Received: 3 February 2021 | Accepted: 18 February 2021

DOI: 10.1002/hbm.25397

RESEARCH ARTICLE

WILEY

Retrospective confidence judgments: Meta-analysis of functional magnetic resonance imaging studies

Beatriz Martín-Luengo[1] | **Oksana Zinchenko**[1,2] | **Aleksandra Dolgoarshinnaia**[1] | **Alina Leminen**[3,4]

[1]Centre for Cognition and Decision Making, Institute for Cognitive Neuroscience, National Research University Higher School of Economics, Moscow, Russian Federation

[2]International Laboratory for Social Neurobiology, Institute for Cognitive Neuroscience, National Research University Higher School of Economics, Moscow, Russian Federation

[3]Cognitive Science, Department of Digital Humanities, Faculty of Arts, University of Helsinki, Helsinki, Finland

[4]Cognitive Brain Research Unit, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Correspondence
Beatriz Martín-Luengo, Centre for Cognition and Decision Making, Institute for Cognitive Neuroscience, National Research University Higher School of Economics, Krivokolenniy Sidewalk 3, Entrance 2, 101000 Moscow, Russian Federation.
Email: bmartinluengo@hse.ru

Funding information
Russian Science Foundation, Grant/Award Number: 19 18 00534

Abstract

Confidence in our retrieved memories, that is, retrospective confidence, is a metamemory process we perform daily. There is an abundance of applied research focusing on the metamemory judgments and very diverse studies including a wide range of clinical populations. However, the neural correlates that support its functioning are not well defined impeding the implementation of noninvasive neuromodulatory clinical interventions. To address the neural basis of metamemory judgments, we ran a meta-analysis, where we used the activation likelihood estimation method on the 19 eligible functional magnetic resonance imaging studies. The main analysis of retrospective confidence revealed concordant bilateral activation in the parahippocampal gyrus, left middle frontal gyrus, and right amygdala. We also run an analysis between the two extreme levels of confidence, namely, high and low. This additional analysis was exploratory, since the minimum amount of articles reporting these two levels was not reached. Activations for the exploratory high > low confidence subtraction analysis were the same as observed in the main analysis on retrospective confidence, whereas the exploratory low > high subtraction showed distinctive activations of the right precuneus. The involvement of the right precuneus emphasizes its role in the evaluation of low confidence memories, as suggested by previous studies. Overall, our study contributes to a better understanding of the specific brain structures involved in confidence evaluations. Better understanding of the neural basis of metamemory might eventually lead to designing more precise neuromodulatory interventions, significantly improving treatment of patients suffering from metamemory problems.

KEYWORDS

activation likelihood estimate, brain mapping, functional magnetic resonance imaging, metamemory, retrospective confidence

1 | INTRODUCTION

Metamemory can be defined as the knowledge we have about our memory, how it works and which strategies we need to implement in order to maximize it (Nelson & Narens, 1994). Metamemory is one of the most important superior cognitive functions due to its relevance in our daily life, particularly in the selection of encoding and retrieval strategies (Pannu & Kaszniak, 2005). For example, when we are learning new material (e.g., vocabulary of another language) we have to decide whether we have studied enough to remember it later. If we believe that we have invested enough time in learning, we will stop studying, whereas in the contrary case we will continue studying it or try a different approach. Moreover, understanding metamemory is essential for clinical populations—wrong beliefs about how memory works are transformed into a wrong perception of reality (Moritz, Woodward, Jelinek, & Klige, 2008). In the rehabilitation of memory processes, awareness of their deficit in the metacognitive processes increases the benefits from rehabilitation (Clare & Woods, 2004).

The seminal work about the neural basis for metamemory was initiated by Shimamura and Squire (1986) Korsakoff syndrome patients using "feeling of knowing judgments," that is, the confidence in recognizing current inaccessible information. Later on, many studies on metamemory measures have been conducted with schizophrenia patients, reporting that these patients are unaware of their damage on memory (Moritz & Woodward, 2006; Moritz, Woodward, & Chen, 2006). In particular, patients with schizophrenia suffer from the so-called "confidence gap," that is the impairment to distinguish between correct and incorrect answers, which is reflected in high confidence ratings for their incorrect answers and low confidence for the correct ones. This phenomenon has been observed with remember-known tasks (Moritz & Woodward, 2006), numeric confidence scales (Moritz et al., 2008), and it has been replicated even in patients that only suffered from a single schizophrenic episode (Moritz, Woodward, & Chen, 2006). The robustness of this effect with different measures and particularly at different degree levels pointed to the prefrontal cortex as the hub for confidence assessment, since it is an area also affected in Korsakoff patients.

These and other early works on clinical patients were crucial in advancing research on the neural basis of metamemory, and also in understanding that metamemory is a process per se, independent of the memory performance (Shimamura & Squire, 1986). However, using only clinical populations, it is impossible to accurately determine the neural basis of metamemory during its normal functioning.

The number of functional magnetic resonance imaging (fMRI) studies on metamemory increased over the past few decades. However, different studies vary greatly in the experimental tasks used to measure metamemory, making it difficult to draw unified conclusions. For example, one can make judgments of metamemory prospectively, that is, estimations of the recallability of an item just presented (judgments of learning) in a later memory test. Kao, Davis, and Gabrieli (2005) presented their participants with scenes and asked them to evaluate if they would recognize those images later. The results point to a distinctive activation of the medial temporal lobe, ventromedial prefrontal cortex, as well as lateral and dorsomedial prefrontal cortex in association with the prediction of future memory. Moreover, one can also make judgments monitoring current performance, that is, subjective evaluations of the correctness of an answer—retrospective confidence—, retrospective confidence is purportedly subserved by the anterior and dorsolateral prefrontal cortex (DLPFC) (Fleming & Dolan, 2012). Hence, while both prospective and retrospective metamemory judgments are metacognitive measures, fMRI findings suggest that the temporality of the judgments of metamemory recruit different brain structures (Fleming & Dolan, 2012; Le Berre et al., 2016; Pannu, Kaszniak, & Rapcsak, 2005).

Within the metamemory judgments the retrospective confidence judgments have been investigated in the various contexts, such as academic performance (Arnold, Higham, & Martín-Luengo, 2013; Mihalca & Mengelkamp, 2020; Mihalca, Mengelkamp, & Schnotz, 2017; Robey, Dougherty, & Buttaccio, 2017), eyewitness testimony and identification (Greenspan & Loftus, 2020; Luna & Martín-Luengo, 2012; Luna, Martín-Luengo, & Brewer, 2015; Semmler, Brewer, & Wells, 2004), as well as various clinical populations (multiple sclerosis: Mazancieux, Souchay, Casez, & Moulin, 2019; psychosis: Dietrichkeit, Grzella, Nagel, & Moritz, 2020; Alzheimer: Bertrand et al., 2019; obsessive compulsive disorder: Moritz & Jaeger, 2018). This suggests that metacognition and particularly retrospective confidence judgments are important in a great deal of tasks in daily life, and thus, this fMRI meta-analysis focuses on the neural underpinnings of retrospective confidence judgments.

To date, there are two fMRI meta-analyses on metacognition in which retrospective confidence judgments have been investigated to some extent. In White, Engen, Sørensen, Overgaard, and Shergill (2014) examined whether the brain activation was similar in what they defined as Type-A research and Type-B research. Type A-research focused on the levels of uncertainty given by the environment (e.g., discriminability of perceptual information), while Type-B research oriented to studies where confidence was measured (subjective assessment of confidence or certainty based on a particular experience). This distinction allows one to determine to which extent the objective and subjective evaluations share their brain substrates. Despite the value of the conceptualization and results of this seminal fMRI meta-analysis, only nine studies were included in the Type-B category. Recent research indicates that such a low number is insufficient for meta-analysis and may biased results due to the over-influence of a single study (Eickhoff, Laird, Fox, Lancaster, & Fox, 2017).

More recently, Vaccaro and Fleming (2018) published an fMRI meta-analysis on metacognition, aiming at distinguishing between the type of metacognitive evaluations (memory vs. decision-making), analysis type (judgment, confidence, and predictors of metacognitive sensitivity), and metamemory temporality (prospective vs. retrospective). The authors identified a cluster for the retrospective confidence judgments, involving activation of the parahippocampal the left inferior frontal gyri (see table 5 and figure 4 in Vaccaro & Fleming, 2018), that is, regions associated with judgments of familiarity in the recognition of novel associations (Du et al., 2019; Haskins, Yonelinas, Quamme, &

Ranganath, 2008). Vaccaro and Fleming also reported the activations between different confidence levels, but for this contrast, all types of metamemory (prospective and retrospective) were collapsed. As mentioned before, the brain activations for different temporal metamemory measures tend to be different (Chua, Schacter, & Sperling, 2009; Fleming & Dolan, 2012; Le Berre et al., 2016; Pannu et al., 2005). Therefore, the results of collapsing different metamemory measures to investigate the differences between high and low evaluations can convey some distortion or lack of clarity.

Thus, in the current meta-analysis, we aimed for the most general representation of the brain activation corresponding to confidence judgment, and included quite heterogeneous studies using different experimental tasks, which lead to the representation of the activations of confidence assessment less influenced by a particular cognitive process (Müller et al., 2018). In other words, the experiments (i.e., contrasts) included in the meta-analysis are all characterized for being estimations of confidence once the main task has been completed.

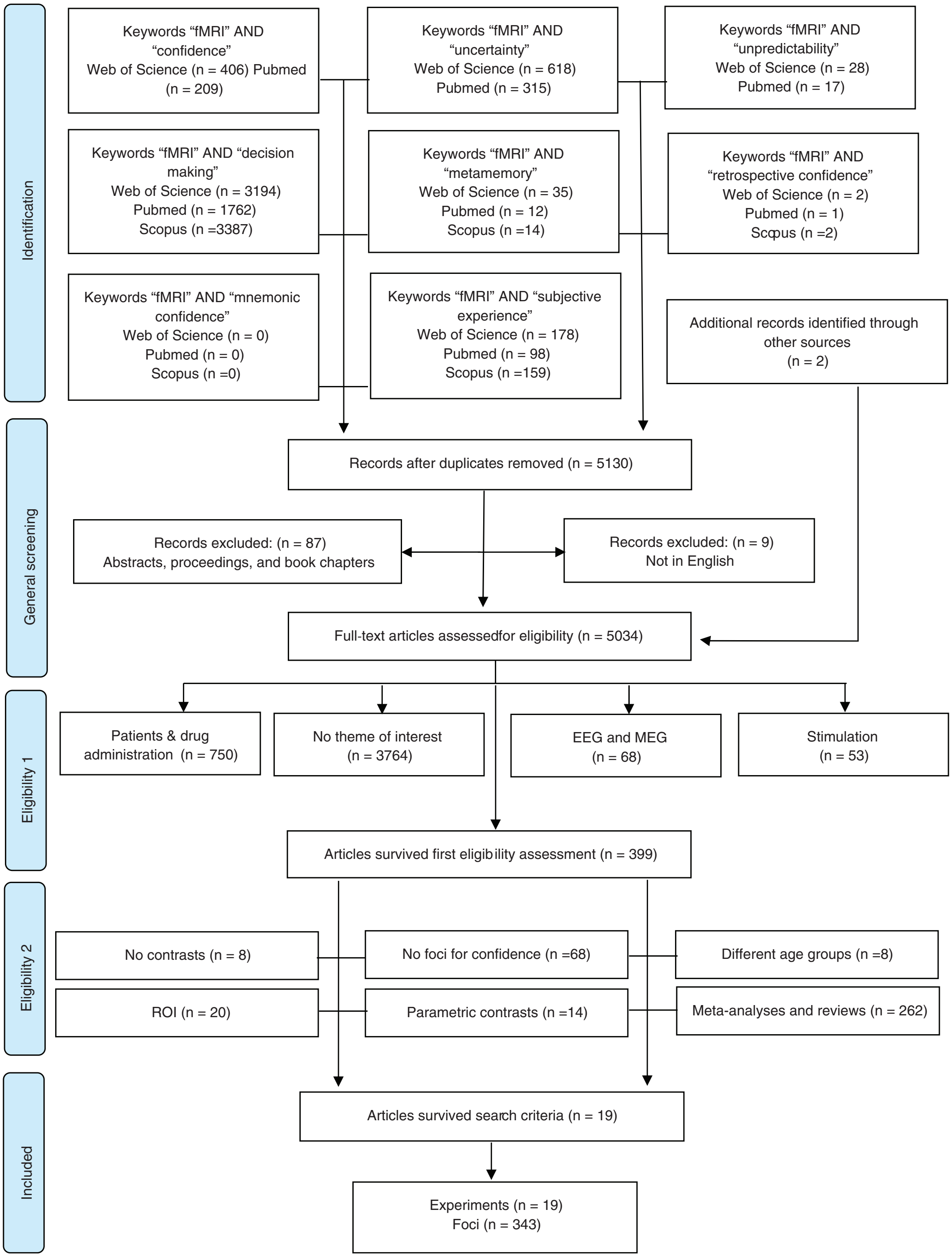

FIGURE 1 PRISMA chart illustrating the selection procedure for the articles used

TABLE 1 Descriptive information on articles used in ALE meta-analysis on retrospective confidence

Short reference	*N*	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
Chua et al., 2004	16	n/a	R	20–33	High confidence correct vs. low confidence correct, Table 3; high confidence	Face-name associative memory paradigm + confidence	Intentional encoding task, in which each subject was instructed to try to remember the name associated with each face for later testing. Subjects were also instructed to indicate if they thought the name "fit" the face. Post scan recognition test: a forced-choice recognition task in which subjects viewed each face seen during encoding paired with the correct name and an incorrect name + their confidence.	8
Chua et al., 2009	32	n/a	n/a	21–29	High confidence > low confidence (in young), Figure 2; low confidence > high confidence, Figure 2	Face-name associative memory paradigm + confidence	Encoding: Participants were instructed to try to remember the name associated with the face for later testing and also to make a purely subjective decision about whether or not the name fit the face. Recognition–confidence blocks: Alternating recognition stimuli and confidence judgment stimuli with four trials of each type; participants viewed each face seen during encoding with three names (correct, two incorrect); high/low confidence	8
Chua et al., 2006	20	n/a	R	20–33	High confidence > low confidence; high confidence > low confidence, for correct answers. Table 2; high confidence > low confidence for incorrect answers, Table 2	Face-name associative memory paradigm + confidence	Encoding: Subjects viewed four novel face–name pairs. Recognition/confidence assessment: Subjects viewed four faces that were encoded in the previous run. In the recognition task (R), subjects chose the correct name among three names. Then, in the confidence assessment task (C), subjects indicated whether they had high or low confidence that they chose the correct name.	24
de Zubicaray, McMahon, Dennis, & Dunn, 2011	16	n/a	R	25.6 ± 7.0	High confidence items > misses, Figure 2	Old vs. new + confidence	Focused attention condition: Words presented in isolation. Divided attention task: a word +2 flanking digits varied in size and quantity (followed by the question about either size or value). Recognition: Old vs. new combined with confidence (e.g., probably old). Five study-test phases were conducted, each separated by a brief retention interval. Instructions as to the nature of each phase were given at the start of the block with "learn" or "remember." Each study phase consisted of 24 words presented one, two, or four times for a total of 56 trials.	3
Fleck, 2005	14	n/a	R	21.0 ± 2,4	Low confidence > high confidence in episodic retrieval, Table 1	Old vs. new + confidence	Prior to scanning, subjects viewed an intermixed list of English words and pseudo-words. For each item, participants performed a lexical decision task. They were also informed that memory for the words would be tested later. During the scan session, participants performed an ER task (recognition memory) and a VP task (area judgment) in form of an "old/new" judgment on the presented stimuli and report of their confidence	11
Hayes et al., 2011	16	14.4 ± 1.7 (years)	n/a	21.6 ± 2.3	High confidence > low confidence, Table 2; low	Old vs. new + confidence	Participants made smaller/bigger than a shoebox or pleasant/unpleasant judgments during encoding, which occurred before the MRI session. During item retrieval, participants made	29

TABLE 1 (Continued)

Short reference	*N*	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
					confidence > high confidence, Table 2		old/new judgments on a 4-point scale (definitely old, probably old, probably new, definitely new). During source retrieval, participants made source judgments on a 4-point scale	
Henson, Rugg, Shallice, & Dolan, 2000	12	n/a	R	21–32	Correct low confidence > correct high confidence	Old vs. new + confidence	Study: Study list, had to indicate whether the word was "pleasant" or "unpleasant"; for semantic encoding of stimuli. Test: Recognition + confidence (high-old, low-old, low-new, and high-new judgments)	2
Kim & Cabeza, 2007	11	n/a	R	18–30	High confidence-TR > low confidence-TR and low confidence-FR > high confidence-FR, Table 3; low confidence-TR > high confidence-TR and high confidence-FR > low confidence -FR, Table 3	DRM paradigm (old vs. new) + confidence	The encoding task was a category judgment task. The retrieval task was an old–new recognition test with confidence ratings that included studied words (true words), nonstudied words from studied categories (false words), and nonstudied, unrelated words (new words)	26
Kim & Cabeza, 2009	12	n/a	R	18–31	High confidence-R activity, Table 2; low confidence-R activity, Table 2	Old vs. new + confidence	Study: Word lists one by one, each consisting of a category name and 4 of the most typical members of the category; the subjects' task was to decide whether all 4 or only 3 instances belonged to the category. Test: Old/new + confidence	9
Kuchinke et al., 2013	20	Students	R	18–45	High-confidence > low-confidence, Table 4	Old vs. new + confidence	Study: Study list. Test: Participants gave memory judgments on a 6-point-rating scale(from (1) "sure new" to "unsure" to (6) "sure old")	18
Leiker & Johnson, 2015	16	Students	R	19 ± 14	High confidence > moderate/low confidence, Table 2; moderate/low confidence > high confidence, Table 2	Old vs. new; source-memory + confidence	Encoding: Presentation of a series of words that subjects were to think about and rate in the context of one of three tasks (the artist task - imagine how an artist would draw the object denoted by the word and rate the difficulty of the drawing; the function task - to think of as many different functions as they could for the object and respond with the number of functions generated; the cost task - to think about and rate the relative cost of the object). Retrieval: Old/new words; source-memory (artist, function, cost, new) + confidence (high, moderate, low)	22
Molenberghs et al., 2016	308	n/a	283 R	41	More confident > less confident, Table 1; less confident > more confident, Table 1	EmpaToM: Empathy, compassion, theory of mind (ToM), and confidence	EmpaToM trial sequence. Following a 2 (emotionality of the video) × 2 (ToM requirements) design, four different video types were presented for each actor: Emotionally negative and neutral videos; videos with and without ToM demands, thereby leading to ToM vs. factual reasoning questions. After each video, participants rated their own affect and their compassion for the person in the video. Subsequently, they answered a ToM or non-ToM (i.e., factual reasoning) question about the video. After each question, participants rated their confidence regarding their performance in the question	25

(Continues)

TABLE 1 (Continued)

Short reference	*N*	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
Moritz, Gläscher et al., 2006; Moritz, Woodward, & Chen, 2006	17	n/a	R	27.41 ± 7.51	High-confident > low-confident, Table 1; high-confidence hits > misses; low-confident > high confident, Table 1	DRM paradigm (old vs. new) + confidence	Encoding: Words from the six lists were visually presented, each list contained 12 stimuli that were presented in descending semantic relatedness to the list theme; participants were asked to indicate whether each item was a noun or not. Recognition: Old/new + confidence	20
Mugikura et al., 2016	24	n/a	R	20–25	Destination + item hits with high confidence > destination + item hits with low confidence, Table 3	Old vs. new, context judgment and confidence	Study: The subjects told a series of facts to either a woman or a man. Test: During fMRI scanning, the subjects were asked to judge whether each presented fact was new or old, and if they judged it as old, to indicate, including a confidence rating (high or low), whether the subjects had told that fact to either a man or a woman during the study phase (definitely told to the woman, probably told to the woman, definitely told to the man, probably told to the man, and not presented [new])	20
Pais-Vieira et al., 2015	17	n/a	R	22.7 ± 2.5	High-confident > low-confident, Table 1; high-confidence hits > misses; low-confident > high confident, Table 1	Old vs. new + confidence	Encoding: Scanned; incidental encoding; negative, positive and neutral pictures; three distinct conditions: ISO (how much they were personally moved by each picture), EPO (to rate the visual brightness of picture) and semantic orienting (not reported). Recognition: 2–3 days later; outside; old/new + confidence	9
Risius et al., 2013	29	n/a	R	25 ± 3.2	High confidence > low confidence, Table 2; low confidence > high confidence, Table 2	Memory retrieval and confidence	Encoding: Outside; short emotional film. Retrieval: Scanned; correct or incorrect statement regarding events in the movie, to which subjects had to respond either "yes" (true) or "no"(false). Monitoring: Participants had to rate their confidence about the foregoing response (high vs. low secureness) and had to decide either to volunteer or to withhold their answer (CONTROL phase)	85
Woroch et al., 2019	14	n/a	R	20–35	High confidence for places > low confidence for places, Table 2; high confidence for faces > low confidence for faces, Table 2	Memory recognition and confidence	A recognition memory experiment with stimuli comprising of words, pictures of faces, and pictures of "places." Study: Participants studied words randomly presented one at a time beneath a picture of either a face or a scene test: Whether the word was previously paired with a face or a place image and simultaneously confidence on a scale from one to six	3
Yonelinas, 2005	16	n/a	R	19–33	Confidence increasing with familiarity (1–4), Table 3; confidence decreasing with familiarity (4–1), Table 3	Old vs. new and confidence	Study: Subjects saw a series of critical words, for each word, they had to decide whether the word denotes an abstract or concrete entity. Recognition memory: Critical words intermixed with new words. If they were able to remember something specific about seeing the word at study, subjects were asked to give a remember (R) judgment. If they could not recollect anything specific about experiencing the item, subjects were asked to rate their memory confidence in order of high to low confidence	20

TABLE 1 (Continued)

Short reference	N	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
Yu, Johnson, & Rugg, 2012	16	n/a	R	18–23	Remember-high (confidence) > remember > weak (confidence), Figure 2a	Remember/know, source memory judgments and confidence	Study phase: Outside the scanner, picture of objects. Test: Critical items + new pictures; for each item the requirement was to first make a "remember/know/new" judgment. Instructions were to use the "Remember" ("R") response if recognition was accompanied by retrieval of any reportable detail about its study presentation, to use the "now" ("K") response for items judged to be studied in the absence of the retrieval of study details, and to respond "new" ("N") to items judged to be unstudied or for which study status was uncertain. For any item accorded an "R" or "K" response, a new cue appeared to signal the requirement to judge whether the item had been studied on the left or right side. The source judgment was made on a 3-point confidence scale (high, moderate, low).	1

Abbreviation: ALE, activation likelihood estimation.

2 | METHODS

2.1 | Literature search and article selection

The literature search was performed using Web of Science (http://apps.webofknowledge.com/), Scopus (https://www.scopus.com/home.uri), and PubMed (https://pubmed.ncbi.nlm.nih.gov) search engines on 11.04.2020, with keywords ("confidence" AND "fMRI," "uncertainty" AND "fMRI," "unpredictability" AND "fMRI," "decision-making" AND "fMRI," "metamemory" AND "fMRI," "retrospective confidence" AND "fMRI," "mnemonic confidence" AND "fMRI," "subjective experience" AND "fMRI"), during years 1998–2020, yielding a total of 11,217 articles (see Figure 1 for PRISMA chart on step-by-step analysis procedure). To achieve homogeneity of imaging data, we did not search for articles that reported PET data. Next, we excluded abstracts and non-English text (n = 40), yielding 5,032 full-text articles eligible for further screening. Further screening was conducted in two steps. First, we excluded articles with irrelevant topics (i.e., decision-making, visual search, etc.), articles that reported data on clinical patients or with drug administration, articles that used other neuroimaging techniques (EEG and MEG), and articles that involved stimulation. This initial screening allowed us to assess the articles that were relevant to our topic more efficiently and thoroughly. During the second screening, we excluded meta-analysis and reviews, articles that did not report fMRI coordinates or measured confidence outside of fMRI, articles that did not report whole brain analysis (only ROI), reported parametric contrasts, or compared different age groups. The list with the manuscripts and reasons for exclusion from this second screening can be consulted in Table 7 of the supplemental materials. The eligibility of the article inclusion based on topic of interest was accessed based on the following criteria: (a) subjects performed tasks in the context of retrospective confidence evaluation and meta-memory; (b) statistical models for contrasts of "high>low" or "low>high" confidence were provided. Articles meeting these criteria were reviewed by two researchers of the team (O. Z. and A. D.). We included into the analysis only articles which reported data on healthy controls (i.e., participants were free from psychiatric or neurological disorders and neuropharmacological influence). Next, we did a reference search on the articles included into analysis, yielding two additional articles (see Table 1 for the detailed information on these articles). We included all relevant experiments from each article in the analysis because the activation likelihood estimation (ALE) algorithm uses a correction to avoid summation of within-group effects (Turkeltaub et al., 2012). Unfortunately, due to the low number of experiments that included the contrasts of high > low confidence (n = 15) and low > high confidence (n = 11), the result for these contrasts needs to be cautiously considered by other researchers in the future (see Eickhoff et al., 2017 for the recommended number of experiments).

2.2 | Activation likelihood estimate

The meta-analysis was performed using GingerALE software, version 3.0.2 (freely available at http://www.brainmap.org/ale/) which allows to perform coordinate-based ALE (Eickhoff et al., 2017; Eickhoff, Bzdok, Laird, Kurth, & Fox, 2012). Foci extracted from different experiments across different articles were used to create a probabilistic map comparing the likelihood of activation to random spatial distribution. All coordinates reported in MNI space were converted to Talairach space before the analysis. Statistical significance was assessed using a cluster-level threshold for multiple comparisons at $p = .05$ with a cluster-forming threshold set to $p = .001$ (Eickhoff et al., 2012, 2017).

In the majority of the studies included in the current meta-analysis, the participants had to indicate the confidence in an old/new memory answer ($n = 11$, de Zubicaray et al, 2019; Fleck, 2005; Hayes et al., 2011; Henson et al., 2000; Kim & Cabeza, 2009; Kuchinke et al., 2013; Leiker & Johnson, 2015; Mugikura et al., 2016; Pais-Vieira et al., 2015; Risius et al., 2013; Woroch et al., 2019; Yonelinas, 2005), or to report confidence in their answer to other tasks ($n = 18$, Chua et al., 2004; Chua et al., 2009; Chua et al., 2006; de Zubicaray et al, 2019; Fleck, 2005; Hayes et al., 2011; Henson et al., 2000; Kim & Cabeza, 2007; Kim & Cabeza, 2009; Kuchinke et al., 2013; Leiker & Johnson, 2015; Molenberghs et al., 2016; Moritz, Gläscher, et al., 2006; Moritz, Woodward, & Chen, 2006; Pais-Vieira et al., 2015; Risius et al., 2013; Woroch et al., 2019; Yonelinas, 2005; Yu et al., 2012). Recent meta-analyses (Vaccaro & Fleming, 2018; White et al., 2014) did not include direct contrasts measured confidence (i.e., feeling of knowing, FOK, or JOL, judgment of learning) or nonparametric contrasts (i.e., "significant correlations with subjective certainty in color trials" in Vaccaro & Fleming, 2018). To keep the methodology consistent, we included only experiments that used subtraction contrasts (i.e., high confidence > low confidence).

3 | RESULTS

3.1 | Confidence

Articles included in the main analysis reported data on 626 participants. Across articles, which reported gender, 276 were male participants. Across articles, which reported handedness, 577 were right-handed. For the articles where age range was given, the median of age was 24.2 years (see Table 1 for detailed information). Figure 1 shows the number of articles, number of experiments, and foci included in the meta-analysis.

3.2 | ALE maps for all tasks related to confidence

All tasks related to confidence evaluation show concordance in four clusters (Table 2 and Figure 2). The largest cluster with the highest ALE value is found in the left parahippocampal gyrus (BA 35 and 36). The second cluster is found in the left middle frontal gyrus (BA 46). Subcortical activation is observed in the next two clusters: in the right amygdala and right parahippocampal gyrus (BA 28, 34).

3.3 | Exploratory analysis: High > low confidence and low < high confidence

As an exploratory analysis, we estimated the two conditions included in the meta-analysis: high confidence and low confidence. Previous studies suggest that these two types of confidence are underpinned by different neural substrates, such as parahippocampal gyrus or DLPFC (Kim & Cabeza, 2009). Therefore, we selected 15 studies (foci = 174) out of the main sample, where participants reported high confidence at the retrieval/recognition stage, and 11 studies (foci = 158)—reporting low confidence.

3.3.1 | ALE maps for the contrast high > low

Three subcortical clusters were identified centered at the right parahippocampal gyrus (BA 28, and amygdala), left parahippocampal gyrus (BA 36), and right caudate. See Figure 3 and Table 3 for the ALE results for the high > low contrast. See Table 4 for detailed information of studies included.

3.3.2 | ALE maps for the contrast low > high

The concordant activation for low confidence versus high confidence was found in three significant clusters centered at the following brain

TABLE 2 ALE meta-analysis of all studies on retrospective confidence

	Foci (x, y, z)					
Cluster	x	y	z	ALE value	Brain area	BA
1	−22	−22	−16	0.024563098	Left parahippocampal gyrus	35
1	−28	−32	−12	0.01996939	Left parahippocampal gyrus	36
2	−40	18	22	0.025068823	Left middle frontal gyrus	46
3	24	−24	−8	0.023705766	Right parahippocampal gyrus	28
4	24	−8	−14	0.017792294	Right amygdala	
4	16	−2	−12	0.017452436	Right parahippocampal gyrus	28
4	16	0	−16	0.017243283	Right parahippocampal gyrus	34

Abbreviation: ALE, activation likelihood estimation.

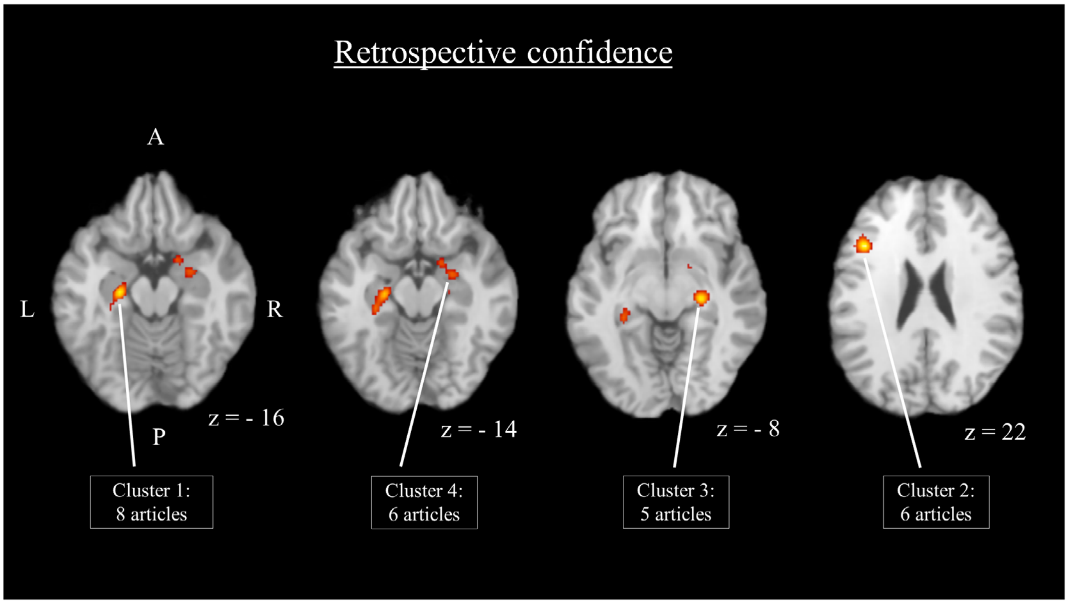

FIGURE 2 Brain maps showing significant activation likelihood estimation (ALE) values for confidence

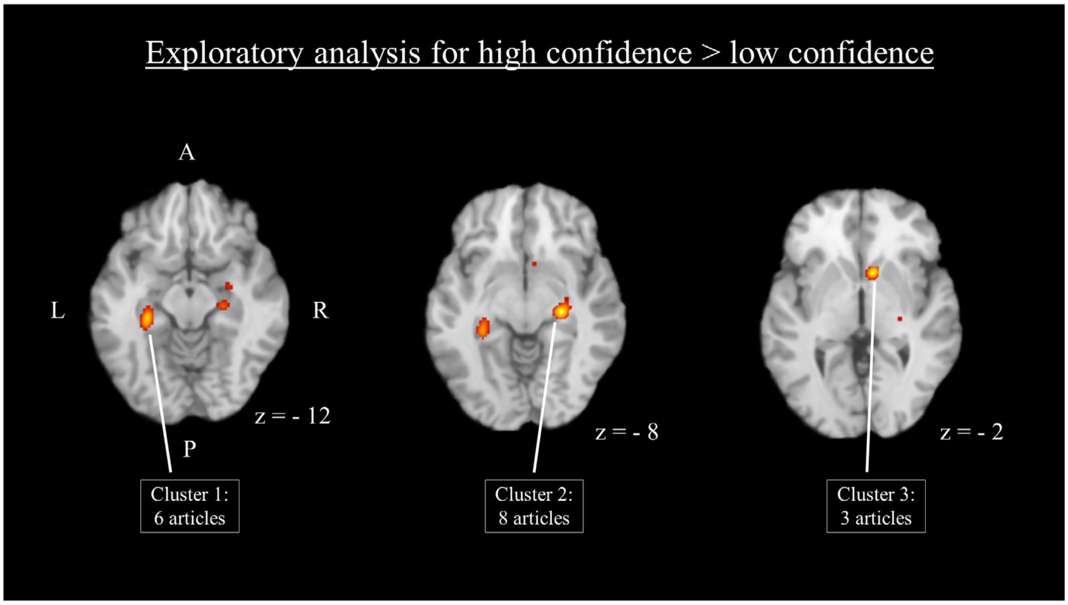

FIGURE 3 Activation likelihood estimation (ALE) results for the high > low contrast

TABLE 3 Supplementary ALE meta-analysis for high > low confidence contrasts (15 papers)

	Foci (x, y, z)					
Cluster	x	y	z	ALE value	Brain area	BA
1	−28	−32	−12	0.019941954	Left parahippocampal gyrus	36
2	26	−24	−8	0.022155292	Right parahippocampal gyrus	28
2	28	−8	−16	0.01427889	Right amygdala	
3	8	8	−2	0.02149668	Right caudate head	

Abbreviation: ALE, activation likelihood estimation.

TABLE 4 Descriptive information on articles used in exploratory analysis on high confidence > low confidence

Short reference	*N*	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
Chua et al., 2004	16	n/a	R	20–33	High confidence correct vs. low confidence correct, Table 3; high confidence incorrect vs. low confidence incorrect, Table 3	Face-name associative memory paradigm + confidence	Intentional encoding task, in which each subject was instructed to try to remember the name associated with each face for later testing. Subjects were also instructed to indicate if they thought the name "fit" the face. Post scan recognition test: a forced-choice recognition task in which subjects viewed each face seen during encoding paired with the correct name and an incorrect name + their confidence.	8
Chua et al., 2009	32	n/a	n/a	21–29	High confidence > low confidence (in young), Figure 2	Face-name associative memory paradigm + confidence	Encoding: Participants were instructed to try to remember the name associated with the face for later testing and also to make a purely subjective decision about whether or not the name fit the face. Recognition–confidence blocks: Alternating recognition stimuli and confidence judgment stimuli with four trials of each type; participants viewed each face seen during encoding with three names (correct, two incorrect); high/low confidence	2
Chua, Schacter, Rand-Giovannetti, & Sperling, 2006	20	n/a	R	20–33	High confidence > low confidence; high confidence > low confidence, for correct answers, Table 2; high confidence > low confidence for incorrect answers, Table 2	Face-name associative memory paradigm + confidence	Encoding: Subjects viewed four novel face–name pairs. Recognition/ confidence assessment: Subjects viewed four faces that were encoded in the previous run. In the recognition task (R), subjects chose the correct name among three names. Then, in the confidence assessment task (C), subjects indicated whether they had high or low confidence that they chose the correct name	24
Hayes, Buchler, Stokes, Kragel, & Cabeza, 2011	16	14.4 ± 1.7 (years)	n/a	21.6 ± 2.3	High confidence > low confidence, Table 2	Old vs. new + confidence	Participants made smaller/bigger than a shoebox or pleasant/ unpleasant judgments during encoding, which occurred before the MRI session. During item retrieval, participants made old/new judgments on a 4-point scale (definitely old, probably old, probably new, definitely new). During source retrieval, participants made source judgments on a 4-point scale	13
Kim & Cabeza, 2007	11	n/a	R	18–30	High confidence-TR > low confidence-TR, Table 3; high confidence-FR > low confidence-FR, Table 3	DRM paradigm (old vs. new) + confidence	The encoding task was a category judgment task. The retrieval task was an old–new recognition test with confidence ratings that included studied words (true words), nonstudied words from studied categories (false words), and nonstudied, unrelated words (new words)	16
Kim & Cabeza, 2009	12	n/a	R	18–31	High confidence-R activity, Table 2	Old vs. new + confidence	Study: Word lists one by one, each consisting of a category name and 4 of the most typical members of the category; the subjects' task was to decide whether all 4 or only 3 instances belonged to the category. Test: Old/new + confidence	5
Kuchinke, Fritzemeier, Hofmann, & Jacobs, 2013	20	Students	R	18–45	High-confidence > low-confidence, Table 4	Old vs. new + confidence	Study: Study list. Test: Participants gave memory judgments on a 6-point-rating scale (from (1) "sure new" to "unsure" to (6) "sure old")	18

TABLE 4 (Continued)

Short reference	N	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
Leiker & Johnson, 2015	16	Students	R	19 ± 1.4	High confidence > moderate/low confidence, Table 2	Old vs. new; source-memory + confidence	Encoding: Presentation of a series of words that subjects were to think about and rate in the context of one of three tasks (the artist task—imagine how an artist would draw the object denoted by the word and rate the difficulty of the drawing; the function task—to think of as many different functions as they could for the object and respond with the number of functions generated; the cost task—to think about and rate the relative cost of the object). Retrieval: Old/new words; source-memory (artist, function, cost, new) + confidence (high, moderate, low)	21
Molenberghs, Trautwein, Böckler, Singer, & Kanske, 2016	308	n/a	283 R	41	More confident > less confident, Table 1	EmpaToM: Empathy, compassion, ToM, and confidence	EmpaToM trial sequence. Following a 2 (emotionality of the video) × 2 (ToM requirements) design, four different video types were presented for each actor: Emotionally negative and neutral videos; videos with and without ToM demands, thereby leading to ToM vs. factual reasoning questions. After each video, participants rated their own affect and their compassion for the person in the video. Subsequently, they answered a ToM or non-ToM (i.e., factual reasoning) question about the video. After each question, participants rated their confidence regarding their performance in the question	10
Moritz, Gläscher et al., 2006; Moritz, Woodward, & Chen, 2006	17	n/a	R	27.41 ± 7.51	High-confident > low-confident, Table 1; high-confidence hits > misses	DRM paradigm (old vs. new) + confidence	Encoding: Words from the six lists were visually presented, each list contained 12 stimuli that were presented in descending semantic relatedness to the list theme; participants were asked to indicate whether each item was a noun or not. Recognition: Old/new + confidence	15
Mugikura et al., 2016	24	n/a	R	20–25	Destination + item hits with high confidence > destination + item hits with low confidence, Table 3	Old vs. new, context judgment and confidence	Study: The subjects told a series of facts to either a woman or a man. Test: During fMRI scanning, the subjects were asked to judge whether each presented fact was new or old, and if they judged it as old, to indicate, including a confidence rating (high or low), whether the subjects had told that fact to either a man or a woman during the study phase (definitely told to the woman, probably told to the woman, definitely told to the man, probably told to the man, and not presented [new])	20
Pais-Vieira, Wing, & Cabeza, 2015	17	n/a	R	22.7 ± 2.5	High-confident > low-confident, Table 1; high-confidence hits > misses	Old vs. new + confidence	Encoding: Scanned; incidental encoding; negative, positive and neutral pictures; three distinct conditions: ISO (how much they were personally moved by each picture), EPO (to rate the visual brightness of picture) and semantic orienting (not reported). Recognition: 2–3 days later; outside; old/new + confidence	6
Risius et al., 2013	29	n/a	R	25 ± 3.2	High confidence > low confidence, Table 2	Memory retrieval and confidence	Encoding: Outside; short emotional film. Retrieval: Scanned; correct or incorrect statement regarding events in the movie, to which subjects had to respond either "yes" (true) or "no" (false). Monitoring: Participants had to rate their confidence about the foregoing response (high vs. low secureness) and had to decide either to volunteer or to withhold their answer (CONTROL phase)	4

(Continues)

TABLE 4 (Continued)

Short reference	N	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
Woroch, Konkel, & Gonsalves, 2019	14	n/a	R	20–35	High confidence for places > low confidence for places, Table 2; high confidence for faces > low confidence for faces, Table 2	Memory recognition and confidence	A recognition memory experiment with stimuli comprising of words, pictures of faces, and pictures of "places." Study: Participants studied words randomly presented one at a time beneath a picture of either a face or a scene test: Whether the word was previously paired with a face or a place image and simultaneously confidence on a scale from one to six	3
Yonelinas, 2005	16	n/a	R	19–33	Confidence increasing with familiarity (1–4), Table 3; confidence decreasing with familiarity (4–1), Table 3	Old vs. new and confidence	Study: Subjects saw a series of critical words, for each word, they had to decide whether the word denotes an abstract or concrete entity. Recognition memory: Critical words intermixed with new words. If they were able to remember something specific about seeing the word at study, subjects were asked to give a remember (R) judgment. If they could not recollect anything specific about experiencing the item, subjects were asked to rate their memory confidence in order of high to low confidence	9

Abbreviation: ToM, theory of mind.

regions: left middle frontal gyrus (BA 46), right precuneus (BA7), and left inferior frontal gyrus (BA 47). See Figure 4 and Table 5 for the ALE results for the low > high contrast. See Table 6 for detailed information of the studies included.

4 | DISCUSSION

Metamemory is the online process, for which we monitor the ongoing activity of memory processes and implement strategies to control its performance depending on our objectives (Nelson & Narens, 1990). This high cognitive process is important for many different processes in both healthy and clinical populations. It is particularly vital in the clinical population, where awareness of the memory may be a prerequisite for the success of the interventions to improve memory (Clare & Woods, 2004). The variety of different measures in assessing the level of metamemory complicates the identification of the underlying brain structures, which makes it crucial to run specific meta-analyses for each main category (i.e., prospective vs. retrospective) in which metamemory measures are classified (Chua et al., 2009; Fleming & Dolan, 2012; Le Berre et al., 2016; Pannu & Kaszniak, 2005).

In this study, we examined concordance across studies in brain areas responding to the assessment of confidence. We also ran the exploratory analysis for the directionality of confidence (high > low, low > high) in order to observe the tendency of the data, which can orient future studies. The main analysis of confidence revealed concordant bilateral activation in the parahippocampal gyrus, left middle frontal gyrus, and right amygdala. Notice that the concordant activation we observed in our analysis closely corresponds to the concordant activation of the parahippocampal gyrus reported by Vaccaro and Fleming (2018) for metamemory judgments and, specifically, retrospective metamemory judgments. However, the rest of our results differ from theirs since we only found this concordant bilateral activation in the parahippocampal gyrus and activation of the left inferior frontal gyrus. This difference most probably stems from our criteria of inclusion—we only included whole-brain fMRI articles, and only those reporting subtraction contrasts. These criteria guarantee that the analyses only reveal regions associated with retrospective confidence and no other general functions involved (Müller et al., 2018).

Overall, our main results confirmed the involvement of the parahippocampal gyrus and the left inferior frontal gyrus in the assessment of confidence (Fleming & Dolan, 2012; Metcalfe & Schwartz, 2015; Vaccaro & Fleming, 2018). We also confirmed the role of the medial temporal lobe in the processes of recollection of information during or prior to the confidence assessment (Cohen & Eichenbaum, 1993; Eichenbaum & Cohen, 2001). The act of retrieving information is undoubtedly required in the assessment of the confidence (Nelson, 1990); otherwise, we would not be able to evaluate the certainty of our memories. MTL has been extensively related in the recollection for declarative memories (i.e., memories for knowledge and events) and retrieval processes (Eichenbaum, Yonelinas, & Ranganath, 2007). It should be noted that the activation of structures

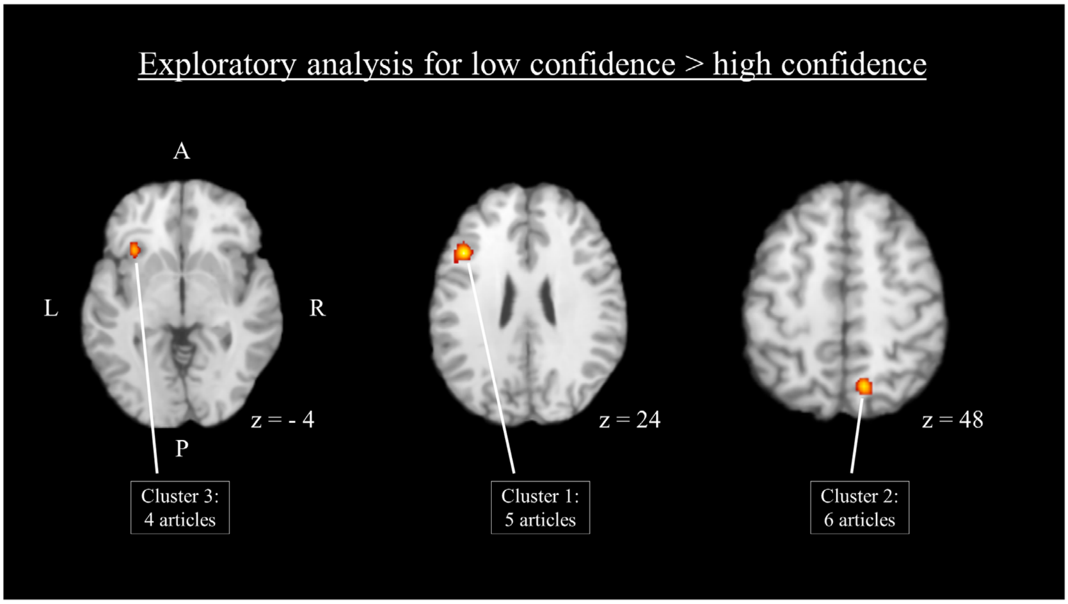

FIGURE 4 Activation likelihood estimation (ALE) results for the low > high contrast

TABLE 5 Supplementary ALE meta-analysis for low > high confidence contrasts (11 papers)

Cluster	Foci (x, y, z) x	y	z	ALE value	Brain area	BA
1	−42	18	24	0.021589866	Left middle frontal gyrus	46
2	12	−66	48	0.020475976	Right precuneus	7
3	−32	18	−4	0.02492742	Left inferior frontal gyrus	47

Abbreviation: ALE, activation likelihood estimation.

included in the MTL only appears in the main analyses and the high > low contrast. The absence of activation of the MTL for the contrast low > high may be explained by the fact that low confidence judgments are related to the inability to retrieve the requested information. This emphasizes the involvement of MTL in retrieval and evaluation of the confidence.

In both, the main and the additional low > high confidence analyses, we found an activation of the left middle frontal gyrus (BA 46). This area is presumably in charge of the response inhibition and error monitoring in go/no-go tasks (Heitzeg et al., 2014; Le et al., 2020), as well as in multiple different cognitive tasks spanning from working memory (Leung, Gore, & Goldman-Rakic, 2002) to motor executive control (Kübler, Dixon, & Garavan, 2006). Hence, the observation of the BA 46 activation for retrospective confidence in the main and in the low > high confidence analyses but not in the high > low analysis, suggests its involvement in the negative outcome of monitoring, leading to low confidence assessment of confidence. The activation BA46 in these divergent tasks where monitoring is required, suggests that it is the most prominent area related to monitoring.

Our findings for the low > high contrast are also consistent with previous work that has associated the activation of the right precuneus with low confidence in memory (Ye, Zou, Lau, Hu, & Kwok, 2018). Ye and colleagues disrupted the precuneus by the use of transcranial magnetic stimulation to examine its causal role in the meta-cognition of perceptual and memory processes. They found differences in the assessment of confidence, particularly, the impairment to adequately rate low confident answers with low confidence on the memory domain. This result not only corroborates the hypothesis of the involvement of the precuneus in metamemory processes (see work of Simons, Peers, Mazuz, Berryhill, & Olson, 2010 on patients with bilateral parietal lesions), but also strengthens the view of a domain-specificity in the assessment of metacognition (Morales, Lau, & Fleming, 2018).

5 | ANALYSIS LIMITATIONS

Our study focused on the brain structures involved in the evaluation of retrospective confidence. Ideally, we should have run the analyses balancing the perceptual modality of the tasks. First, we included studies that used diverse types of materials (images or semantic information). Such a variability in underlying cognitive processes may introduce certain bias in the results, since there is an open debate on whether metacognition is a domain-specific or domain-general process, and whether they possibly co-exist (Morales et al., 2018). Second, as repeatedly mentioned above, the results for high > low and

TABLE 6 Descriptive information on articles used in exploratory analysis on low confidence > high confidence

Short reference	*N*	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
Chua et al., 2009	32	n/a	n/a	21–29	Low confidence > high confidence, Figure 2	Face-name associative memory paradigm + confidence	Encoding: Participants were instructed to try to remember the name associated with the face for later testing and also to make a purely subjective decision about whether or not the name fit the face. Recognition–confidence blocks: Alternating recognition stimuli and confidence judgment stimuli with four trials of each type; participants viewed each face seen during encoding with three names (correct, two incorrect); high/low confidence	6
Fleck, 2005	14	n/a	R	21.0 ± 2,4	Low confidence > high confidence in episodic retrieval, Table 1	Old vs. new + confidence	Prior to scanning, subjects viewed an intermixed list of English words and pseudo-words. For each item, participants performed a lexical decision task. They were also informed that memory for the words would be tested later. During the scan session, participants performed an ER task (recognition memory) and a VP task (area judgment) in form of an "old/new" judgment on the presented stimuli and report of their confidence	11
Hayes et al., 2011	16	14.4 ± 1.7 (years)	n/a	21.6 ± 2.3	Low confidence > high confidence, Table 2	Old vs. new + confidence	Participants made smaller/bigger than a shoebox or pleasant/ unpleasant judgments during encoding, which occurred before the MRI session. During item retrieval, participants made old/new judgments on a 4-point scale (definitely old, probably old, probably new, definitely new). During source retrieval, participants made source judgments on a 4-point scale	16
Henson et al., 2000	12	n/a	R	21–32	Correct low confidence > correct high confidence	Old vs. new + confidence	Study: Study list, had to indicate whether the word was "pleasant" or "unpleasant"; for semantic encoding of stimuli. Test: Recognition + confidence (high-old, low-old, low-new, and high-new judgments)	2
Kim & Cabeza, 2007	11	n/a	R	18–30	Low confidence-FR > high confidence-FR, Table 3; low confidence-TR > high confidence-TR, Table 3	DRM paradigm (old vs. new) + confidence	The encoding task was a category judgment task. The retrieval task was an old–new recognition test with confidence ratings that included studied words (true words), nonstudied words from studied categories (false words), and nonstudied, unrelated words (new words)	10
Kim & Cabeza, 2009	12	n/a	R	18–31	Low confidence-R activity, Table 2	Old vs. new + confidence	Study: Word lists one by one, each consisting of a category name and 4 of the most typical members of the category; the subjects' task was to decide whether all 4 or only 3 instances belonged to the category. Test: Old/new + confidence	4
Leiker & Johnson, 2015	16	Students	R	19 ± 14	Moderate/low confidence > high confidence, Table 2	Old vs. new; source-memory + confidence	Encoding: Presentation of a series of words that subjects were to think about and rate in the context of one of three tasks (the artist task - imagine how an artist would draw the object denoted by the word and rate the difficulty of the drawing; the function task - to think of as many different functions as they could for the object and respond with the number of functions generated; the cost task - to think about and rate the relative cost of the object). Retrieval: Old/new words; source-memory (artist, function, cost, new) + confidence (high, moderate, low)	1

TABLE 6 (Continued)

Short reference	*N*	Education	Hand	Age	Contrast	Task (short)	Task (description)	Foci
Molenberghs et al., 2016	308	n/a	283 R	41	Less confident > more confident, Table 1	EmpaToM: Empathy, compassion, ToM, and confidence	EmpaToM trial sequence. Following a 2 (emotionality of the video) × 2 (ToM requirements) design, four different video types were presented for each actor: Emotionally negative and neutral videos; videos with and without ToM demands, thereby leading to ToM vs. factual reasoning questions. After each video, participants rated their own affect and their compassion for the person in the video. Subsequently, they answered a ToM or non-ToM (i.e., factual reasoning) question about the video. After each question, participants rated their confidence regarding their performance in the question	15
Moritz, Gläscher et al., 2006; Moritz, Woodward, & Chen, 2006	17	n/a	R	27.41 ± 7.51	Low-confident > high confident, Table 1	DRM paradigm (old vs. new) + confidence	Encoding: Words from the six lists were visually presented, each list contained 12 stimuli that were presented in descending semantic relatedness to the list theme; participants were asked to indicate whether each item was a noun or not. Recognition: Old/new + confidence	1
Risius et al., 2013	29	n/a	R	25 ± 3.2	Low confidence > high confidence, Table 2	Memory retrieval and confidence	Encoding: Outside; short emotional film. Retrieval: Scanned; correct or incorrect statement regarding events in the movie, to which subjects had to respond either "yes" (true) or "no" (false). Monitoring: Participants had to rate their confidence about the foregoing response (high vs. low secureness) and had to decide either to volunteer or to withhold their answer (CONTROL phase)	81
Yonelinas, 2005	16	n/a	R	19–33	Confidence decreasing with familiarity (4–1), Table 3	Old vs. new and confidence	Study: Subjects saw a series of critical words, for each word, they had to decide whether the word denotes an abstract or concrete entity. Recognition memory: Critical words intermixed with new words. If they were able to remember something specific about seeing the word at study, subjects were asked to give a remember (R) judgment. If they could not recollect anything specific about experiencing the item, subjects were asked to rate their memory confidence in order of high to low confidence	11

Abbreviation: ToM, theory of mind.

low > high contrasts were preliminary and should be considered with caution, because none of them reach the recommended minimum number of studies to guarantee sufficient power (17 studies; see Eickhoff et al., 2012; Müller et al., 2018). Despite these shortcomings, the current meta-analysis presents new knowledge on the topic of confidence judgments, which is more advantageous compared to standard reviews.

6 | CONCLUSION

We aimed to further clarify the particular brain structures sub serving metamemory, particularly on retrospective confidence. Our results indicated the involvement of four clusters required for the assessment of confidence (see Table 2). Some of the regions included in these clusters have also been related to memory processes as well as to monitoring of errors and inhibition of answers, but all of them are consistent with the proposed functioning of metamemory processes (Fleming & Frith, 2014). Finally, our results can better inform future clinical research focusing on retrospective confidence failures and emphasize the need for similar studies with other types of metamemory measures.

ACKNOWLEDGMENT

This article was fully funded by the Russian Foundation grant number RSF 19 18 00534.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All the files and information used for analysis are available upon request.

ORCID

Beatriz Martín-Luengo https://orcid.org/0000-0003-3642-5337

REFERENCES

Arnold, M. M., Higham, P. A., & Martín-Luengo, B. (2013). A little bias goes a long way: The effects of feedback on the strategic regulation of accuracy on formula-scored tests. *Journal of Experimental Psychology: Applied*, *19*(4), 383–402. https://doi.org/10.1037/a0034833

Bertrand, J. M., Mazancieux, A., Moulin, C. J., Béjot, Y., Rouaud, O., & Souchay, C. (2019). In the here and now: Short term memory predictions are preserved in Alzheimer's disease. *Cortex*, *119*, 158–164. https://doi.org/10.1016/j.cortex.2019.03.027

Chua, C., Rand-Giovannetti, E., Schacter, D. L., Albert, M. S., Chua, E. F., & Sperling, R. A. (2004). Dissociating confidence and accuracy: Functional magnetic resonance imaging shows origins of the subjective memory experience. *Journal of Cognitive Neuroscience*, *16*(7), 1131–1142. https://doi.org/10.1162/0898929041920568

Chua, E. F., Schacter, D. L., Rand- Giovannetti, E., & Sperling, R. A. (2006). Understanding metamemory: Neural correlates of the cognitive process and subjective level of confidence in recognition memory. *NeuroImage*, *29*(4), 1150–1160. https://doi.org/10.1016/j.neuroimage.2005.09.058

Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009). Neural basis for recognition confidence in younger and older adults. *Psychology and Aging*, *24*(1), 139–153. https://doi.org/10.1037/a0014029

Clare, L., & Woods, R. T. (2004). Cognitive training and cognitive rehabilitation for people with early-stage Alzheimer's disease: A review. *Neuropsychologycal Rehabilitation*, *14*(4), 385–401. https://doi.org/10.1080/09602010443000074

Cohen, N. J., & Eichenbaum, H. B. (1993). *Memory, amnesia, and hippocampal function*. Cambridge: MIT Press.

de Zubicaray, G. I., McMahon, K. L., Dennis, S., & Dunn, J. C. (2011). Memory strength effects in fMRI studies: A matter of confidence. *Journal of Cognitive Neuroscience*, *23*(9), 2324–2335. https://doi.org/10.1007/s11682-016-9545-2

Dietrichkeit, M., Grzella, K., Nagel, M., & Moritz, S. (2020). Using virtual reality to explore differences in memory biases and cognitive insight in people with psychosis and healthy controls. *Psychiatry Research*, *285*, 112787. https://doi.org/10.1016/j.psychres.2020.112787

Du, X., Zhan, L., Chen, G., Guo, D., Li, C., Moscovitch, M., & Yang, J. (2019). Differential activation of the medial temporal lobe during item and associative memory across time. *Neuropsychologia*, *135*, 107252. https://doi.org/10.1016/j.neuropsychologia.2019.107252

Eichenbaum, H., & Cohen, N. J. (2001). *From conditioning to conscious recollection: Memory systems in the brain*. New York, NY: Oxford University Press.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation revisited. *NeuroImage*, *59*(3), 2349–2361. https://doi.org/10.1016/j.neuroimage.2011.09.017

Eickhoff, S. B., Laird, A. R., Fox, P. M., Lancaster, J. L., & Fox, P. T. (2017). Implementation errors in the GingerALE software: Description and recommendations. *Human Brain Mapping*, *38*(1), 7–11. https://doi.org/10.1002/hbm.23342

Fleck, M. S. (2005). Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cerebral Cortex*, *16*(11), 1623–1630. https://doi.org/10.1093/cercor/bhj097

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1338–1349. http://doi.org/10.1098/rstb.2011.0417

Fleming, S. M., & Frith, C. D. (Eds.). (2014). *The cognitive neuroscience of metacognition*. London, England: Springer.

Greenspan, R. L., & Loftus, E. F. (2020). Eyewitness confidence malleability: Misinformation as post-identification feedback. *Law and Human Behavior*, *44*(3), 194–208. https://doi.org/10.1037/lhb0000369

Haskins, A. L., Yonelinas, A. P., Quamme, J. R., & Ranganath, C. (2008). Perirhinal cortex supports encoding and familiarity-based recognition of novel associations. *Neuron*, *59*(4), 554–560. https://doi.org/10.1016/j.neuron.2008.07.035

Hayes, S. M., Buchler, N., Stokes, J., Kragel, J., & Cabeza, R. (2011). Neural correlates of confidence during item recognition and source memory retrieval: Evidence for both dual-process and strength memory theories. *Journal of Cognitive Neuroscience*, *23*(12), 3959–3971. https://doi.org/10.1162/jocn_a_00086

Heitzeg, M. M., Nigg, J. T., Hardee, J. E., Soules, M., Steinberg, D., Zubieta, J. K., & Zucker, R. A. (2014). Left middle frontal gyrus response to inhibitory errors in children prospectively predicts early problem substance use. *Drug and Alcohol Dependence*, *141*, 51–57. https://doi.org/10.1016/j.drugalcdep.2014.05.002

Henson, R. N. A., Rugg, M. D., Shallice, T., & Dolan, R. J. (2000). Confidence in recognition memory for words: Dissociating right prefrontal roles in episodic retrieval. *Journal of Cognitive Neuroscience*, *12*(6), 913–923. https://doi.org/10.1162/08989290051137468

Kao, Y. C., Davis, E. S., & Gabrieli, J. D. (2005). Neural correlates of actual and predicted memory formation. *Nature Neuroscience*, *8*(12), 1776–1783. https://doi.org/10.1038/nn1595

Kim, H., & Cabeza, R. (2007). Trusting our memories: Dissociating the neural correlates of confidence in veridical versus illusory memories. *Journal of Neuroscience*, *27*(45), 12190–12197. https://doi.org/10.1523/JNEUROSCI.3408-07.2007

Kim, H., & Cabeza, R. (2009). Common and specific brain regions in high-versus low-confidence recognition memory. *Brain Research*, *1282*, 103–113. https://doi.org/10.1016/j.brainres.2009.05.080

Kübler, A., Dixon, V., & Garavan, H. (2006). Automaticity and reestablishment of executive control—An fMRI study. *Journal of Cognitive Neuroscience*, *18*(8), 1331–1342. https://doi.org/10.1162/jocn.2006.18.8.1331

Kuchinke, L., Fritzemeier, S., Hofmann, M. J., & Jacobs, A. M. (2013). Neural correlates of episodic memory: Associative memory and confidence drive hippocampus activations. *Behavioural Brain Research*, *254*, 92–101. https://doi.org/10.1016/j.bbr.2013.04.035

le Berre, A. P., Müller-Oehring, E. M., Kwon, D., Serventi, M. R., Pfefferbaum, A., & Sullivan, E. V. (2016). Differential compromise of prospective and retrospective metamemory monitoring and their dissociable structural brain correlates. *Cortex*, *81*, 192–202. https://doi.org/10.1016/j.cortex.2016.05.002

Le, T. M., Wang, W., Zhornitsky, S., Dhingra, I., Zhang, S., & Li, C. S. R. (2020). Interdependent neural correlates of reward and punishment sensitivity during rewarded action and inhibition of action. *Cerebral Cortex*, *30*(3), 1662–1676. https://doi.org/10.1093/cercor/bhz194

Leiker, E. K., & Johnson, J. D. (2015). Pattern reactivation co-varies with activity in the core recollection network during source memory. *Neuropsychologia*, *75*, 88–98. https://doi.org/10.1016/j.neuropsychologia.2015.05.021

Leung, H. C., Gore, J. C., & Goldman-Rakic, P. S. (2002). Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. *Journal of Cognitive Neuroscience*, *14*(4), 659–671. https://doi.org/10.1162/08989290260045882

Luna, K., & Martín-Luengo, B. (2012). Confidence–accuracy calibration with general knowledge and eyewitness memory cued recall questions. *Applied Cognitive Psychology*, *26*(2), 289–295. https://doi.org/10.1002/acp.1822

Luna, K., Martín-Luengo, B., & Brewer, N. (2015). Are regulatory strategies necessary in the regulation of accuracy? The effect of direct-access answers. *Memory & Cognition*, *43*, 1180–1192. https://doi.org/10.3758/s13421-015-0534-2

Mazancieux, A., Souchay, C., Casez, O., & Moulin, C. J. (2019). Metacognition and self-awareness in multiple sclerosis. *Cortex*, *111*, 238–255. https://doi.org/10.1016/j.cortex.2018.11.012

Mihalca, L., & Mengelkamp, C. (2020). Effects of induced levels of prior knowledge on monitoring accuracy and performance when learning from self-regulated problem solving. *Journal of Educational Psychology*, *112*(4), 795–810. https://doi.org/10.1037/edu0000389

Mihalca, L., Mengelkamp, C., & Schnotz, W. (2017). Accuracy of metacognitive judgments as a moderator of learner control effectiveness in problem-solving tasks. *Metacognition and Learning*, *12*, 357–379. https://doi.org/10.1007/s11409-017-9173-2

Molenberghs, P., Trautwein, F.-M., Böckler, A., Singer, T., & Kanske, P. (2016). Neural correlates of metacognitive ability and of feeling confident: A large-scale fMRI study. *Social Cognitive and Affective Neuroscience*, *11*(12), 1942–1951. https://doi.org/10.1093/scan/nsw093

Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *Journal of Neuroscience*, *38*(14), 3534–3546. https://doi.org/10.1523/JNEUROSCI.2360-17.2018

Moritz, S., Gläscher, J., Sommer, T., Büchel, C., & Braus, D. F. (2006). Neural correlates of memory confidence. *NeuroImage*, *33*(4), 1188–1193. https://doi.org/10.1016/j.neuroimage.2006.08.003

Moritz, S., & Jaeger, A. (2018). Decreased memory confidence in obsessive–compulsive disorder for scenarios high and low on responsibility: Is low still too high? *European Archives of Psychiatry and Clinical Neuroscience*, *268*(3), 291–299. https://doi.org/10.1007/s00406-017-0783-0

Moritz, S., & Woodward, T. S. (2006). The contribution of metamemory deficits to schizophrenia. *Journal of Abnormal Psychology*, *15*(1), 15–25. https://doi.org/10.1037/0021-843X.15.1.15

Moritz, S., Woodward, T. S., & Chen, E. (2006). Investigation of metamemory dysfunctions in first-episode schizophrenia. *Schizophrenia Research*, *81*(2–3), 247–252. https://doi.org/10.1016/j.schres.2005.09.004

Moritz, S., Woodward, T. S., Jelinek, L., & Klige, R. (2008). Memory and metamemory in schizophrenia: A liberal acceptance account of psychosis. *Psychological Medicine*, *38*(6), 825–832. https://doi.org/10.1017/S0033291707002553

Mugikura, S., Abe, N., Ito, A., Kawasaki, I., Ueno, A., Takahashi, S., & Fujii, T. (2016). Medial temporal lobe activity associated with the successful retrieval of destination memory. *Experimental Brain Research*, *234*(1), 95–104. https://doi.org/10.1007/s00221-015-4415-5

Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., … Wager, T. D. (2018). Ten simple rules for neuroimaging meta-analysis. *Neuroscience & Biobehavioral Reviews*, *84*, 151–161. https://doi.org/10.1016/j.neubiorev.2017.11.012

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In *Psychology of learning and motivation* (Vol. 26, pp. 125–173). Cambridge, MA: Academic Press.

Nelson, T.O., & Narens, L. (1994). Why Investigate Metacognition?. In A.P. Shimamura (Ed.), *Metacognition*. (pp. 1–26). Cambridge, MA.: The MIT Press.

Pais-Vieira, C., Wing, E. A., & Cabeza, R. (2015). The influence of self-awareness on emotional memory formation: An fMRI study. *Social Cognitive and Affective Neuroscience*, *11*(4), 580–592. https://doi.org/10.1093/scan/nsv141

Pannu, J. K., & Kaszniak, A. W. (2005). Metamemory experiments in neurological populations: A review. *Neuropsychology Review*, *15*(3), 105–130. https://doi.org/10.1007/s11065-005-7091-6

Pannu, J. K., Kaszniak, A. W., & Rapcsak, S. Z. (2005). Metamemory for faces following frontal lobe damage. *Journal of the International Neuropsychological Society*, *11*(6), 668–676. https://doi.org/10.1017/S1355617705050873

Risius, U. M., Staniloiu, A., Piefke, M., Maderwald, S., Schulte, F., Brand, M., & Markowitsch, H. J. (2013). Retrieval, monitoring, and control processes: A 7 tesla FMRI approach to memory accuracy. *Frontiers in Behavioral Neuroscience*, *7*, 24. https://doi.org/10.3389/fnbeh.2013.00024

Robey, A. M., Dougherty, M. R., & Buttaccio, D. R. (2017). Making retrospective confidence judgments improves learners' ability to decide what not to study. *Psychological Science*, *28*(11), 1683–1693. https://doi.org/10.1177/0956797617718800

Semmler, C., Brewer, N., & Wells, G. L. (2004). Effects of postidentification feedback on eyewitness identification and nonidentification confidence. *Journal of Applied Psychology*, *89*(2), 334–346. https://doi.org/10.1037/0021-9010.89.2.334

Shimamura, A. P., & Squire, L. R. (1986). Memory and metamemory: a study of the feeling-of-knowing phenomenon in amnesic patients. *Journal of Experimental Psychology*, *12*(3), 452.

Simons, J. S., Peers, P. V., Mazuz, Y. S., Berryhill, M. E., & Olson, I. R. (2010). Dissociation between memory accuracy and memory confidence following bilateral parietal lesions. *Cerebral Cortex*, *20*(2), 479–485. https://doi.org/10.1093/cercor/bhp116

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Human Brain Mapping*, *33*, 1–13. https://doi.org/10.1002/hbm.21186

Vaccaro, A. G., & Fleming, S. M. (2018). Thinking about thinking: A coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain and Neuroscience Advances*, *2*, 1–14. https://doi.org/10.1177/2398212818810591

White, T. P., Engen, N. H., Sørensen, S., Overgaard, M., & Shergill, S. S. (2014). Uncertainty and confidence from the triple-network perspective: Voxel-based meta-analyses. *Brain and Cognition*, *85*, 191–200. https://doi.org/10.1016/j.bandc.2013.12.002

Woroch, B., Konkel, A., & Gonsalves, B. D. (2019). Activation of stimulus-specific processing regions at retrieval tracks the strength of relational memory. *AIMS Neuroscience*, *6*(4), 250–265. https://doi.org/10.3934/Neuroscience.2019.4.250

Ye, Q., Zou, F., Lau, H., Hu, Y., & Kwok, S. C. (2018). Causal evidence for mnemonic metacognition in human precuneus. *Journal of Neuroscience*, *38*(28), 6379–6387. https://doi.org/10.1523/JNEUROSCI.0660-18.2018

Yonelinas, A. P. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, *25*(11), 3002–3008. https://doi.org/10.1523/JNEUROSCI.5295-04.2005

Yu, S. S., Johnson, J. D., & Rugg, M. D. (2012). Dissociation of recollection-related neural activity in ventral lateral parietal cortex. *Cognitive Neuroscience*, *3*(3–4), 142–149. https://doi.org/10.1080/17588928.2012.669363

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Martín-Luengo B, Zinchenko O, Dolgoarshinnaia A, Leminen A. Retrospective confidence judgments: Meta-analysis of functional magnetic resonance imaging studies. *Hum Brain Mapp*. 2021;1–18. https://doi.org/10.1002/hbm.25397