integrative OT, non-integrative OT, and no-OT condition of Experiment 2. Source memory was measured by parameter *d* of the 2HTSM (Bayen et al., 1996). Error bars indicate one standard error of estimate.

This dissertation was supported by the Research Training Group "Statistical Modeling in Psychology" (SMiP), funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation).