Cognitive, Affective, & Behavioral Neuroscience (2021) 21:172–190
https://doi.org/10.3758/s13415-020-00858-6

Interactions of Emotion and Self-reference in Source Memory: An ERP Study

Diana R. Pereira[1] · Adriana Sampaio[1] · Ana P. Pinheiro[1,2]

Accepted: 8 December 2020 / Published online: 19 February 2021

Abstract

The way emotional information is encoded (e.g., deciding whether it is self-related or not) has been found to affect source memory. However, few studies have addressed how the emotional quality and self-referential properties of a stimulus interactively modulate brain responses during stimulus encoding and source memory recognition. In the current study, 22 participants completed five study-test cycles with negative, neutral, and positive words encoded in self-referential versus non-self-referential conditions, while event-related potentials of the electroencephalogram were recorded. An advantage of self-referential processing in source memory performance, reflected in increased recognition accuracy, was shown for neutral and positive words. At the electrophysiological level, self-referential words elicited increased amplitudes in later processing stages during encoding (700–1,200 ms) and were associated with the emergence of old/new effects in the 300–500 ms latency window linked to familiarity effects. In the 500–800 ms latency window, old/new effects emerged for all valence conditions except for negative words studied in the non-self-referential condition. Negative self-referential words also elicited a greater mobilization of post-retrieval monitoring processes, reflected in an enhanced mean amplitude in the 800–1,200 ms latency window. Together, the current findings suggest that valence and self-reference interactively modulate source memory. Specifically, negative self-related information is more likely to interfere with the recollection of source memory features.

Keywords Source memory · Valence · Self-referential processing · Event-related potentials · Emotion · Recognition memory

Introduction

The effects of emotion on source memory, i.e., the ability to remember features associated with a given study event (Johnson, Hashtroudi, & Lindsay, 1993), vary across studies (see Pereira, Sampaio, & Pinheiro, 2019 for a summary table). Whereas some studies report a deleterious (Cook, Hicks, & Marsh, 2007; Ferré, Comesaña, & Guasch, 2019, Experiment 2 and 3; MacKenzie, Powell, & Donaldson, 2015; Maddock & Frein, 2009) or a beneficial effect of emotion (Kensinger & Corkin, 2003; Kensinger & Schacter, 2006), others show no differences between emotional and neutral stimuli on source memory (Ferré et al., 2019, Experiment 1; Gallant, Spaniol, & Yang, 2019; Gallant & Yang, 2014; Kensinger & Schacter, 2006; Koenig & Mecklinger, 2008; Wang & Fu, 2011). These conflicting findings may be accounted for by different factors: the way information is encoded is a relevant candidate to consider (Leynes & Crawford, 2018; Leynes & Mok, 2017). In this regard, self-referential processing has been one of the encoding strategies examined in the context of source memory by testing its influence on features, such as stimulus location and color (Yin, Ma, Xu, & Yang, 2019) and the encoding task itself (Dulas, Newsome, & Duarte, 2011; Durbin, Mitchell, & Johnson, 2017; Hou, Grilli, & Glisky, 2019; Leshikar & Duarte, 2012; Leshikar, Dulas, & Duarte, 2015; Mao, Wang, Wu, & Guo, 2017; Pereira et al., 2019; Zhang, Pan, Li, & Guo, 2018). Nonetheless, few studies have probed the interactive effects of self-referential encoding and emotion on source memory and their electrophysiological correlates. As such, the current study specified whether and how self-referential encoding (vs. non-self-referential encoding) and

✉ Ana P. Pinheiro
appinheiro@psicologia.ulisboa.pt

Diana R. Pereira
diana.r.pereira@gmail.com

Adriana Sampaio
adriana.sampaio@psi.uminho.pt

[1] Psychological Neuroscience Laboratory, CIPsi, School of Psychology, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

[2] Faculdade de Psicologia, CICPSI, Universidade de Lisboa, Alameda da Universidade, 1649-013 Lisboa, Portugal

stimulus valence (negative vs. neutral vs. positive) modulate overt encoding and source memory recognition processes, combining behavioral (performance accuracy), and electrophysiological (event-related potentials [ERPs]) measures.

Effects of emotion and self-reference during encoding: ERP evidence

During information encoding, ERPs afford excellent temporal resolution to specify when and how the processing of emotionally salient information diverges from the processing of neutral information (Citron, 2012; Hinojosa, Méndez-Bértolo, & Pozo, 2010; Schupp, Flaisch, Stockburger, & Junghöfer, 2006). Accordingly, the affective properties of a stimulus, including valence (its degree of pleasantness/unpleasantness) and arousal (the degree of calmness/excitement elicited by the stimulus), were found to modulate distinct ERP components, such as the early posterior negativity (EPN) (Citron, 2012; Schupp et al., 2006), the late positive complex (LPC) (Citron, 2012; Foti, Hajcak, & Dien, 2009; Schupp et al., 2006), and the slow wave (SW), a late positivity that follows the LPC (Foti et al., 2009; Fig. 1). The amplitude of these components tends to be consistently larger for emotional compared to neutral stimuli (Barnacle, Tsivilis, Schaefer, & Talmi, 2018; Codispoti, Ferrari, & Bradley, 2007; Diedrich, Naumann, Maier, Becker, & Bartussek, 1997; Dolcos & Cabeza, 2002; Schacht & Sommer, 2009; Schupp et al., 2006). Notwithstanding, emotion effects on later components, such as the LPC, appear to be particularly sensitive to different task conditions, including attentional focus (i.e., if participants are explicitly required to attend to the emotional properties of the stimuli) and specific task demands (i.e., if a task requires stronger perceptual vs. semantic processing; Fischler & Bradley, 2006). Specifically, emotion effects were enhanced when participants were explicitly asked to evaluate the emotional quality of the stimuli and when semantic processing was required (Diedrich et al., 1997; Fischler & Bradley, 2006; Hajcak, Moser, & Simons, 2006; Hinojosa et al., 2010; Schacht & Sommer, 2009). Therefore, the LPC might be particularly sensitive to interactive effects of emotion and self-reference on encoding processes (Gutchess & Kensinger, 2018).

In this context, whereas some studies revealed increased LPC amplitude for self-related positive stimuli (Auerbach, Stanton, Proudfit, & Pizzagalli, 2015; Herbert, Herbert, Ethofer, & Pauli, 2010; Shestyuk & Deldin, 2010), others demonstrated a negativity bias, i.e., increased LPC amplitude for self-related negative stimuli (Grundy, Benarroch, Lebarr, & Shedden, 2015; Herbert, Pauli, & Herbert, 2011; Katyal, Hajcak, Flora, Bartlett, & Goldin, 2020; Zhou et al., 2017). Several factors might account for these discrepant findings, including methodological (e.g., implicit vs. explicit self-referential tasks; stimulus properties such as self-relevance; mixed vs. blocked design; Fan et al., 2013; Hudson, Wilson, Green, Itier, & Henderson, 2020; H. Zhang, Guan, Qi, & Yang, 2013) and participant-related variables (e.g., self-esteem; depression; meditation practices; Auerbach et al., 2015; Katyal et al., 2020; Nowicka, Wójcik, Kotlewska, Bola, & Nowicka, 2018; Shestyuk & Deldin, 2010; H. Zhang et al., 2013). However, it remains unclear how explicit instructions to memorize target stimuli may affect the interactive effects between emotion and self-reference. Specifically, the use of intentional encoding ensures that emotional, neutral, self-referential, and non-self-referential conditions do not differ in task-relevance, thereby involving similar levels of prioritization and allocation of attentional resources (see Barnacle et al., 2018 for a similar argument).

So far, the existing studies using intentional learning conditions also revealed increased amplitudes of the LPC and SW for negative and positive stimuli compared with neutral ones (Barnacle et al., 2018; Dolcos & Cabeza, 2002; Palomba,

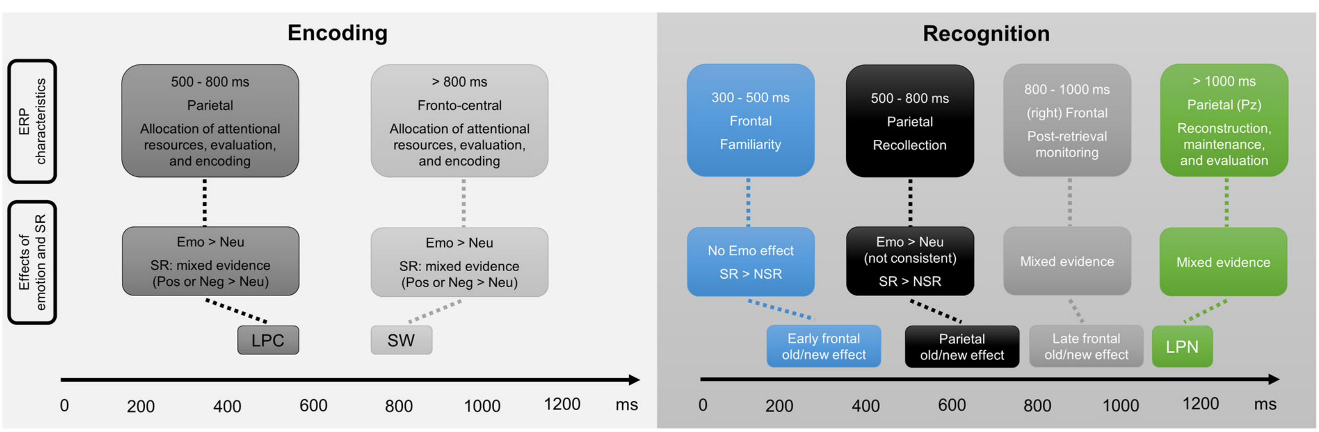

Fig. 1 Event-related potentials of interest in the current study. Their characteristics (latency in ms; topography; associated cognitive operations) are briefly described, as well as the main findings reported in the literature concerning the effects of emotion and self-reference on the amplitude of each component. Emo = Emotion; ERP = Event-related potential; LPC = Late Positive Complex; LPN = Late Posterior Negativity; Neg = Negative; Neu = Neutral; NSR = Non-self-reference; Pos = Positive; SR = Self-reference; SW = Slow Wave

Angrilli, & Mini, 1997; Wirkner, Ventura-Bort, Schulz, Hamm, & Weymar, 2018). Nonetheless, some studies reported valence-specific effects, namely enhanced LPC amplitude for positive relative to both negative and neutral stimuli (Herbert, Kissler, Junghöfer, Peyk, & Rockstroh, 2006) or for negative relative to both positive and neutral stimuli (Gallant, Pun, & Yang, 2018). Notwithstanding, none of these studies probed the interactive effects between valence and self-reference under explicit encoding instructions. Indirect evidence from studies comparing self-referential processing in depressive and healthy participants revealed that, in the later group, increased LPC amplitude for positive self-related words compared with negative words was observed (Shestyuk & Deldin, 2010). Given the paucity of research on this topic, the current study examined how valence and self-reference modulate the LPC and SW under intentional encoding conditions.

Effects of emotion and self-reference on recognition memory: ERP evidence

In a typical recognition memory paradigm, the encoding phase is followed by a memory test, in which the previously presented stimuli are repeated and intermixed with new items, and participants are invited to make old/new judgments. Given that source memory involves testing memory for features such as perceptual (e.g., stimulus color or background), contextual (e.g., temporal: list 1 vs. list 2; spatial: location on the screen), semantic (e.g., organization of items into specific thematic units), affective (e.g., emotional reactions during the event), and cognitive operations (e.g., encoding tasks; verbal thoughts; mental images; Johnson et al., 1993), the source memory test usually requires the discrimination between different details manipulated during encoding.

Differences in brain electrical activity derived from correctly recognized old items and correctly rejected new items were found to be associated with old/new effects characterized by distinct spatiotemporal distributions, polarities, and putative functional roles (see Voss & Paller, 2017; see Fig. 1 for a schematic illustration). The first well-described component is the early frontal old/new effect (also known as frontal N400-like component: FN400; 300-500 ms post-stimulus onset; maximum voltage over frontal electrodes), which is characterized by an enhanced negativity for new compared to old items (Curran, 2000). This effect has been proposed as an index of familiarity, i.e., the impression that certain information was previously seen even though no further details from the study episode are available (Curran, 2000; Curran & Dien, 2003; Czernochowski, Mecklinger, Johansson, & Brinkmann, 2005; but see Voss & Federmeier, 2011 for a different proposal). In a subsequent time window, the parietal old/new effect is typically observed (Fig. 1), i.e., old compared to new items elicit an enhanced positivity (Curran, 2000; Leynes & Phillips, 2008). This effect has been related to operations involving the recollection of information from the study episode as its amplitude is sensitive not only to accurate source responses (vs. incorrect source responses and responses attributed to familiarity; Addante, Ranganath, & Yonelinas, 2012; Cansino & Trejo-Morales, 2008; Leynes, Askin, & Landau, 2017; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999; Wilding & Rugg, 1997), but also to the number of recoverable details (Leynes & Crawford, 2018; Leynes, Crawford, & Bink, 2005; Wilding, 2000; Woroch & Gonsalves, 2010).

The affective properties of the target stimuli were found to modulate the parietal old/new effect more strongly than the early frontal old/new effect: whereas the parietal old/new effect seems to be characterized by an increased positivity in response to emotional (negative and/or positive) compared with neutral items (Langeslag & Van Strien, 2008; McNeely, Dywan, & Segalowitz, 2004; Weymar, Löw, & Hamm, 2011; Weymar, Löw, Melzig, & Hamm, 2009; Wirkner et al., 2018), the early frontal old/new effect seems to be relatively unaffected by emotion manipulations (Johansson, Mecklinger, & Treese, 2004; Koenig & Mecklinger, 2008; Lavoie & O'Connor, 2013; Maratos, Allan, & Rugg, 2000; Minor & Herzmann, 2019; Newsome, Dulas & Duarte, 2012; Weymar et al., 2009; but see Inaba, Nomura, & Ohira, 2005; Mao, You, Li, & Guo, 2015; Meng et al., 2017; Schaefer, Pottage, & Rickart, 2011; Xu, Zhang, Li, & Guo, 2015). Nonetheless, emotion effects on the parietal old/new effect are not fully consistent across studies (lack of effect: Lavoie & O'Connor, 2013; Maratos et al., 2000; Windmann & Kutas, 2001; enhanced amplitude for negative items: Inaba et al., 2005; Johansson et al., 2004; Newsome, Dulas, & Duarte, 2012; Weymar et al., 2011). Also, when self-referential encoding conditions were employed during encoding, amplitude enhancements of both early frontal and parietal old/new effects have been documented in the context of source memory attributions (Mao et al., 2017; Dulas et al., 2011).

During recognition, two other later electrophysiological responses might emerge, especially when source discriminations are required: the late right frontal old/new effect and the late posterior negativity (LPN) (Cycowicz & Friedman, 2003; Donaldson & Rugg, 1998; Johansson & Mecklinger, 2003; Mecklinger, Rosburg, & Johansson, 2016). The first is a enhanced positivity in response to old compared to new items (Friedman & Johnson, 2000; Fig. 1). Although its functional role remains to be specified, this effect has been related to post-retrieval monitoring processes that allow the evaluation and search of information retrieved from memory (Donaldson & Rugg, 1998; Leynes, 2002; Leynes, Grey, & Crawford, 2006; Trott et al., 1999; Van Petten, Senkfor, & Newberg, 2000; Wilding & Rugg, 1996). The second, the LPN, is a negative-going deflection whose amplitude is enhanced for studied compared with unstudied items (Friedman, Cycowicz, & Bersick, 2005; Mecklinger, Johansson, Parra, & Hanslmayr, 2007; Mecklinger et al., 2016; Fig. 1). It has been associated with the reconstruction, maintenance, and evaluation of item and contextual features of a prior learning episode, which may be required during difficult source memory decisions or when task-relevant details are not readily available (Johansson & Mecklinger, 2003; Leynes et al., 2006; Mecklinger et al., 2016; Wilding & Rugg, 1997).

With respect to the effects of emotion on these later components, the evidence is mixed and scarce. Whereas some studies reported no right frontal old/new effect for emotional stimuli (Maratos et al., 2000; Newsome et al., 2012), other studies documented an amplitude enhancement in response to emotional stimuli (Langeslag & Van Strien, 2008; McNeely et al., 2004; Xu et al., 2015). In the case of the LPN, some studies failed to show modulatory effects of emotion on its amplitude (Koenig & Mecklinger, 2008), whereas others indicated that this effect might emerge only under specific emotional conditions (see Newsome et al., 2012 for an example regarding negative stimuli). Thus, the studies conducted so far suggest that the presence or absence of these old/new effects may depend on the type of source details tested during recognition and on how emotion affects the memory of such source-specifying details. Taken together, the modulatory effects of emotion and self-reference on the old/new effects in the context of source memory remain to be specified.

The current study and hypotheses

This study aimed to probe the interactive effects of self-reference and valence on both recognition of internal source memory details (the encoding task used in the study phase) and processing of target words during encoding. Specifically, in the self-referential condition, participants were instructed to evaluate whether the words described them, whereas in the non-self-referential condition, participants evaluated whether the words were perceived as being commonly used by other people in daily living (i.e., common condition).

At the behavioral level, we anticipated that source memory performance would be improved for both positive and neutral compared to negative words in the self-referential condition, following prior studies (Durbin et al., 2017; Leshikar et al., 2015; Pereira et al., 2019; Y. Zhang et al., 2018). In the case of the non-self-referential condition, source memory performance was expected to be enhanced for neutral compared to negative and positive words (e.g., Cook et al., 2007; Ferré et al., 2019, Experiment 2 and 3; Pereira et al., 2019, Experiment 2). Nonetheless, a lack of effect might also be anticipated (e.g., Ferré et al., 2019, Experiment 1; Kensinger & Schacter, 2006; Sharot & Yonelinas, 2008; see also Pereira et al., 2019 for a summary table).

During encoding, we predicted that words studied in the self-referential condition would elicit larger LPC and SW mean amplitudes regardless of valence (Hudson et al., 2020; Nowicka et al., 2018; Zhao et al., 2016), as self-referential encoding is expected to lead to greater elaboration due to its motivational relevance (Symons & Johnson, 1997). Based on the assumption that there is a processing bias toward positive self-referent information, larger LPC and SW amplitudes were anticipated in response to positive relative to negative and neutral words when self-referentially encoded (Herbert et al., 2006; Herbert et al., 2010; Shestyuk & Deldin, 2010). In the case of the non-self-referential condition, the amplitudes of these components were expected to be larger for emotional relative to neutral words (Barnacle et al., 2018; Dolcos & Cabeza, 2002; Palomba et al., 1997; Wirkner et al., 2018).

With respect to the old/new electrophysiological effects, we hypothesized the following:

a) Typical early frontal old/new effects were expected for old items associated with accurate source memory responses (vs. correct rejections of new items), irrespective of self-reference and valence, considering prior evidence (Johansson et al., 2004; Koenig & Mecklinger, 2008; Lavoie & O'Connor, 2013; Maratos et al., 2000; Minor & Herzmann, 2019; Newsome et al., 2012; Weymar et al., 2009).

b) Typical parietal old/new effects were expected to be reduced in response to negative self-referential source hits compared to both positive and neutral self-referential source hits. This was specifically based on the assumption that, in regular conditions, the amount of retrieved details might be lower for negative words as participants tend to disregard negative self-referent information (D'Argembeau & Van der Linden, 2008; Derry & Kuiper, 1981; Lewis, Critchley, Rotshtein, & Dolan, 2007; Moran, Macrae, Heatherton, Wyland, & Kelley, 2006; Pauly, Finkelmeyer, Schneider, & Habel, 2013). No valence-related modulations were expected for the amplitude of source hits in the non-self-referential condition as the amount of source features retrieved, even if qualitatively distinct, may not vary as much across valence categories, especially considering the use of intentional encoding instructions.

c) Based on the functional roles of the late right frontal old/new effect and the LPN (e.g., post-retrieval; monitoring; evaluative processes; Johansson & Mecklinger, 2003; Leynes et al., 2006; Mecklinger et al., 2016; Rugg, Allan, & Birch, 2000), differences between experimental conditions were expected to emerge for those more likely associated with difficult source memory decisions, such as words studied in the non-self-referential condition or negative words studied in the self-referential condition. Additionally, such differences were expected to be more pronounced in the late right frontal old/new effect- than in the LPN-time window, as it is likely that the details of the study episode recovered during the test phase are diagnostic and relevant to source discrimination, resulting in a reduced mobilization of reconstructive processes typically related to the LPN (Johansson & Mecklinger, 2003; Mecklinger et al., 2016).

Methods

Participants

Twenty-two right-handed European Portuguese speakers (19 females) aged between 19 and 37 (M = 25.81, SD = 6.35) years, with an average of 16.14 years of formal education participated in the study. All participants had normal or corrected-to-normal vision and hearing. Also, they reported no current psychiatric or neurological conditions, no history of drug abuse, and no psychotropic medication. To rule out the presence of mood disorders, participants completed the Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996; Coelho, Martins, & Barros, 2002; M = 3.68, SD = 5.58) and the State-Trait Anxiety Inventory - Form Y (STAI-Y; Silva & Spielberger, 2007; state subscale: M = 27.64, SD = 5.79; trait subscale: M = 28.95, SD = 9.07). Additionally, the participants included in this study had at least 12 artifact-free trials in each condition of interest (see Supplementary Table S2). Participants provided written informed consent prior to their enrollment in the study. The study was approved by a local Ethics committee (University of Minho, Braga, Portugal; SECVS 105/2016).

Stimuli

Stimuli were 360 European Portuguese (EP) adjectives (120 neutral, 120 positive, 120 negative), selected from the EP version of the Affective Norms for English Words (ANEW; Bradley & Lang, 1999; Soares, Comesaña, Pinheiro, Simões, & Frade, 2012), and from a pilot study (n = 125 EP native speakers) that aimed to expand the ANEW adjectives set (see Supplementary Table S1). Details regarding the pilot study were described elsewhere (Pereira, Sampaio, & Pinheiro, 2020). Adjectives differed in valence (negative < neutral < positive; all $p < 0.001$) but not in arousal ($p = 0.12$). The decision to keep arousal at medium levels across valence categories followed prior literature showing that ERPs obtained during encoding and recognition may be differentially modulated by valence and arousal (Delaney-Busch, Wilkie, & Kuperberg, 2016; Meng et al., 2017; Xu et al., 2015). For instance, an inverted-U relationship was found between the amplitude of the parietal old/new effect and both stimulus valence (Schaefer, Fletcher, Pottage, Alexander, & Brown, 2009) and arousal (Schaefer et al., 2011). Thus, we ensured that only stimulus valence differed across conditions following previous studies (Gallant et al., 2019; Gallant & Yang, 2014). The words in the three valence conditions also were matched for frequency ($p = 0.20$), number of letters ($p = 0.83$), and number of syllables ($p = 0.61$; Supplementary Table S1). Stimuli were randomly divided into 10 lists to create five different study-test cycles. These lists did not differ in mean valence and arousal ratings, frequency, number of letters, and number of syllables ($p > 0.05$).

Procedure

The experiment took place in a sound-attenuated, dimly lit, electrically shielded room. Participants were told that they were about to perform a memory task with words, but first they had to perform two-choice judgments for each stimulus. For some words, they were instructed to evaluate whether the word described or related to them in any way by responding "yes" or "no" (self-referential condition). For other words, they had to evaluate whether the words were commonly used by other people in their everyday lives, also by responding "yes" or "no" (non-self-referential or common condition). Participants were informed that the self-reference and the common judgment tasks would appear in a mixed fashion. It was emphasized that the words and how they were evaluated should be memorized for a later memory test (intentional encoding). Moreover, after each word, they were required to rate in a 6-point scale how likely they were to remember in a later memory test that the study word was evaluated according to the self-reference or the common judgment task (judgments of source - JOSs: 1 = "Sure I will not remember the judgment"; 6 = "Sure I will remember the judgment"). In the memory test, the studied words were intermixed with new words and participants had to select one of four options: "self-description"; "common"; "evaluated, but do not know if self-description/common"; or "new." In both phases, participants were instructed to respond as fast as possible. During the test phase, response accuracy was highlighted. Participants performed two short study-test cycles to become familiarized with the task requirements. Each study-test cycle included six trials during the study phase (3 self-referential, 3 common) and nine trials during the test phase (6 old and 3 new words). None of these words were repeated in the main experiment.

The experimental task was composed of five study-test cycles, which were counterbalanced and randomized across participants. Additionally, the word lists assigned to each study and test phase were counterbalanced across participants so that the words used during the study phase for one participant could be new words during the test phase for another participant. Each study phase included 48 trials: half required a self-referential judgment, and the other half required a common judgment. The assignment of each word to one of those tasks also was counterbalanced across participants so that the same word could be evaluated following the common judgment for one participant and the self-referential judgment for another participant. For each encoding task, eight words were negative, eight were neutral, and eight were positive. The words from each experimental condition—self-negative, self-neutral, self-positive, common-negative, common-neutral, common-positive—were intermixed and randomly presented, one by one, in the center of the computer screen (white Arial; font size 88; black background). The structure of an experimental trial and associated details are presented in Fig. 2.

Each test phase was composed of 72 trials: 48 studied words randomly mixed with 24 new words (8 from each valence category). Ideally, the number of old and new items should be balanced due to effects of differences in the proportion of old/new items on behavioral and brain measures (Vilberg & Rugg, 2009). However, due to several constraints (e.g., limitations of the available EP words; length of the recognition phase and overall experiment) and because we were particularly interested in the modulatory role of valence in successful self-referential source memory attributions, we decided to increase the signal-to-noise ratio in favor of old items, following previous studies (Mao et al., 2015; Minor & Herzmann, 2019; Newsome et al., 2012). The structure of an experimental trial in the test phase is illustrated in Fig. 2. Of note, the "do not know" option was provided to mitigate possible response contaminations due to guesses and biases (Addante et al., 2012; Dulas & Duarte, 2013; Minor & Herzmann, 2019; Newsome et al., 2012; Yin et al., 2019), which is ideal to isolate the electrophysiological activity underlying accurate source memory responses (Leynes et al., 2017; Wilding, 2000; Wilding & Rugg, 1997).

At the end of the experiment, participants were debriefed about their performance and the implementation of specific encoding strategies. Overall, the experimental session lasted for approximately 2 hours. Presentation® software (Version 18.3, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) was used to control stimulus presentation.

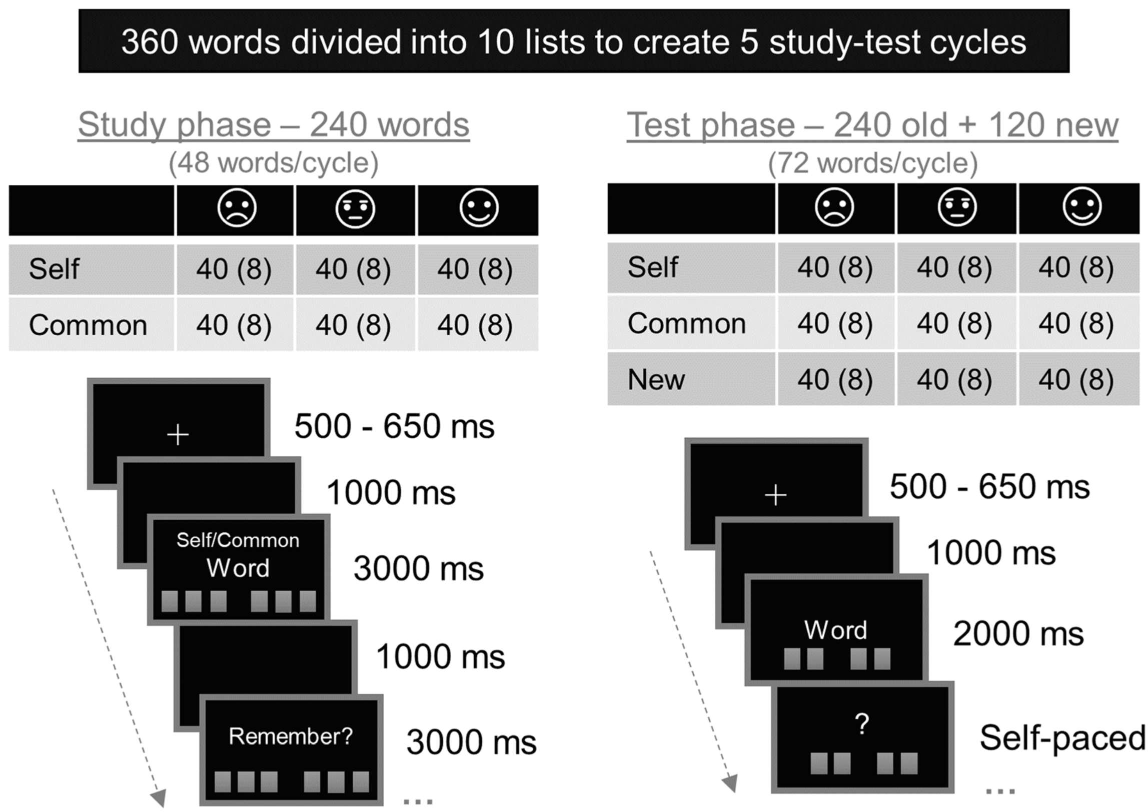

Fig. 2 Experimental trial from the study and test phases. The number and type of stimuli regarding each encoding task (self vs. common) x valence (negative vs. neutral vs. positive) condition are summarized. Note: During the study phase, each trial began with a fixation cross (500–650 ms) followed by a blank screen (1,000 ms). Then, the word appeared together with an instruction at the top of the screen, which informed about the ongoing judgment task (3,000 ms). At the bottom of the same screen, six small rectangles were shown to represent the position of the fingers required for the study phase. Particularly, participants were asked to place the left ring, middle, and index fingers on the keyboard keys "Z"(1), "X"(2), "C"(3), and the right ring, middle, and index fingers on the keyboard keys "B"(4), "N"(5), "M"(6). These keys were covered with stickers numbered from 1 to 6 to facilitate finger-key mapping and the response during the JOSs ratings, and to minimize cognitive demands. The "yes"/"no" responses were restricted to the index fingers and counterbalanced across participants. Another blank screen was introduced (1,500 ms) followed by the JOSs rating which was prompted by the word "remember?" (3,000 ms). In the test phase, each trial was composed of a fixation cross (500-650 ms), a black blank screen (1,000 ms), the old/new word (2,000 ms), and a response screen with the prompt "?" containing four response options at the bottom of the screen ("self-description"; "common"; "evaluated, but do not know if self-description/common"; "new"). At this stage, participants used four keys/fingers: "X" (2; left middle), "C" (3; left index), "B" (4; right index), "N" (5; right middle). The finger-response mapping was counterbalanced across participants

ERP data acquisition and analysis

EEG data were recorded continuously at a digitization rate of 512 Hz using a 64-channel BioSemi Active Two system (Biosemi, Amsterdam, Netherlands; http://www.biosemi.com/products.htm) and stored on hard disk for later analysis. An additional five external electrodes were used: two were placed on the left and right mastoids (reference); two electrodes were placed on the outer canthi of the eyes to register the horizontal electrooculogram (HEOG); and one was placed below the left to record the vertical electrooculogram (VEOG). Electrode offset was kept below 30 mV. The EEG was recorded during the study and test phases of all five experimental blocks.

The offline EEG processing was conducted using two free-open source MATLAB packages suited for EEG/ERP data analysis: the EEGLAB toolbox version 13.5.4b (Delorme & Makeig, 2004; http://sccn.ucsd.edu/eeglab/) and the ERPLAB Toolbox version 7.0.0 (Lopez-Calderon & Luck, 2014; http://erpinfo.org/erplab/). Data were first band-pass filtered at 0.1-40 Hz using an IIR Butterworth filter (12 dB/octave roll-of slope) and referenced to the averaged of the right and left mastoids. An independent component analysis (ICA; runica algorithm) was run to detect and correct eye-related artifacts. To guide the rejection of components, three EEGLAB data processing extensions were used: SASICA (Chaumon, Bishop, & Busch, 2015; https://github.com/dnacombo/SASICA), ADJUST (Mognon, Jovicich, Bruzzone, & Buiatti, 2011; https://www.nitrc.org/docman/view.php/739/2101/TutorialADJUST1_1.pdf), and ICLabel (Pion-Tonachini, Makeig, & Kreutz-Delgado, 2017; https://labeling.ucsd.edu/tutorial/overview). On average, 2.82 components (SD = 0.66; range 2-4) were removed per participant. Noisy channels were interpolated using the spherical interpolation method (M = 3.95; SD = 1.29; range 0-5). Event-based epochs were created including a baseline of 200 ms before stimulus onset and 1,200 ms (word encoding) or 2,000 ms (recognition phase) post-stimulus onset. All epochs were baseline corrected considering the 200 ms period before stimulus onset. Epochs with ±100 μV amplitude were discarded from the analysis before averaging.

Following prior studies (Barnacle et al., 2018; Dolcos & Cabeza, 2002; Fischler & Bradley, 2006; Gallant et al., 2018; Herbert et al., 2006; Koenig & Mecklinger, 2008), mean amplitude (μV) during word encoding was extracted in the following time windows: 500–700 (LPC); 700–950; 950–1,200 ms (SWs). We were only able to run the comparison between experimental conditions for general encoding due to an insufficient number of trials associated with incorrect responses or hits in some participants (Supplementary Table S2).

Regarding the recognition stage, the mean amplitude of well-described ERP components was extracted for both source hits and correct rejections of new items (Cansino, Hernández-Ramos, & Trejo-Morales, 2012; Dulas et al., 2011; Koenig & Mecklinger, 2008; Mao et al., 2015; Newsome et al., 2012; Wilding, 1999): early frontal old/new effect (300–500 ms), parietal old/new effect (500–800 ms), right frontal old/new effect (800–1,200; 1,200–1,800 ms), LPN (800–1,200; 1,200–1,800 ms). Again, the ERP responses related to incorrect responses were not examined further due to the low number of trials. Mean amplitude was extracted from distinct regions of interest (ROIs) covering anterior to posterior and left to right locations, following previous studies (Dulas et al., 2011; Langeslag & Van Strien, 2008): left frontal (F3/F5), middle frontal (F1/Fz/F2), right frontal (F4/F6), left fronto-central (FC3/FC5), middle fronto-central (FC1/FCz/FC2), right fronto-central (FC4/FC6), left central (C3/C5), middle central (C1/Cz/C2), right central (C4/C6), left centro-parietal (CP3/CP5), middle centro-parietal (CP1/CPz/CP2), right centro-parietal (CP4/CP6), left parietal (P3/P5), middle parietal (P1/Pz/P2), right parietal (P4/P6).

Statistical analysis

Behavioral data

Following Snodgrass and Corwin (1988), the recognition measure Pr = ([p(hits) – p(false alarms)]) was used as an index of item recognition accuracy, wherein p(hits) was obtained by collapsing the "do not know" responses with both correct and incorrect source responses. To index response bias, the Br measure was used - Br = ([p(false alarms)/(1 – Pr)]) - wherein values above 0.50 indicate a more liberal response criterion, i.e., in cases of uncertainty the decision is biased to choose "old"; whilst values below 0.50 indicate a more conservative response criterion, that is, the decision is biased to select "new" in the face of uncertainty. Regarding source memory recognition, we computed the difference between source hits and incorrect source responses: [p(correct source) – p (incorrect source)] (Dulas & Duarte, 2012, 2014; Hou et al., 2019; Leshikar et al., 2015; Newsome et al., 2012). Data were submitted to a 2 (self-reference: self vs. common) x 3 (valence: negative vs. neutral vs. positive) repeated-measures analysis of variance (RMANOVA).

ERP data

Mean amplitudes during encoding were subjected to a 2 (self-reference: self *vs.* common) x 3 (valence: negative vs. neutral vs. positive) x 5 (ROI: frontal vs. fronto-central vs. central vs. centro-parietal vs. parietal) x 3 (laterality: left vs. middle vs. right) RMANOVA. In the recognition phase, a 3 (self vs. common vs. new) x 3 (valence) x 5 (ROI) x 3 (laterality) RMANOVA was computed. Statistically significant effects were described and followed up if they included the experimental conditions (self-reference and/or valence). For all the RMANOVAS, Bonferroni-corrected post-hoc tests were run to qualify main and interaction effects, and the Greenhouse-Geisser correction was considered when violations of sphericity were present. Partial eta squared (η^2_p) was used as a measure of effect size.[1]

Results

Table 1 presents a summary of the main behavioral results, and Table 2 presents a summary of the RMANOVA ERP results.

Behavioral results (recognition accuracy)

Item recognition The 2 × 3 RMANOVA revealed a main effect of self-reference, $F(1, 21) = 44.68$, $p < 0.001$, η^2_p = 0.68, and valence, $F(2, 42) = 10.49$, $p < 0.001$, $\eta^2_p = 0.33$, as well as an interaction effect between factors, $F(2, 42) = 7.98$, $p = 0.001$, $\eta^2_p = 0.28$. Specifically, the interaction revealed that, irrespective of valence, there was an item memory benefit for words studied self-referentially compared with words studied in the common condition (negative: $p = 0.019$, 95% CI_{diff} [0.01, 0.09]; neutral: $p < 0.001$, 95% CI_{diff} [0.10, 0.20]; positive: $p < 0.001$, 95% CI_{diff} [0.08, 0.16]). Whereas no differences in item recognition were observed between negative ($M = 0.59$, $SE = 0.03$), neutral ($M = 0.63$, $SE = 0.03$), and positive words ($M = 0.59$, $SE = 0.02$) in the common condition, there were statistically significant differences between all valence categories for words in the self-referential condition: neutral ($M = 0.78$, $SE = 0.03$; $p < 0.001$, 95% CI_{diff} [0.09, 0.18]) and positive words ($M = 0.71$, $SE = 0.03$; $p = 0.019$, 95% CI_{diff} [0.01, 0.12]) were better recognized than negative words ($M = 0.64$, $SE = 0.03$), and neutral words were also better recognized than positive words ($p = .014$, 95% CI_{diff} [0.01, 0.13]; Fig. 3a).

In the case of the response bias measures, the 2 x 3 RMANOVA showed a main effect of self-reference, $F(1, 21) = 27.05$, $p < 0.001$, $\eta^2_p = 0.56$, and valence, $F(2, 42) = 4.98$, $p = 0.019$, $\eta^2_p = 0.19$, $\varepsilon = 0.79$, as well as an interaction effect, $F(2, 42) = 5.64$, $p = 0.007$, $\eta^2_p = 0.21$. The Br value was higher in the self-referential condition than in the common condition for neutral ($p < 0.001$, 95% CI_{diff} [0.11, 0.30]) and positive words ($p < 0.001$, 95% CI_{diff} [0.10, 0.28]), but not for negative words ($p = 0.064$, 95% CI_{diff} [−0.004, 0.13]). This indicates that for both neutral and positive words, participants used a more conservative response criterion in the common condition than in the self-referential condition. In fact, on average, the Br values were above 0.50 in the self-referential condition, which is indicative of a more liberal response criterion, whilst the same was not observed in the common condition (see Table 1; Snodgrass & Corwin, 1988). Regarding the self-referential condition, the response bias was higher for positive ($M = 0.64$, $SE = 0.06$) than for negative words ($M = 0.54$, $SE = 0.04$; $p = 0.046$, 95% CI_{diff} [0.001, 0.19]). For the common condition, the response criterion was more conservative for neutral words ($M = 0.34$, $SE = 0.05$) relative to both negative ($M = 0.48$, $SE = 0.04$; $p = 0.004$, 95% CI_{diff} [0.04, 0.25]) and positive words ($M = 0.45$, $SE = 0.05$; $p = 0.031$, 95% CI_{diff} [0.01, 0.21]).

Source recognition The main effects and the interaction effect were statistically significant [self-reference: $F(1, 21) = 13.34$, $p = 0.001$, $\eta^2_p = 0.39$; valence: $F(2, 42) = 10.71$, $p < 0.001$, η^2_p = 0.34; self-reference x valence: $F(2, 42) = 10.62$, $p < 0.001$, η^2_p = 0.34]. The planned comparisons revealed that for neutral ($p = 0.005$, 95% CI_{diff} [0.04, 0.19]) and positive words ($p < 0.001$, 95% CI_{diff} [0.15, 0.36]), source memory recognition was improved in the self-referential as compared to the common condition (Table 1). Furthermore, whereas self-referential processing benefited source memory for neutral ($M = 0.64$, $SE = 0.04$; $p < 0.001$, 95% CI_{diff} [0.07, 0.25]) and positive words ($M = 0.67$, $SE = 0.04$; $p < 0.001$, 95% CI_{diff} [0.09, 0.28]) compared with negative words ($M = 0.48$, $SE = 0.04$), the source memory for words studied in the common condition was improved when words had a neutral ($M = 0.53$, $SE = 0.04$) compared with a positive valence ($M = 0.41$, $SE = 0.04$; $p = 0.007$, 95% CI_{diff} [0.03, 0.21]; Fig. 3b), with no differences in the case of negative words ($M = 0.46$, $SE = 0.04$).

ERP results

Word encoding *500–700 ms.* No self-reference- and/or valence-related modulations were found in this time window (Fig. 4).

700-950 and 950-1,200 ms. In the 700–950 ms latency window, two interaction effects including the self-reference factor were observed: self-reference x ROI; self-reference x ROI x laterality. Specifically, the mean amplitude obtained for words studied self-referentially was more positive-going than the mean amplitude for words studied in the common condition irrespective of valence (all $p < 0.05$; Fig. 4) and over most ROIs (exceptions were detected in the left and middle parietal locations). In the 950–1,200 ms latency window, similar interaction effects emerged (Table 2). The mean amplitude for words studied in the self-referential condition was again more

[1] In the *Supplementary Material* section, we present additional analyses, namely: one sample *t*-tests computed for item and source recognition measures; differences between experimental conditions with respect to the proportion of misses, false alarms, correct rejections, "do not know" responses, incorrect source responses, "yes" responses, and mean response times collected during the word encoding and JOSs ratings; mean JOSs ratings, electrophysiological responses, and Goodman-Kruskal gamma correlations computed for each experimental condition.

Table 1 Descriptive statistics of the behavioral performance as a function of encoding task and valence

Measure	Encoding task / Valence					
	Self – *M* (*SD*)			Common – *M* (*SD*)		
	Negative	Neutral	Positive	Negative	Neutral	Positive
Source hit	0.60 (0.15)	0.74 (0.13)	0.75 (0.14)	0.55 (0.14)	0.59 (0.16)	0.53 (0.12)
Incorrect source	0.12 (0.08)	0.10 (0.07)	0.08 (0.07)	0.10 (0.07)	0.06 (0.04)	0.12 (0.09)
DNK source	0.11 (0.07)	0.06 (0.04)	0.06 (0.05)	0.13 (0.09)	0.10 (0.06)	0.12 (0.07)
Source miss	0.17 (0.10)	0.10 (0.09)	0.11 (0.09)	0.22 (0.11)	0.25 (0.14)	0.23 (0.10)
Item Pr measure	0.64 (0.12)	0.78 (0.12)	0.71 (0.13)	0.59 (0.13)	0.63 (0.15)	0.59 (0.11)
Item Br measure	0.54 (0.21)	0.54 (0.30)	0.64 (0.28)	0.48 (0.20)	0.34 (0.21)	0.45 (0.24)
SM measure	0.48 (0.21)	0.64 (0.18)	0.67 (0.20)	0.46 (0.18)	0.53 (0.18)	0.41 (0.19)
RT - Word encoding	1772 (341)	1932 (277)	1770 (236)	1804 (259)	1793 (275)	1767 (309)
	New negative		New neutral		New positive	
CR	0.81 (0.10)		0.88 (0.08)		0.81 (0.12)	
FA	0.19 (0.10)		0.12 (0.08)		0.19 (0.12)	

CR Correct Rejections, *DNK* Do Not Know, *FA* False Alarms, *RT* Response Time, *SM* Source Memory.

positive than for words studied in the common condition for most ROIs (all $p < 0.05$; exceptions were detected over right centro-parietal and all parietal electrode sites; Fig. 4).

Test phase *300–500 ms.* The analysis considering self-source hits, common-source hits, and correct rejections yielded a main effect of self-reference. The mean amplitude for self-source hits ($M = 1.51$, $SE = 0.57$) was larger than for correct rejections ($M = 0.51$, $SE = 0.75$; $p = 0.014$, 95% CI_{diff} [0.17, 1.82]), whereas the mean amplitude associated with common-source hits did not differ from the mean amplitude of both self-source hits and correct rejections ($p > 0.05$). Thereby, this early time window showed a general differentiation between old self-referential and new words, characterized by a more positive-going amplitude for old compared with new items (Fig. 5).

500–800 ms. A self-reference x valence x laterality interaction was observed (Table 2). The follow-up of the three-way interaction indicated that, in the case of negative words, the mean amplitude for self-source hits (left: $M = 3.93$, $SE = 0.68$; middle: $M = 4.08$, $SE = 0.99$; right: $M = 3.90$, $SE = 0.89$) was larger than the mean amplitude for correct rejections irrespective of laterality (left: $M = 2.33$, $SE = 0.68$; middle: $M = 2.30$, $SE = 0.98$; right: $M = 2.39$, $SE = 0.79$; all $p < 0.05$). Considering neutral words, the mean amplitude for both self- and common-source hits (self left: $M = 4.40$, $SE = 0.69$;

Table 2 Summary of the ERP RMANOVA results that include the experimental conditions under study (Self-Reference and Valence)

Effects	Word encoding									Test phase											
	500 - 700			700 - 950			950 - 1,200			300 - 500			500 - 800			800 - 1,200			1,200 - 1,800		
	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
SR	3.13	0.091	0.13	9.87	0.005	0.32	16.21	0.001	0.44	5.22	0.009	0.20	7.85	0.001	0.27	1.63	0.208	0.07	1.51	0.233	0.07
Valence	0.25	0.778	0.01	0.98	0.385	0.04	0.27	0.763	0.01	2.30	0.113	0.10	0.49	0.619	0.02	0.65	0.529	0.03	1.15	0.232	0.07
SR x valence	1.71	0.193	0.08	2.21	0.122	0.10	1.23	0.302	0.10	1.81	0.134	0.08	1.76	0.145	0.08	1.76	0.145	0.08	0.92	0.458	0.04
SR x ROI	0.42	0.644	0.02	5.45	0.005	0.21	21.73	<0.001	0.51	0.28	0.781	0.01	1.36	0.262	0.06	10.82	<0.001	0.34	10.81	<0.001	0.34
Valence x ROI	1.03	0.388	0.05	0.43	0.702	0.02	0.26	0.817	0.01	0.62	0.612	0.03	0.53	0.683	0.03	0.84	0.491	0.04	0.44	0.786	0.02
SR x valence x ROI	1.98	0.113	0.09	1.70	0.165	0.08	1.20	0.320	0.05	1.23	0.300	0.06	1.41	0.214	0.06	1.28	0.279	0.06	1.61	0.146	0.07
SR x laterality	0.48	0.599	0.02	2.67	0.081	0.11	2.24	0.119	0.10	1.43	0.247	0.06	1.00	0.413	0.05	2.23	0.105	0.10	6.93	0.001	0.25
Valence x laterality	0.51	0.732	0.02	0.63	0.642	0.03	1.47	0.218	0.07	2.10	0.088	0.09	0.99	0.418	0.05	1.51	0.208	0.07	2.61	0.041	0.11
SR x valence x laterality	1.03	0.399	0.05	1.12	0.353	0.05	0.91	0.463	0.04	1.82	0.076	0.08	2.89	0.005	0.12	2.00	0.049	0.09	1.36	0.216	0.06
SR x ROI x laterality	1.73	0.096	0.08	3.24	0.002	0.13	3.71	0.001	0.15	1.01	0.444	0.05	2.88	0.004	0.12	3.33	0.001	0.14	2.93	0.007	0.12
Valence x ROI x laterality	0.60	0.747	0.03	0.67	0.702	0.03	0.79	0.597	0.04	0.50	0.862	0.02	0.53	0.930	0.03	0.88	0.592	0.04	0.99	0.466	0.05
SR x valence x ROI x laterality	0.77	0.611	0.04	1.01	0.450	0.05	0.64	0.742	0.03	0.89	0.541	0.04	0.73	0.691	0.04	0.89	0.548	0.04	1.07	0.388	0.05

ERP Event-Related Potential, *RMANOVA* Repeated-Measures Analysis of Variance, *ROI* Region-of-interest, *SR* Self-reference. Statistically significant effects ($p < 0.050$) are shaded in gray.

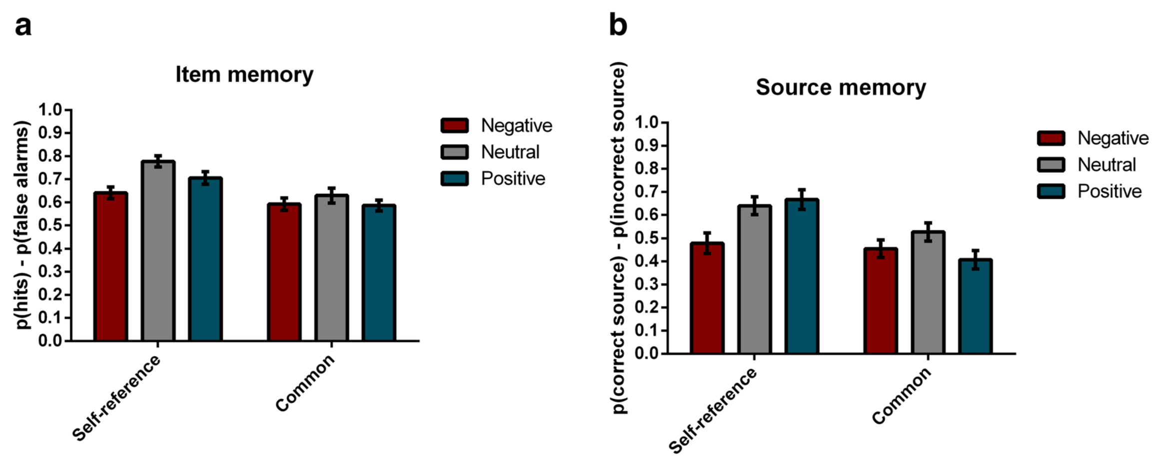

Fig. 3 Behavioral results for item memory (**a**) and source memory (**b**). The recognition scores are plotted on the y-axis as a function of encoding task (self-reference vs. common) and valence (negative vs. neutral vs. positive). The error bars represent the standard error of the mean

self middle: $M = 4.26$, $SE = 1.03$; common left: $M = 3.77$, $SE = 0.61$; common middle: $M = 3.87$, $SE = 0.96$) was larger than for correct rejections over left and middle electrode sites (new left: $M = 2.37$, $SE = 0.65$; new middle: $M = 2.36$, $SE = 0.98$; all $p < 0.05$). In the case of positive words, the mean amplitude for both self- ($M = 3.51$, $SE = 0.79$) and common-source hits ($M = 3.95$, $SE = 1.02$) was larger than for correct rejections over right electrode sites ($M = 2.09$, $SE = 0.79$), whereas only the mean amplitude for common-source hits (left: $M = 4.12$, $SE = 0.91$; middle: $M = 4.45$, $SE = 1.23$) was larger than for correct rejections over left and middle electrode sites (left: $M = 2.51$, $SE = 0.66$; middle: $M = 2.29$, $SE = 0.93$; all $p < 0.05$; Fig. 5).

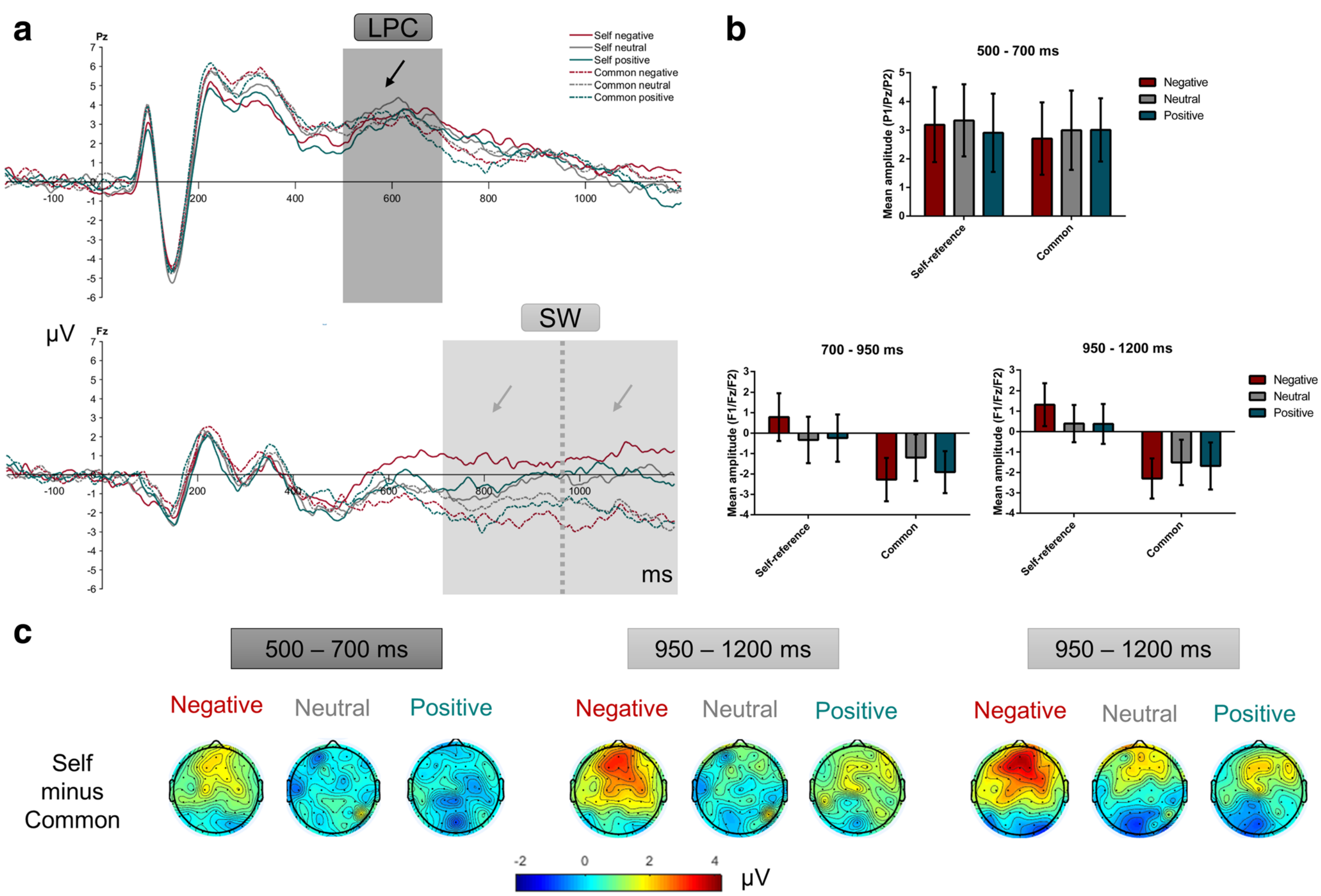

Fig. 4 Grand average waveforms obtained during word encoding (**a**) and graphic depiction of the mean amplitude in specific regions-of-interest (**b**) considering each encoding task (self vs. common) x valence (negative vs. neutral vs. positive) condition. The topographic maps (**c**) result from the subtraction between encoding tasks (self-minus-common). LPC = Late Positive Complex; SW = Slow Wave

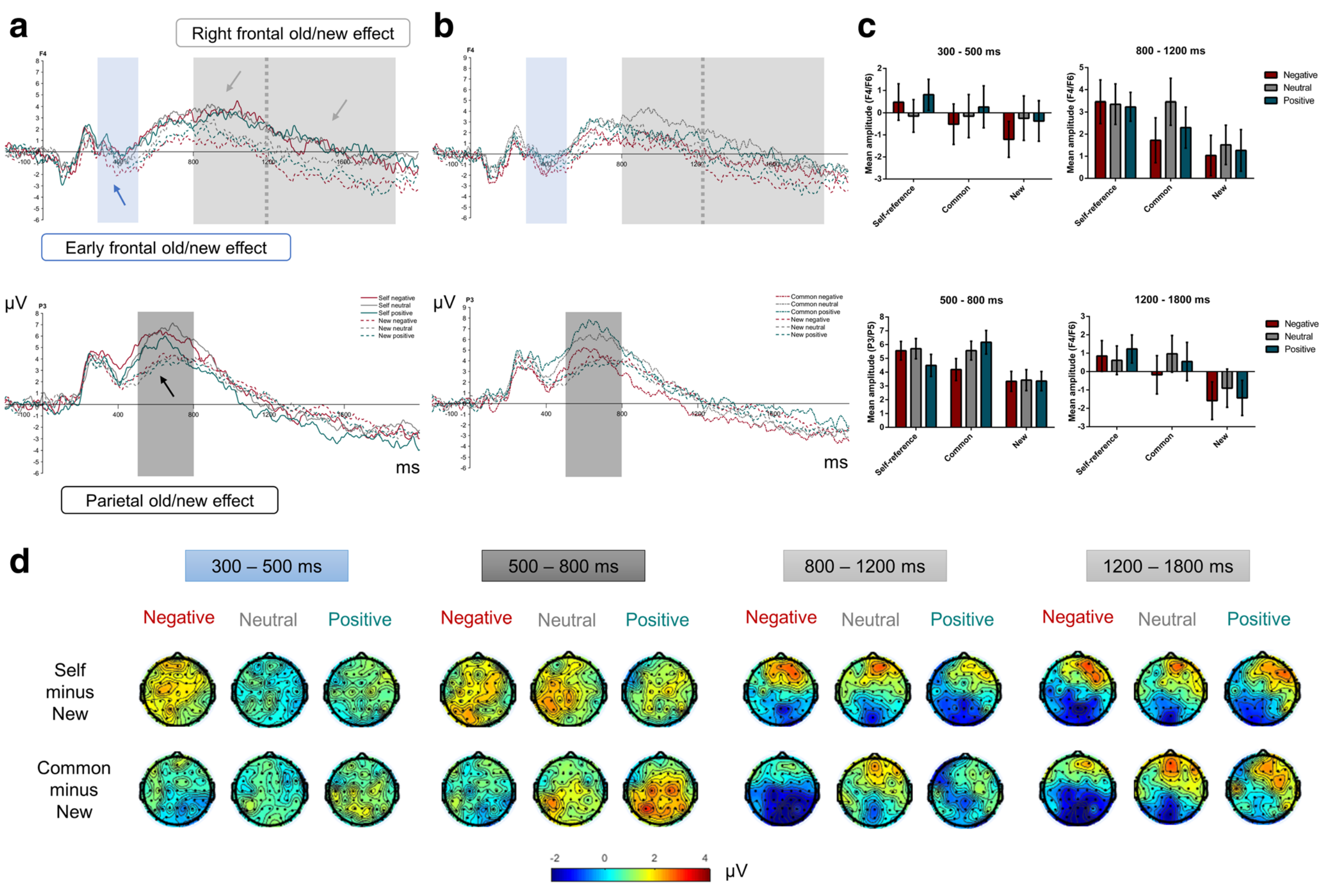

Fig. 5 Grand average waveforms and topographical maps obtained during the test phase considering the four-time windows analyzed. **a.** Self-source hits (old items) and correct rejections (new items) as a function of valence (negative vs. neutral vs. positive). **b.** Common-source hits and correct rejections. **c.** Mean amplitude in specific regions-of-interest. **d.** Topographic maps resulting from the subtraction between self-source/common-source hits and correct rejections as a function of valence

800-1,200 ms. Two interaction effects including the experimental factors of interest were observed: self-reference x valence x laterality; self-reference x ROI x laterality. Regarding the first interaction, post-hoc tests showed that, over middle electrode sites, the mean amplitude for negative correct rejections ($M = 1.38$, $SE = 0.84$) and negative self-source hits ($M = 2.05$, $SE = 1.01$) was larger than the mean amplitude for negative common-source hits ($M = 0.33$, $SE = 0.95$; all $p < 0.05$). Over right electrode sites, only the difference between negative common- ($M = 1.42$, $SE = 0.85$) and self-source hits ($M = 3.02$, $SE = 0.86$) reached statistical significance ($p = 0.031$); additionally, considering the same electrode sites and common-source hits, the mean amplitude for neutral words ($M = 3.24$, $SE = 0.81$) was larger than for negative words ($M = 1.42$, $SE = 0.85$; $p = .030$; Fig. 5). When decomposing the self-reference x ROI x laterality interaction, a typical right frontal old/new effect was found, showing a larger mean amplitude for both self- ($M = 3.35$, $SE = 0.72$) and common-source hits ($M = 2.49$, $SE = 0.87$) compared with correct rejections ($M = 1.27$, $SE = 0.86$; all $p < 0.05$; Fig. 5).

1,200-1,800 ms. In this time window, a significant self-reference x ROI x laterality interaction effect was observed (Table 2). The decomposition of the interaction revealed a typical right frontal old/new effect, which was detected over middle/right frontal and right fronto-central sites, showing that the mean amplitude for correct rejections (middle frontal: $M = -2.65$, $SE = 0.96$; right frontal: $M = -1.31$, $SE = 0.94$; right fronto-central: $M = -1.22$, $SE = 0.82$) was less positive than the mean amplitude for both self- (middle frontal: $M = -0.94$, $SE = 0.87$; right frontal: $M = 0.90$, $SE = 0.70$; right fronto-central: $M = 0.80$, $SE = 0.89$) and common-source hits (middle frontal: $M = -0.73$, $SE = 1.00$; right frontal: $M = 0.45$, $SE = 0.92$; right fronto-central: $M = 0.28$, $SE = 0.85$; all $p < 0.05$; Fig. 5).

Discussion

The main goal of the current study was twofold: 1) to probe whether and how valence and self-reference interactively modulate internal source memory, combining behavioral and electrophysiological measures during recognition (old/new effects); 2) to specify how these variables also may influence stimulus processing during intentional encoding, reflected in the LPC and SW amplitudes. With respect to behavioral performance, we observed that self-referential encoding granted an advantage not only in terms of general recognition (item memory) but, more importantly, in terms of source memory. Also, the effects of valence on source memory differed as a function of the encoding task (Fig. 3). With respect to electrophysiological measures, self-referential conditions yielded larger SW amplitudes during encoding regardless of valence. During recognition, the advantage of self-referential encoding was observed in an earlier time window typically associated with familiarity and irrespective of valence. Interactive effects began to emerge in the 500–800 ms time window as old/new effects were observed for most conditions except for negative words studied in the common condition. Also, in the 800–1,200 ms time window, the mean amplitude for negative words studied self-referentially was larger than for negative words studied in the common condition (see Fig. 6 for a summary of the main findings).

Behavioral findings

When positive and neutral words were encoded self-referentially, both item and source memory were boosted as compared to negative words, corroborating our initial hypothesis and prior findings (item memory: Leshikar et al., 2015; Pauly et al., 2013; Yang, Truong, Fuss, & Bislimovic, 2012; source memory: Durbin et al., 2017; Leshikar et al., 2015; Pereira et al., 2019; Y. Zhang et al., 2018). Such results support the notion that while neutral and especially positive information is more likely to be deemed as self-referential, the same does not occur for negative information, whose content is less likely to fit into the current self-view (D'Argembeau & Van der Linden, 2008; Derry & Kuiper, 1981; Lewis, Cairney, Manning, & Critchley, 2011; Moran et al., 2006; Pauly et al., 2013). Other behavioral indices analyzed here also support the idea that negative words were perceived as less self-relevant (see *Supplementary Materials* section); for example: self-reference effects on source memory were restricted to neutral and positive stimuli; the proportion of responses that matched the self-referential task during encoding (i.e., "yes" responses) was the lowest for negative words encoded self-referentially, whilst the proportion of misses during recognition was the highest for these words. Also, the improved source memory for neutral and positive words agrees with the notion that successful access to information stored in memory is modulated by how well this information is integrated into prior knowledge structures and how well it fits long-term goals such as maintaining a favorable and coherent self-view, while hindering the access to less relevant information (e.g., negative self-referential information; Conway & Pleydell-Pearce, 2000; Conway, Singer, & Tagini, 2004; D'Argembeau & Van der Linden, 2008).

In the case of the non-self-referential (common) condition, valence was expected to impair source memory (Pereira et al., 2019, Experiment 2). This hypothesis was partially confirmed as source memory performance for neutral words was improved compared with positive words. However, the performance for negative words was no different from neutral words, which stands in contrast with prior studies testing memory for non-self-referential and semantic encoding tasks (Cook et al., 2007; Ferré et al., 2019, Experiment 2; Mao et al., 2015; Newsome et al., 2012; Otani, Jaffa, Libkuman, Goernert, & Kato, 2012a; Otani et al., 2012b). The lack of differences in internal source memory for emotional versus neutral stimuli also was reported in prior studies and thus expected (Ferré et al., 2019, Experiment 1; Kensinger & Schacter, 2006; Pereira et al., 2019, Experiment 1; Sharot & Yonelinas, 2008). An alternative explanation is that the current study is underpowered to detect differences between negative and neutral stimuli as compared to prior studies (Pereira et al., 2019).

Emotion and self-reference effects during encoding

Contrary to our initial predictions, the LPC and SW amplitudes were not modulated by valence. Rather, differences in word processing only emerged in the SW latency windows: words studied in the self-referential condition elicited a more positive-going amplitude than words studied in the common condition. Although this effect shows that self-reference grants a processing advantage in late processing stages (Herbert et al.,

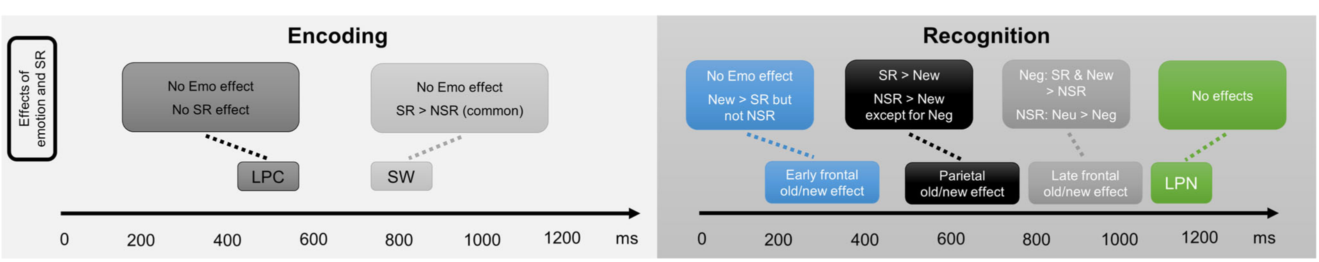

Fig. 6 Event-related potentials of interest in the current study, including the main findings. Emo = Emotion; ERP = Event-related potential; LPC = Late Positive Complex; LPN = Late Posterior Negativity; Neg = Negative; Neu = Neutral; NSR = Non-self-reference; Pos = Positive; SR = Self-reference; SW = Slow Wave

2011; Hudson et al., 2020; Mu & Han, 2010; Nowicka et al., 2018; Zhao et al., 2016), the LPC amplitude also was expected to be sensitive to stimulus valence (Gutchess & Kensinger, 2018). Even though intentional learning instructions might have contributed to this result, previous studies that used intentional conditions also reported emotion effects on both LPC and SW (Barnacle et al., 2018; Dolcos & Cabeza, 2002; Palomba et al., 1997; Wirkner et al., 2018). These contradictory findings underscore the need for assessing the effects of intentional versus incidental learning conditions on the ERP responses related to stimulus encoding as a function of valence. Nonetheless, other experimental factors, such as the use of words (vs. pictures) and of similar arousal levels across valence categories also may offer a plausible explanation.

Specifically, several authors have argued that verbal stimuli do not convey the same emotional intensity as pictures, and as a result, null findings are more likely to occur when words are used as experimental stimuli (Citron, 2012; Dolcos & Cabeza, 2002; Herbert et al., 2006; Mathews, Ridgeway, & Holmes, 2013). Instead, emotional pictures appear to be characterized by enhanced biological relevance and physiological arousal compared with emotional words (Hinojosa, Carretié, Valcárcel, Méndez-Bértolo, & Pozo, 2009). Furthermore, valence and arousal properties may distinctly modulate ERP components as a function of attention focus. For example, Delaney-Busch et al. (2016) demonstrated that, when participants were overtly instructed to attend to the affective properties of the stimuli, the LPC was modulated by stimulus valence but not by arousal. On the contrary, when the affective properties of the stimuli were task-irrelevant (as in the current study), the LPC was modulated by arousal but not by valence. Given that words were controlled for arousal in the current study, the lack of a valence effect could be related to the implicit (vs. explicit) focus on the affective properties of the stimuli. This possibility agrees with the notion that the impact of emotion on the LPC and subsequent late components is likely to be attenuated when participants' attention is diverted away from the affective content of the stimuli (Hajcak et al., 2006; Hajcak & Nieuwenhuis, 2006). This also explains why the only difference observed during encoding was between encoding tasks as participants were overtly instructed to pay attention to the stimulus and encoding task, which were consistently highlighted and tested in distinct study-test cycles.

Emotion and self-reference effects during recognition

In the early frontal old/new effect time window, which has been typically associated with familiarity (Curran, 2000; Curran & Dien, 2003; Czernochowski et al., 2005), no valence-related modulations were observed, supporting previous studies (Johansson et al., 2004; Koenig & Mecklinger, 2008; Lavoie & O'Connor, 2013; Maratos et al., 2000; Minor & Herzmann, 2019; Newsome et al., 2012; Weymar et al., 2009). Nevertheless, the current study showed a self-reference effect, as old/new distinctions were only seen for words studied in the self-referential condition. Additionally, the electrophysiological response associated with common-source hits was similar to both self-source hits and correct rejections. Although prior studies showed that familiarity processes are sensitive to the type of encoding task (Nyhus & Curran, 2009; Peters & Daum, 2009) and source details (Mollison & Curran, 2012), the lack of difference between common-source and correct rejections was not anticipated, especially considering the use of intentional encoding conditions and evidence from prior studies using encoding tasks with deep processing demands (Dulas et al., 2011; Marzi & Viggiano, 2010). Instead, the results appear to support the view that self-referential processing enhances familiarity processes associated with source monitoring (Mao et al., 2017). This advantage dovetails with the unitization account, which occur when item and source features are combined and integrated into a meaningful unit of information during encoding, resulting in a facilitated recognition of source features driven by familiarity processes (Diana, Van Den Boom, Yonelinas, & Ranganath, 2011; Diana, Yonelinas, & Ranganath, 2008; Kuo & Van Petten, 2006). Nonetheless, more evidence is required to support these claims as in the current study the electrophysiological response in the first latency window (300-500 ms) did not differ as a function of self-relevance (i.e., was similar for self- and common-source hits), which might be attributed to limited statistical power.

When considering the parietal old/new effect, typically associated with the recollection of information from the study episode including source features (Addante et al., 2012; Allan & Rugg, 1998; Cansino & Trejo-Morales, 2008; Leynes et al., 2017; MacLeod & Donaldson, 2017; Peters & Daum, 2009; Trott et al., 1999; Wilding & Rugg, 1997; Woroch & Gonsalves, 2010), the expected advantage for words encoded self-referentially was not observed. Because the amplitude of the parietal old/new effect is sensitive to variations in the amount of recoverable information from the study episode (Leynes et al., 2005; Leynes & Crawford, 2018; Wilding, 2000; Woroch & Gonsalves, 2010), this result suggests that the amount of information retrieved was likely similar across encoding tasks, regardless of whether it was or not source-specifying (i.e., diagnostic; see Leynes & Crawford, 2018; Leynes & Mok, 2017).

Nonetheless, whereas old/new effects were observed for neutral and positive words irrespective of encoding task, in the case of negative words the old/new effect was detected in the self-referential condition. The absence of a similar effect in the common condition demonstrates that emotional information is not always associated with enhanced recollection, as already documented by prior studies (Koenig & Mecklinger, 2008;

Lavoie & O'Connor, 2013; Mao et al., 2015; Maratos et al., 2000). Instead, old/new effects may or may not occur for distinct emotional conditions depending on the type of source features tested during recognition (Koenig & Mecklinger, 2008). Specifically, negative valence might hinder the recollection of extrinsic source details, i.e., attributes that are not embedded in the stimulus (see the object-based binding theory by Mather, 2007), such as the encoding task used during the study phase (Mao et al., 2015). It could be argued that the same was not observed for negative words encoded self-referentially as stimulus and encoding task were likely unitized during encoding, thus favoring the recollection of contextual information.

Importantly, the analysis of right-frontal old/new effect (800–1,200 ms) revealed that self-referential negative words and negative correct rejections elicited larger mean amplitudes than negative words studied in the common condition. Furthermore, in the common condition, the mean amplitude for neutral words was larger than for negative words. The old/new effect has been associated with post-retrieval monitoring processes that are sensitive to retrieval demands (Donaldson & Rugg, 1998; Leynes, 2002; Leynes et al., 2006; Trott et al., 1999; Van Petten et al., 2000; Wilding & Rugg, 1996), including difficult memory decisions (Rugg et al., 2000). Thereby, the former differences seem to suggest that decisions associated with non-self-referential neutral words and self-referential negative words more strongly engage post-retrieval monitoring resources.

The difference between negative self- and common-source hits may be related to the fact that self-referential conditions allowed the retrieval of details containing less source-specifying or diagnostic of the encoding task. Thus, negative stimuli in such conditions demanded a greater involvement of monitoring processes (see Koenig & Mecklinger, 2008; Newsome et al., 2012 for similar arguments). Considering that source memory performance for negative self-referential (vs. neutral and positive self-referential) words was found to be impaired, it is possible that participants conducted a secondary memory search to resolve possible uncertainties regarding the association between negative information and the self (Dulas & Duarte, 2013; Senkfor & Van Petten, 1998). Such uncertainties are likely to arise if we consider that negative information tends to be less relatable to the self (D'Argembeau & Van der Linden, 2008; Derry & Kuiper, 1981; Lewis et al., 2011; Moran et al., 2006; Pauly et al., 2013). Additionally, the fact that negative words were generally associated with more difficult memory decisions during recognition might explain the larger ERP amplitude in response to negative new words, reflecting a strategic monitoring of these items to ensure no prior links to the self.

An LPN effect was not observed in the 1,200–1,800 ms latency window, which suggests that the discrimination between the two source features did not impose further evaluation or reconstruction of task-relevant elements from the study episode. This outcome was expected given the use of intentional encoding conditions that emphasized both stimulus and encoding task, and the fact that these elements were also targeted during recognition, which might have facilitated memory retrieval according to the encoding specificity principle (Tulving & Thomson, 1973). In such experimental conditions, it is likely that source-specifying information is readily available for most retrieval attempts, thus resulting in an absent LPN effect (Mecklinger et al., 2016).

Limitations

One of the main limitations of this study was the number of stimuli per experimental condition, which restricted several aspects of the analysis during encoding and recognition. For instance, it was not possible to compare electrophysiological activity between remembered and forgotten items during word encoding (e.g., difference due to subsequent memory - Dm effects: Dolcos & Cabeza, 2002; Leynes & Crawford, 2018). Importantly, it was not possible to contrast ERP responses elicited by correct versus incorrect source attributions, which should be examined in future studies. This would clarify how valence and self-referential encoding influence source memory recognition in accurate (containing source-specifying features) versus erroneous source decisions (characterized by retrieval attempts containing other features that are not source-specifying; Cansino et al., 2012; Cansino & Trejo-Morales, 2008; Leynes & Mok, 2017; Mao et al., 2015). Together with fact that the sample was mainly composed by female participants (see Glaser, Mendrek, Germain, Lakis, & Lavoie, 2012 for an example of sex-related memory variations), these factors limit the generalization of the current findings. Additionally, given that self-referential conditions grant an advantage in terms of memory performance, this directly influenced the number of artifact-free trials available for the ERP analysis during recognition, creating an imbalance between encoding tasks (Supplementary Table S2). In this regard, future studies may use a self-referential encoding strategy and probe its impact on other source features, such as spatial location, stimulus color, or even other encoding tasks (Hou et al., 2019; Yin et al., 2019).

Conclusions

The current study revealed that valence and self-reference interactively modulate both behavioral and ERP indices of source memory. Specifically, self-referential processing resulted in improved internal source memory for positive and neutral words. Electrophysiologically, the self-reference advantage was supported by processes occurring during encoding and recognition. During encoding, self-referential conditions led to amplitude enhancements (700–950 and 950–1,200 ms latency windows), possibly associated with a greater mobilization of sustained elaboration and rehearsal resources. During recognition, self-reference effects were manifested in an early processing stage typically related to familiarity, supporting previous claims that familiarity can contribute to source memory recognition (Diana et al., 2008, 2011; Kuo & Van Petten, 2006; Mao et al., 2015). Nonetheless, interactions between self-reference and valence were particularly observed for negative information in subsequent latency windows (500–800 ms; 800–1,200 ms), corroborating the view that negative valence may affect the retrieval of extrinsic source features, such as the task used during encoding (Mao et al., 2015). Ultimately, unraveling how emotional self-referential information affects distinct source memory features (e.g., perceptual; spatio-temporal; semantic; affective; cognitive operations; see Johnson et al., 1993) may contribute to a better understanding of the neurocognitive mechanisms underpinning memory dysfunction in clinical conditions that implicate alterations of the positive self-view, such as such anxiety and depression (Auerbach et al., 2015; Pauly et al., 2013; Shestyuk & Deldin, 2010).

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13415-020-00858-6.

Acknowledgments This work was supported by a PhD Fellowship (PD/BD/105964/2014), awarded to DRP, funded by the Portuguese Foundation for Science and Technology (FCT) through national funds and co-funded by the European Social Fund (ESF) through the Operational Programme for Human Capital (POCH). It also was supported by a research grant (PTDC/MHC-PCN/0101/2014; PTDC/PSI-GER/32152/2017) funded by FCT and awarded to APP and AS, respectively. This study was conducted at the Psychology Research Centre (PSI/01662), School of Psychology, University of Minho, supported by the FCT through the Portuguese State Budget (UIDB/PSI/01662/2020). The authors acknowledge the participants for their valuable collaboration in this study and our colleagues from the Psychological Neuroscience Lab for the help provided during data acquisition.

Compliance with ethical standards

Conflict of interest None declared.

References

Addante, Richard J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: evidence of accurate source recognition without recollection. *NeuroImage*, *62*(1), 439–450. https://doi.org/10.1016/j.neuroimage.2012.04.031

Allan, K., & Rugg, M. D. (1998). Neural correlates of cued recall with and without retrieval of source memory. *NeuroReport*, *9*(15), 3463–3466. https://doi.org/10.1097/00001756-199810260-00023

Auerbach, R. P., Stanton, C. H., Proudfit, G. H., & Pizzagalli, D. A. (2015). Self-referential processing in depressed adolescents: A high- density ERP study. *Journal of Abnormal Psychology*, *124*(2), 233–245. https://doi.org/10.1037/abn0000023

Barnacle, G. E., Tsivilis, D., Schaefer, A., & Talmi, D. (2018). Local context influences memory for emotional stimuli but not electrophysiological markers of emotion-dependent attention. *Psychophysiology*, 55(4), e13014. https://doi.org/10.1111/psyp.13014

Beck, A. T., Steer, R., & Brown, G. (1996). *Beck Depression Inventory-II manual*. New York: The Psychological Corporation.

Bradley, M. M., & Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction manual and affective ratings*. Gainesville, FL: Center for Research in Psychophysiology, University of Florida.

Cansino, S., Hernández-Ramos, E., & Trejo-Morales, P. (2012). Neural correlates of source memory retrieval in young, middle-aged and elderly adults. *Biological Psychology*, *90*(1), 33–49. https://doi.org/10.1016/j.biopsycho.2012.02.004

Cansino, S., & Trejo-Morales, P. (2008). Neurophysiology of successful encoding and retrieval of source memory. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(1), 85–98. https://doi.org/10.3758/CABN.8.1.85

Chaumon, M., Bishop, D. V. M., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*, *250*, 47–63. https://doi.org/10.1016/j.jneumeth.2015.02.025

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

Codispoti, M., Ferrari, V., & Bradley, M. M. (2007). Repetition and event-related potentials: Distinguishing early and late processes in affective picture perception. *Journal of Cognitive Neuroscience*, *19*(4), 577–586. https://doi.org/10.1162/jocn.2007.19.4.577

Coelho, R., Martins, A., & Barros, H. (2002). Clinical profiles relating gender and depressive symptoms among adolescents ascertained by the Beck Depressive Inventory II. *European Psychiatry, 17*(4), 222-226. https://doi.org/10.1016/S0924-9338(02)00663-6

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

Conway, M. A., Singer, J. F., & Tagini, A. (2004). The self and autobiographical memory: Correspondence and coherence. *Social Cognition*, *22*(5), 491–529. https://doi.org/10.1521/soco.22.5.491.50768

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

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*(6), 923–938. https://doi.org/10.3758/BF03209340

Curran, T., & Dien, J. (2003). Differentiating amodal familiarity from modality-specific memory processes: An ERP study. *Psychophysiology*, *40*(6), 979–988. https://doi.org/10.1111/1469-8986.00116

Cycowicz, Y., & Friedman, D. (2003). Source memory for the color of pictures: Event-related brain potentials(ERPs) reveal sensory-specific retrieval-related activity. *Psychophysiology*, *40*(3), 455–464. https://doi.org/10.1111/1469-8986.00047

Czernochowski, D., Mecklinger, A., Johansson, M., & Brinkmann, M. (2005). Age-related differences in familiarity and recollection: ERP evidence from a recognition memory study in children and young adults. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(4), 417–433. https://doi.org/10.3758/CABN.5.4.417

D'Argembeau, A., & Van der Linden, M. (2008). Remembering pride and shame: Self-enhancement and the phenomenology of autobiographical memory. *Memory*, *16*(5), 534–547. https://doi.org/10.1080/09658210802010463

Delaney-Busch, N., Wilkie, G., & Kuperberg, G. (2016). Vivid: How valence and arousal influence word processing under different task demands. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(3), 415–432. https://doi.org/10.3758/s13415-016-0402-y

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Derry, P. A., & Kuiper, N. A. (1981). Schematic processing and self-reference in clinical depression. *Journal of Abnormal Psychology*, *90*(4), 286–297. https://doi.org/10.1037/0021-843X.90.4.286

Diana, R. A., Van Den Boom, W., Yonelinas, A. P., & Ranganath, C. (2011). ERP correlates of source memory: Unitized source information increases familiarity-based retrieval. *Brain Research*, *1367*, 278–286. https://doi.org/10.1016/j.brainres.2010.10.030

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

Diedrich, O., Naumann, E., Maier, S., Becker, G., & Bartussek, D. (1997). A frontal positive slow wave in the ERP associated with emotional slides. *Journal of Psychophysiology, 11*(1), 71-84.

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: Encoding pleasant, unpleasant, and neutral pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(3), 252–263. https://doi.org/10.3758/CABN.2.3.252

Donaldson, D. I., & Rugg, M. D. (1998). Recognition memory for new associations: Electrophysiological evidence for the role of recollection. *Neuropsychologia*, *36*(5), 377–395. https://doi.org/10.1016/S0028-3932(97)00143-7

Dulas, M. R., & Duarte, A. (2012). The effects of aging on material-independent and material-dependent neural correlates of source memory retrieval. *Cerebral Cortex*, *22*(1), 37–50. https://doi.org/10.1093/cercor/bhr056

Dulas, M. R., & Duarte, A. (2013). The influence of directed attention at encoding on source memory retrieval in the young and old: An ERP study. *Brain Research*, *1500*, 55–71. https://doi.org/10.1016/j.brainres.2013.01.018

Dulas, M. R., & Duarte, A. (2014). Aging affects the interaction between attentional control and source memory: An fMRI study. *Journal of Cognitive Neuroscience*, *26*(12), 2653–2669. https://doi.org/10.1162/jocn_a_00663

Dulas, M. R., Newsome, R. N., & Duarte, A. (2011). The effects of aging on ERP correlates of source memory retrieval for self-referential information. *Brain Research*, *1377*, 84–100. https://doi.org/10.1016/j.brainres.2010.12.087

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

Fan, W., Chen, J., Wang, X. Y., Cai, R., Tan, Q., Chen, Y., …Zhong, Y. (2013). Electrophysiological correlation of the degree of self-reference effect. *PLoS ONE*, *8*(12), e80289. https://doi.org/10.1371/journal.pone.0080289

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

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: Words, phrases, and task effects. *Progress in Brain Research*, *156*, 185–203. https://doi.org/10.1016/S0079-6123(06)56009-1

Foti, D., Hajcak, G., & Dien, J. (2009). Differentiating neural responses to emotional pictures: Evidence from temporal-spatial PCA. *Psychophysiology*, *46*(3), 521–530. https://doi.org/10.1111/j.1469-8986.2009.00796.x

Friedman, D., Cycowicz, Y. M., & Bersick, M. (2005). The late negative episodic memory effect: The effect of recapitulating study details at test. *Cognitive Brain Research*, *23*(2–3), 185–198. https://doi.org/10.1016/j.cogbrainres.2004.10.005

Friedman, D., & Johnson, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy Research and Technique*, *51*(1), 6–28. https://doi.org/10.1002/1097-0029(20001001)51:1<6::AID-JEMT2>3.0.CO;2-R

Gallant, S. N., Pun, C., & Yang, L. (2018). Age differences in the neural correlates underlying control of emotional memory: An event-related potential study. *Brain Research*, *1697*, 83–92. https://doi.org/10.1016/j.brainres.2018.06.019

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

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

Glaser, E., Mendrek, A., Germain, M., Lakis, N., & Lavoie, M. E. (2012). Sex differences in memory of emotional images: A behavioral and electrophysiological investigation. *International Journal of Psychophysiology*, *85*(1), 17–26. https://doi.org/10.1016/j.ijpsycho.2012.01.007

Grundy, J. G., Benarroch, M. F. F., Lebarr, A. N., & Shedden, J. M. (2015). Electrophysiological correlates of implicit valenced self-processing in high vs. low self-esteem individuals. *Social Neuroscience*, *10*(1), 100–112. https://doi.org/10.1080/17470919.2014.965339

Gutchess, A., & Kensinger, E. A. (2018). Shared mechanisms may support mnemonic benefits from self-referencing and emotion. *Trends in Cognitive Sciences*, *22*(8), 712–724. https://doi.org/10.1016/j.tics.2018.05.001

Hajcak, G., Moser, J. S., & Simons, R. F. (2006). Attending to affect: Appraisal strategies modulate the electrocortical response to arousing pictures. *Emotion*, *6*(3), 517–522. https://doi.org/10.1037/1528-3542.6.3.517

Hajcak, G., & Nieuwenhuis, S. (2006). Reappraisal modulates the electrocortical response to unpleasant pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(4), 291–297. https://doi.org/10.3758/CABN.6.4.291

Herbert, C., Herbert, B. M., Ethofer, T., & Pauli, P. (2010). His or mine? The time course of self-other discrimination in emotion processing. *Social Neuroscience*, *6*(3), 277–288. https://doi.org/10.1080/17470919.2010.523543

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

Herbert, C., Pauli, P., & Herbert, B. M. (2011). Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Social Cognitive and Affective Neuroscience*, *6*(5), 653–661. https://doi.org/10.1093/scan/nsq082

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

Hinojosa, J. A., Méndez-Bértolo, C., & Pozo, M. A. (2010). Looking at emotional words is not the same as reading emotional words: Behavioral and neural correlates. *Psychophysiology*, *47*(4), 748–757. https://doi.org/10.1111/j.1469-8986.2010.00982.x

Hou, M., Grilli, M. D., & Glisky, E. L. (2019). Self-reference enhances relational memory in young and older adults. *Aging,* *Neuropsychology, and Cognition, 26*(1), 105-120. https://doi.org/10.1080/13825585.2017.1409333

Hudson, A., Wilson, M. J. G., Green, E. S., Itier, R. J., & Henderson, H. A. (2020). Are you as important as me? Self-other discrimination within trait-adjective processing. *Brain and Cognition, 142*, 105569. https://doi.org/10.1016/j.bandc.2020.105569

Inaba, M., Nomura, M., & Ohira, H. (2005). Neural evidence of effects of emotional valence on word recognition. *International Journal of Psychophysiology*, *57*(3), 165–173. https://doi.org/10.1016/j.ijpsycho.2005.01.002

Johansson, M., & Mecklinger, A. (2003). The late posterior negativity in ERP studies of episodic memory: Action monitoring and retrieval of attribute conjunctions. *Biological Psychology*, *64*(1–2), 91–117. https://doi.org/10.1016/S0301-0511(03)00104-2

Johansson, M., Mecklinger, A., & Treese, A. C. (2004). Recognition memory for emotional and neutral faces: An event-related potential study. *Journal of Cognitive Neuroscience*, *16*(10), 1840–1853. https://doi.org/10.1162/0898929042947883

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

Katyal, S., Hajcak, G., Flora, T., Bartlett, A., & Goldin, P. (2020). Event-related potential and behavioural differences in affective self-referential processing in long-term meditators versus controls. *Cognitive, Affective, & Behavioral Neuroscience, 20***,** 326–339. https://doi.org/10.3758/s13415-020-00771-y

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

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

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

Kuo, T. Y., & Van Petten, C. (2006). Prefrontal engagement during source memory retrieval depends on the prior encoding task. *Journal of Cognitive Neuroscience*, *18*(7), 1133–1146. https://doi.org/10.1162/jocn.2006.18.7.1133

Langeslag, S. J. E., & Van Strien, J. W. (2008). Age differences in the emotional modulation of ERP old/new effects. *International Journal of Psychophysiology*, *70*(2), 105–114. https://doi.org/10.1016/j.ijpsycho.2008.07.022

Lavoie, M. E., & O'Connor, K. P. (2013). Effect of emotional valence on episodic memory stages as indexed by event-related potentials. *World Journal of Neuroscience*, *03*(04), 250–262. https://doi.org/10.4236/wjns.2013.34034

Leshikar, E. D., & Duarte, A. (2012). Medial prefrontal cortex supports memory accuracy for self-referenced items. *Social Neuroscience*, *7*(2), 126–145. https://doi.org/10.1080/17470919.2011.585242

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

Lewis, P. A., Cairney, S., Manning, L., & Critchley, H. D. (2011). The impact of overnight consolidation upon memory for emotional and neutral encoding contexts. *Neuropsychologia*, *49*(9), 2619–2629. https://doi.org/10.1016/j.neuropsychologia.2011.05.009

Lewis, P. A., Critchley, H. D., Rotshtein, P., & Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cerebral Cortex*, *17*(3), 742–748. https://doi.org/10.1093/cercor/bhk024

Leynes, P. A. (2002). The effect of specific test queries on source-monitoring event-related potentials. *Brain and Cognition*, *50*(2), 218–233. https://doi.org/10.1016/S0278-2626(02)00505-5

Leynes, P. A., Askin, B., & Landau, J. D. (2017). Visual perspective during remembering: ERP evidence of familiarity-based source monitoring. *Cortex*, *91*, 157–168. https://doi.org/10.1016/j.cortex.2016.12.003

Leynes, P. A., & Crawford, C. J. (2018). Event-related potential (ERP) evidence that encoding focus alters recollected features. *Brain and Cognition*, *127*, 42–50. https://doi.org/10.1016/j.bandc.2018.09.005

Leynes, P. A., Crawford, J. T., & Bink, M. L. (2005). Interrupted actions affect output monitoring and event-related potentials (ERPs). *Memory*, *13*(7), 759–772. https://doi.org/10.1080/09658210444000377

Leynes, P. A., Grey, J. A., & Crawford, J. T. (2006). Event-related potential (ERP) evidence for sensory-based action memories. *International Journal of Psychophysiology*, *62*(1), 193–202. https://doi.org/10.1016/j.ijpsycho.2006.04.003

Leynes, P. A., & Mok, B. A. (2017). Encoding focus alters diagnostic recollection and event-related potentials (ERPs). *Brain and Cognition*, *117*, 1–11. https://doi.org/10.1016/j.bandc.2017.06.011

Leynes, P. A., & Phillips, M. C. (2008). Event-related potential (ERP) evidence for varied recollection during source monitoring. *Journal of Experimental Psychology: Learning Memory and Cognition*, *34*(4), 741–751. https://doi.org/10.1037/0278-7393.34.4.741

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

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

MacLeod, C. A., & Donaldson, D. I. (2017). Investigating the functional utility of the left parietal ERP old/new effect: Brain activity predicts within but not between participant variance in episodic recollection. *Frontiers in Human Neuroscience*, *11*, 580. https://doi.org/10.3389/fnhum.2017.00580

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

Mao, X., Wang, Y., Wu, Y., & Guo, C. (2017). Self-referential information alleviates retrieval inhibition of directed forgetting effects—an ERP evidence of source memory. *Frontiers in Behavioral Neuroscience*, *11*, 187. https://doi.org/10.3389/fnbeh.2017.00187

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

Maratos, E. J., Allan, K., & Rugg, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, *38*(11), 1452–1465. https://doi.org/10.1016/S0028-3932(00)00061-0

Marzi, T., & Viggiano, M. P. (2010). Deep and shallow encoding effects on face recognition: An ERP study. *International Journal of Psychophysiology*, *78*(3), 239–250. https://doi.org/10.1016/j.ijpsycho.2010.08.005

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

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

McNeely, H. E., Dywan, J., & Segalowitz, S. J. (2004). ERP indices of emotionality and semantic cohesiveness during recognition judgments. *Psychophysiology*, *41*(1), 117–129. https://doi.org/10.1111/j.1469-8986.2003.00137.x

Mecklinger, A., Johansson, M., Parra, M., & Hanslmayr, S. (2007). Source-retrieval requirements influence late ERP and EEG memory effects. *Brain Research*, *1172*, 110–123. https://doi.org/10.1016/j.brainres.2007.07.070

Mecklinger, A., Rosburg, T., & Johansson, M. (2016). Reconstructing the past: The late posterior negativity (LPN) in episodic memory studies. *Neuroscience and Biobehavioral Reviews*, *68*, 621–638. https://doi.org/10.1016/j.neubiorev.2016.06.024

Meng, X., Zhang, L., Liu, W., Ding, X. S., Li, H., Yang, J., & Yuan, J. J. (2017). The impact of emotion intensity on recognition memory: Valence polarity matters. *International Journal of Psychophysiology*, *116*, 16–25. https://doi.org/10.1016/j.ijpsycho.2017.01.014

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

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*(2), 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x

Mollison, M. V., & Curran, T. (2012). Familiarity in source memory. *Neuropsychologia*, *50*(11), 2546–2565. https://doi.org/10.1016/j.neuropsychologia.2012.06.027

Moran, J. M., Macrae, C. N., Heatherton, T. F., Wyland, C. L., & Kelley, W. M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive Neuroscience*, *18*(9), 1586–1594. https://doi.org/10.1162/jocn.2006.18.9.1586

Mu, Y., & Han, S. (2010). Neural oscillations involved in self-referential processing. *NeuroImage*, *53*(2), 757–768. https://doi.org/10.1016/j.neuroimage.2010.07.008

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

Nowicka, M. M., Wójcik, M. J., Kotlewska, I., Bola, M., & Nowicka, A. (2018). The impact of self-esteem on the preferential processing of self-related information: Electrophysiological correlates of explicit self vs. other evaluation. *PLoS ONE*, *13*(7), e0200604. https://doi.org/10.1371/journal.pone.0200604

Nyhus, E., & Curran, T. (2009). Semantic and perceptual effects on recognition memory: Evidence from ERP. *Brain Research*, *1283*, 102–114. https://doi.org/10.1016/j.brainres.2009.05.091

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

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

Palomba, D., Angrilli, A., & Mini, A. (1997). Visual evoked potentials, heart rate responses and memory to emotional pictorial stimuli. *International Journal of Psychophysiology*, *27*(1), 55–67. https://doi.org/10.1016/S0167-8760(97)00751-4

Pauly, K., Finkelmeyer, A., Schneider, F., & Habel, U. (2013). The neural correlates of positive self-evaluation and self-related memory. *Social Cognitive and Affective Neuroscience*, *8*(8), 878–886. https://doi.org/10.1093/scan/nss086

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

Pereira, D. R., Sampaio, A., & Pinheiro, A. P. (2020). Is internal source memory recognition modulated by emotional encoding contexts? *Psychological Research*. Advance online publication. https://doi.org/10.1007/s00426-020-01294-4

Peters, J., & Daum, I. (2009). Frontal but not parietal positivity during source recollection is sensitive to episodic content. *Neuroscience Letters*, *454*(3), 182–186. https://doi.org/10.1016/j.neulet.2009.03.019

Pion-Tonachini, L., Makeig, S., & Kreutz-Delgado, K. (2017). Crowd labeling latent Dirichlet allocation. *Knowledge and Information Systems*, *53*(3), 749–765. https://doi.org/10.1007/s10115-017-1053-1

Rugg, M., Allan, K., & Birch, C. (2000). Electrophysiological evidence for the modulation of retrieval orientation. *Journal of Cognitive Neuroscience*, *12*(4), 664–678. https://doi.org/10.1162/089892900562291

Schacht, A., & Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(1), 28–43. https://doi.org/10.3758/CABN.9.1.28

Schaefer, A., Fletcher, K., Pottage, C. L., Alexander, K., & Brown, C. (2009). The effects of emotional intensity on ERP correlates of recognition memory. *NeuroReport*, *20*(3), 319–324. https://doi.org/10.1097/WNR.0b013e3283229b52

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

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: Event-related brain potential studies. *Progress in Brain Research, 156*, 31-51. https://doi.org/10.1016/S0079-6123(06)56002-9

Senkfor, A. J., & Van Petten, C. (1998). Who said what? An event-related potential investigation of source and item memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *24*(4), 1005–1025. https://doi.org/10.1037/0278-7393.24.4.1005

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

Shestyuk, A. Y., & Deldin, P. J. (2010). Automatic and strategic representation of the self in major depression: Trait and state abnormalities. *American Journal of Psychiatry*, *167*(5), 536–544. https://doi.org/10.1176/appi.ajp.2009.06091444

Silva, D. R., & Spielberger, C. D. (2007). *Manual do Inventário de Estado Traço de Ansiedade (STAI)*. USA: Mind Garden, Inc.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Soares, A. P., Comesaña, M., Pinheiro, A. P., Simões, A., & Frade, C. S. (2012). The adaptation of the Affective Norms for English Words (ANEW) for European Portuguese. *Behavior Research Methods, 44*(1), 256-269. https://doi.org/10.3758/s13428-011-0131-7

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Trott, C. T., Friedman, D., Ritter, W., Fabiani, M., & Snodgrass, J. G. (1999). Episodic priming and memory for temporal source: Event-related potentials reveal age-related differences in prefrontal functioning. *Psychology and Aging*, *14*(3), 390–413. https://doi.org/10.1037/0882-7974.14.3.390

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

Van Petten, C., Senkfor, A. J., & Newberg, W. M. (2000). Memory for drawings in locations: Spatial source memory and event-related potentials. *Psychophysiology*, *37*(4), 551–564. https://doi.org/10.1017/S0048577200990541

Vilberg, K. L., & Rugg, M. D. (2009). An investigation of the effects of relative probability of old and new test items on the neural correlates of successful and unsuccessful source memory. *NeuroImage*, *45*(2), 562–571. https://doi.org/10.1016/j.neuroimage.2008.12.020

Voss, J. L., & Federmeier, K. D. (2011). FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. *Psychophysiology*, *48*(4), 532–546. https://doi.org/10.1111/j.1469-8986.2010.01085.x

Voss, J. L., & Paller, K. A. (2017). Neural substrates of remembering: Event-related potential studies. In *Learning and memory: A comprehensive reference* (pp. 81-98). Elsevier Inc. https://doi.org/10.1016/b978-0-12-809324-5.21070-5

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

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

Weymar, M., Löw, A., Melzig, C. A., & Hamm, A. O. (2009). Enhanced long-term recollection for emotional pictures: Evidence from high-density ERPs. *Psychophysiology*, *46*(6), 1200–1207. https://doi.org/10.1111/j.1469-8986.2009.00869.x

Wilding, E. L. (1999). Separating retrieval strategies from retrieval success: An event-related potential study of source memory. *Neuropsychologia*, *37*(4), 441–454. https://doi.org/10.1016/S0028-3932(98)00100-6

Wilding, E. L. (2000). In what way does the parietal ERP old/new effect index recollection? *International Journal of Psychophysiology*, *35*(1), 81–87. https://doi.org/10.1016/S0167-8760(99)00095-1

Wilding, E. L., Rugg, M. D. (1996). An event-related potential study of recognition with and without retrieval of source. *Brain: A Journal of Neurology*, *119*(3), 889–905. https://doi.org/10.1093/brain/119.3.889

Wilding, E. L., & Rugg, M. D. (1997). An event-related potential study of memory for spoken and heard words. *Neuropsychologia*, *35*(9), 1185–1195. https://doi.org/10.1016/S0028-3932(97)00048-1

Windmann, S., & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, *13*(5), 577–592. https://doi.org/10.1162/089892901750363172

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

Woroch, B., & Gonsalves, B. D. (2010). Event-related potential correlates of item and source memory strength. *Brain Research*, *1317*, 180–191. https://doi.org/10.1016/j.brainres.2009.12.074

Xu, H., Zhang, Q., Li, B., & Guo, C. (2015). Dissociable effects of valence and arousal on different subtypes of old/new effect: Evidence from event-related potentials. *Frontiers in Human Neuroscience*, *9*, 650. https://doi.org/10.3389/fnhum.2015.00650

Yang, L., Truong, L., Fuss, S., & Bislimovic, S. (2012). The effects of ageing and divided attention on the self-reference effect in emotional memory: Spontaneous or effortful mnemonic benefits? *Memory*, *20*(6), 596–607. https://doi.org/10.1080/09658211.2012.690040

Yin, X., Ma, Y., Xu, X., & Yang, H. (2019). The effect of self-referencing on memory for different kinds of source information. *Memory*, *27*(4), 519–527. https://doi.org/10.1080/09658211.2018.1532009

Zhao, Y., Luo, W., Chen, J., Zhang, D., Zhang, L., Xiao, C., …Tan, S. (2016). Behavioral and neural correlates of self-referential processing deficits in bipolar disorder. *Scientific Reports, 6*, 24075. https://doi.org/10.1038/srep24075

Zhang, H., Guan, L., Qi, M., & Yang, J. (2013). Self-esteem modulates the time course of self-positivity bias in explicit self-evaluation. *PLoS ONE*, *8*(12), e81169. https://doi.org/10.1371/journal.pone.0081169

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

Zhou, H., Guo, J., Ma, X., Zhang, M., Liu, L., Feng, L., …Zhong, N. (2017). Self-reference emerges earlier than emotion during an implicit self-referential emotion processing task: Event-related potential evidence. *Frontiers in Human Neuroscience*, *11*, 451. https://doi.org/10.3389/fnhum.2017.00451

Open practices statement The data on which this study is based will be made available upon reasonable request to the corresponding author's email. However, the data cannot be publicly available as we did not collect participants' consent with respect to this issue. This study was not preregistered.