nature reviews psychology

https://doi.org/10.1038/s44159-023-00245-1

Review article

Check for updates

Towards a common conceptual space for metacognition in perception and memory

Audrey Mazancieux[1,4], Michael Pereira [2,4], Nathan Faivre[2], Pascal Mamassian [3], Chris J. A. Moulin[2] & Céline Souchay [2]

Abstract

Engaging in metacognition (evaluating, controlling, and representing cognitive states) is paramount for efficient behaviour. In this Review, we examine different types of cognitive architectures that might be at play when people provide metacognitive judgements in the domains of memory and perception. Building upon this conceptual framework, we review evidence supporting and challenging domain-general metacognition. We also discuss commonalities in metacognition across domains, focusing on the influence of decisional processes on metacognitive judgements. We emphasize the challenges of isolating metacognitive processes and how these challenges influence conclusions regarding the domain generality of metacognition, including in clinical conditions that are hypothesized to have metacognitive impairments. Finally, we give an overview of 'adecisional' metacognition: evaluations made outside the context of a decisional process. We find no evidence for a strong form of domain generality but outline how such an architecture could be identified in future research.

Sections

[1]Cognitive Neuroimaging Unit, NeuroSpin Center, Institute for Life Sciences Frédéric Joliot, Fundamental Research Division, Commissariat à l'Energie Atomique et aux énergies alternatives, INSERM, Université Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette, France. [2]Laboratoire de Psychologie et NeuroCognition, Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, Grenoble, France. [3]Laboratoire des Systèmes Perceptifs, Département d'Études Cognitives, École Normale Supérieure, Paris Sciences et Lettres University, CNRS, Paris, France. [4]These authors contributed equally: Audrey Mazancieux, Michael Pereira. e-mail: celine.souchay@univ-grenoble-alpes.fr

Introduction

Metacognitive evaluations are commonplace for expressing one's confidence in ideas and decisions in daily life. For example, someone might report being 90% sure that they remembered to turn off the oven after baking a cake. Metacognitive evaluations are also important in perceptual decisions, for instance, being sufficiently confident to cross the road having judged the speed of an approaching car. These two types of evaluations are known as metamemory (metacognition for memory processes) and metaperception (metacognition for perceptual processes).

In the past 15 years there has been an expansion of metacognitive domains: metacognitive processes have been evaluated in episodic and semantic memory, visual and auditory perception, reasoning and motor function. Initially a developmental concept[1], metacognition has been investigated in the context of education[2], eyewitness memory[3] and memory impairments[4]. Metacognition has also been used as a tool to study mechanisms underlying perceptual consciousness[5,6]. Considering this range, a common conceptual space for metacognition is needed. There has also been growing interest in determining whether metacognition obeys domain-specific or domain-general rules[7] and efforts to compare metacognition across modalities and domains[8–13].

Several reviews of metacognition exist of studies conducted with individuals who are neurotypical[14,15] and individuals with pathological conditions[16], but none have directly reviewed evidence for metacognition across different cognitive domains including metamemory and metaperception. Because these two fields developed separately, there is not yet a comprehensive understanding of the cognitive architecture of metacognition broadly construed.

In this Review, we map out the domain generality of metacognition, focusing on metamemory and metaperception. Starting from the distinction between metacognitive knowledge and metacognitive experiences, we first explore possible architectures of metacognition across domains. We next review behavioural evidence for and against domain generality. Then we review relevant models and critical properties of metacognition in the perceptual and memory domains in neurotypical and clinical populations. Our goal is to build towards a common conceptual space in which domain generality can be discussed across fields.

Defining and measuring domain-general metacognition

This section sets the framework for the remainder of the Review and defines the levels of analysis that will be considered to explore the question of domain generality. We also discuss a theoretical distinction between decisional and adecisional metacognitive evaluation, which enables a comparison across perception and memory.

Two levels of representation have been distinguished in metacognition. They are interrelated, with 'metaknowledge' informing 'metacognitive experience'[1]. Metaknowledge or metacognitive knowledge encompasses specific knowledge or beliefs about one's own cognitive capabilities (such as 'I have more success detecting my brother's face in a crowd than I do remembering to send him a birthday card'), the impact of the task and strategy use. By contrast, metacognitive experience is directly related to the task at hand and the decisional process to which it pertains.

We use the terminology of information-based metacognition and experience-based metacognition to discriminate between these two notions[17]. As for metaknowledge, information-based metacognition involves inferential processes from explicit theories or beliefs. Experience-based metacognition involves the experience of a cognitive process giving rise to a metacognitive feeling through the application of heuristics[18].

We focus on experience-based metacognition because that is the most studied type in the field. Experience-based metacognition can be subtended by several types of cognitive architectures with distinct levels of domain generality. We consider four potential architectures (Fig. 1) that link the meta level and the object level, which are inspired by an influential view of metacognition[19]. We define the meta level as involving second-order representations and behaviours and the object level as involving first-order representations and behaviours (although see ref. 20).

In a strongly domain-general architecture, one metacognitive module monitors and controls several independent cognitive domains (Fig. 1a). At the algorithmic level, this meta-level module can instantiate second-order processes across domains that are independent of first-order processes. For instance, in an attempt to propose a multidomain approach to metacognition, the self-consistency model proposes that monitoring is based on a process that samples different representations from a pool of representations[21]. This sampling process can be the same across domains even though representations are specific to each domain, resulting in a domain-general metacognitive system.

A second potential architecture represents domain-specific metacognitive modules that share processes in specific first-order contexts (such as evaluation of reaction time in decision-making tasks). This architecture would correspond to a 'weaker' domain generality because

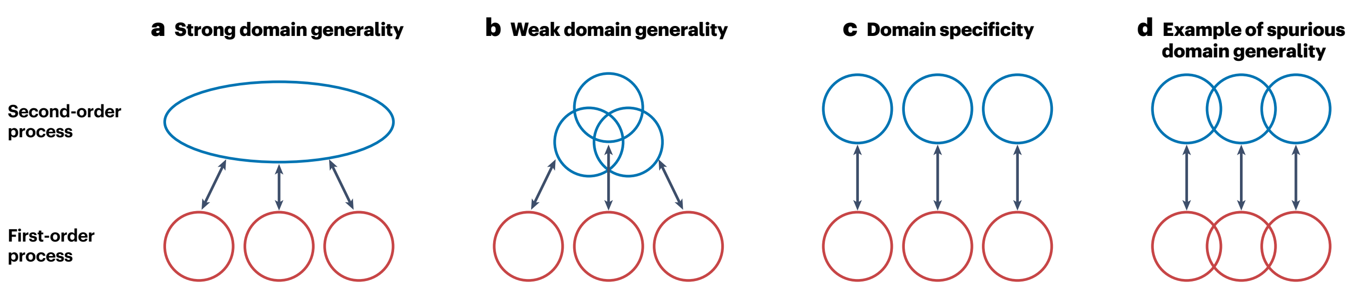


Fig. 1 | Theoretical architectures for domain-general metacognition. **a**, 'Strong' domain generality, in which the same metacognitive mechanisms are involved in the monitoring and control of all cognitive domains. **b**, 'Weak' domain generality, in which representations are specific to each domain but can share mechanisms in specific first-order contexts. **c**, Domain specificity, in which a separate metacognitive level is associated with each cognitive domain. **d**, An example of spurious domain generality in which shared first-order mechanisms impact second-order mechanisms in a domain-general way.

it is limited to some specific aspects of metacognition (Fig. 1b). A third architecture is strict domain specificity, in which no process is shared between the first-order and second-order levels (Fig. 1c). Although we have illustrated three discrete potential architectures, there is probably a continuum of possible weak domain-general architectures ranging from a metacognitive process that is systematically used (strong domain generality) to uniquely used processes (domain specificity).

Table 1 | Measures of metacognition

Direct measures		Indirect measures
Adecisional	**Decisional**	
Global predictions[a]	Judgements of learning	Re-study choice
Ease of learning judgements	Feelings of knowing	Study time allocation
		Opt-out paradigms[a]
	Confidence forced-choice[a]	Post-decision wagering[a]
		Optimal waiting time[a]

[a]Measures that can be used across different domains (notably memory and perception). All remaining measures pertain to the memory domain; there is no specific metaperception measure.

Because of the relationship between the two levels, variables that influence first-order processes in a domain-general way can also influence second-order processes, creating spurious domain generality (Fig. 1d). For instance, working memory, attention, motivation, arousal or mood can influence both cognitive and metacognitive measures across domains, giving the impression of a shared architecture. However, in this case there is not a shared metacognitive component. For example, if a good mood gives rise to better cognitive function and better metacognitive measures to the same extent, the aspects of first-order and second-order processes that share this influence of mood will seem related across domains, although there are no shared processes in the metacognitive architecture. This dependence demonstrates the importance of separating first-order and second-order processes[22].

We focus on the computational (behaviour) and algorithmic (how behaviour is computed) levels of Marr's three levels of analysis[23]. Notions of strong or weak domain generality can also be examined at the implementation level (for a review of the neural correlates of metacognition, see ref. 24). For instance, if the same neurons or regions are involved in the computation of metacognition for several domains, this would suggest a strong domain-general architecture[25]. However, it is possible to have a strong domain-general process at the algorithmic level without having a strong domain-general implementation (such as if several task-specific regions can perform similar computations).

In the remainder of this Review, we review evidence to distinguish between the possible architectures of metacognition described above. We focus on direct measures (Table 1) in which participants are instructed to self-reflect about a first-order task (such as the encoding of information, the recall of information or a decision) and report about their self-reflection (Box 1). These measures include asking participants to judge how difficult it is for them to learn something (ease of learning judgements), how likely they are to retrieve something (judgements of learning or feeling of knowing judgements) or how confident they are in a decision that they made (retrospective confidence judgements). Participants can also be asked which of two consecutive decisions is more likely to be correct (confidence forced-choice), or to evaluate their performance over several decisions (global judgements).

In indirect measures, other behaviours are used to infer metacognitive processes during a given cognitive process (such as the encoding of information or a decision) and participants are not directly asked to introspect and evaluate their performance (Box 2). We favour evidence from direct measures as it is still debated whether indirect measures truly reflect metacognitive behaviour or whether these behaviours could be obtained through first-order processes such as reinforcement learning (although there is some evidence for metacognition in decisions related to motor output; Box 3). We focus on metamemory and metaperception as very little work has been done linking other domains.

Within direct measures, a classical distinction has been used in the metamemory literature between prospective (occurring before memory retrieval) and retrospective (occurring after memory retrieval) judgements[26]. However, this distinction does not help to develop a shared conceptual space that functions across domains. For instance, a judgement of learning[27] is classified as a prospective judgement, because it is a prediction of upcoming recall once a stimulus has been encoded. Nonetheless, the judgement of learning is based on the experience during the encoding stage, and as such pertains to a second-order judgement about a (past) process. Furthermore, these judgements arguably rely on a retrieval attempt, as has been shown by delaying the point at which a judgement of learning is made[28]. Likewise, the feeling of knowing judgement is a prediction of future recognition[29,30]. A retrieval attempt or evaluation of a presented cue is necessary to gauge the likelihood of future recognition. Indeed, partial knowledge available during recall has been shown to correlate with the accuracy of feeling of knowing judgements[31,32]. Describing feelings of knowing as prospective because they are made for an upcoming test overlooks the fact that they result from a decision about the retrievability of information from memory. This judgement is subtly different from the judgements taken in metaperception designs[33] in which participants first give their confidence level and then perform the decisional task (but see ref. 34 for prospective metaperceptual judgements performed before stimulus onset). In short, the prospective–retrospective distinction limits the classification of metacognitive evaluations to metamemory because it is only pertinent to encoding storage-retrieval designs.

Rather than using the prospective–retrospective distinction common to metamemory, we distinguish decisional metacognition as any type of evaluation made around the point at which a specific decision is made during a task qualified as first-order (or sub-task, such as encoding a word for later recall). Our conceptualization of decisional metacognition is critical for constraining the types of cues and experiences that are defined as metacognitive and improves the ability to compare metaperception and metamemory tasks. Adecisional metacognition groups together beliefs, knowledge about functioning and strategic regulation; the evaluation is not derivative of any specific first-order process. Thus, the extent to which metacognition is domain general in the real world rests upon the balance of decisional and adecisional factors, which will change according to context and task demands.

The critical distinction in this conceptualization is that in a decisional metacognitive evaluation a first-order decision is made, which leaves a trace from which one can extract some information to make a second-order evaluation. If no decision has been made, then one has no access to such decisional information on which to base their judgement, and the evaluation is of a different kind. Referring to these former evaluations as decisional judgements enables comparison across common methods and models that pertain to multiple cognitive domains. The most frequently used second-order decisional measure

Box 1

Direct measures

Global judgements

Global judgements are an adecisional measure of metacognition and involve estimating performance at a task level, such as how many words will be correctly recalled from a list[155], how many words from a particular category were generated in a given period[137] or a prediction of the grade achieved in a university examination[191]. A critical factor is that the prediction is typically made before and after completing the task. For example, predictions of future performance on an examination are made before the examination and again once the examination is completed. The initial prediction before the task starts (or before the university course has been taught) acts as a reference point for interpreting the global prediction. Any shift between an initial and an informed prediction is based on the capacity to monitor the task, and the initial pre-task prediction is based on expectancies, beliefs, and metacognitive knowledge[173]. Global judgements do not refer to any one decision, and as such reveal processes used in generalized evaluations of function, rather than pinpointing metacognitive mechanisms and processes.

Ease of learning judgements

Ease of learning judgements[192] are adecisional predictions about what will be easy or difficult to learn and pertain to items that have not yet been learned. Participants are therefore not asked to learn the items for an upcoming memory test when making their judgement. On the presentation of each item, the typical question asked of the participants is 'How likely is it that you will learn this word for the test?'. After making a judgement for each item, participants are asked to study the items and recall them. Despite the importance of this initial assessment of how difficult a material is and its potential impact on learning, most studies suggest that ease of learning judgements poorly or moderately predict the actual learnability of material[193–195] but this prediction increases when the gap between easy-to-lean and hard-to-learn items is higher[196].

Judgements of learning

Judgements of learning take the same form as an ease of learning judgement but are made after having attempted to learn the item, therefore making them decisional judgements. Typically, the judgement of learning is made for cue–target word pairs, with the metacognitive judgement made for the likelihood of retrieving a target when prompted by the cue word. Similar to ease of learning judgements, they are judgements of the likelihood of subsequent recall. Thus, this measure has likewise been used to explore the cues that are used to make metacognitive judgements of memory and factors that influence recall (for meta-analytic reviews, see ref. 197).

Judgements of learning can be made immediately after the encoding of the item or after a delay (either in a second phase after the initial block of encoding of word pairs or, more typically, after several intervening cue–target pairs). There is a robust phenomenon whereby metacognitive sensitivity is higher for judgements made after a delay than immediately[198]. Another well-established judgement of learning phenomenon is the font-size effect, whereby the magnitude of judgements of learning is increased by font size: words written in a large font at encoding are judged more likely to be recalled than words in a smaller font, even though words written in large font are not better recalled than words written in small font (for an explanatory meta-analysis, see ref. 199).

Feeling of knowing judgements

Feeling of knowing judgements[30] are decisional predictions about the likelihood of subsequent recognition of information one currently cannot recall[29,200]. In a feeling of knowing experiment, participants are presented either with new information to learn, such as word pairs (episodic memory task), or are presented with general knowledge questions, such as 'what is the capital of France?' (semantic memory task). In episodic memory tasks, after the learning period, participants are presented with the first word of the pair and asked to recall the second word. In semantic memory tasks, participants are presented with the question. In either case, if they cannot recall the information, the feeling of knowing judgement is to predict whether they will be able to recognize the missing information if presented to them later. Thus, feeling of knowing judgements are predictions about material that participants failed to retrieve and have been found to be relatively predictive of future recognition in young adults[29,201,202].

Retrospective confidence judgements

Retrospective confidence judgements are decisional judgements and the most common measure of metacognition used in the field of metacognition. They refer to the level of confidence that a participant has in a given answer, measured using a multiple-point scale (for example, from one to six). They have been extensively used in decision-making[15], notably to investigate cross-domain comparisons[10,11,39], but also in other tasks such as statistical learning[203].

Confidence forced-choice

The confidence forced-choice paradigm[204] requires participants to choose which of two decisions about two different stimuli is more likely to be correct. For example, participants first have to make a visual decision (for instance, presence or absence of a visual stimulus) followed by a second auditory decision (for instance, presence or absence of an auditory stimulus). They then have to say whether they feel the first decision is more likely to be correct than the second. By varying the difficulty levels within pairs of stimuli, researchers estimate a psychometric function for chosen versus declined decisions. The difference in slopes between these two curves serves as a proxy for metacognitive performance, irrespective of confidence bias. By asking participants to choose between decisions that pertain to different cognitive domains, this method can be used to characterize the domain generality of metacognition[205].

pertains to confidence. Retrospective confidence judgements refer to the level of confidence that a participant has in being correct on a given first-order decision. As they can be performed on any kind of first-order decision, they have been the main measure used to investigate domain-general metacognition. However, metacognition is not only limited to decisional judgements, and we offer some speculative comments on adecisional judgements later on.

Decisional metacognition and adecisional metacognition differ from the information-based metacognition and experience-based metacognition distinction because decisional judgements can be information-based (such as bias or starting point of evidence accumulation) or experience-based (such as drift rate or sensitivity). Likewise, adecisional judgements can be information-based (such as beliefs) or experience-based (such as emotional state). However, decisional metacognition is largely experience-based and adecisional metacognition is largely information-based.

Decisional judgements

In this section, we review evidence for and against domain-general metacognition by considering metacognitive bias and metacognitive sensitivity in decisional judgements. Using models in the metaperception and metamemory fields, we describe potential domain-general processes with reference to the proposed metacognitive architectures (Fig. 1).

Evaluating domain generality

There are different approaches for assessing the domain generality of a metacognitive process in behavioural studies. The first and most common class of methods consists of correlating metacognitive bias or metacognitive sensitivity across domains, typically at the level of a group of participants. Metacognitive bias refers to the overall magnitude of metacognitive evaluations irrespective of performance. Metacognitive sensitivity refers to the ability to discriminate between correct and incorrect responses when performance is neither at ceiling nor at chance levels. The optimal calculation of metacognitive sensitivity has been widely discussed and reviewed[22,35–38] (Box 4). Domain generality implies that evaluations of decisional judgements should be correlated across different domains: individuals with high metacognition in one domain also have high metacognition in other domains.

Correlation studies consistently find that metacognitive bias is stable across many domains including perception and episodic and semantic memory[8,10–13]. However, the pattern appears more complex for metacognitive sensitivity. A previous review concluded that although domain-general metacognitive sensitivity can be identified for perception across different modalities (such as audition, touch and vision), correlations across metamemory and metaperception are low[7]. These low correlations were possibly due to variability in measures of metacognitive sensitivity and the low sample size of the studies. Using more appropriate measures (such as metacognitive efficiency[22]; Box 4) that isolate the confound between first-order and second-order performance as well as more appropriate sample sizes, several studies have found positive correlations across memory and perception tasks[10–13,39–41] (although see ref. 42). These correlations are restricted to two-alternative forced-choice tasks (2AFC or discrimination tasks) and seem to be absent for yes–no tasks (detection tasks[10]), when the two types of task have been compared. These kinds of results might point to different computations underlying confidence for detection and discrimination tasks even within a single domain[43].

Box 2

Indirect measures

In indirect measures of metacognition, participants are not directly asked for a self-evaluation and, instead, other behaviours are used to infer metacognition. Two major measures used in metamemory and metaperception are post-decision wagering and opt-out paradigms. In post-decision wagering, participants have to place bets on the accuracy of their decisions. Early versions of this paradigm were used as an objective measure of subjective visibility[206], assuming that participants would place higher bets following seen stimuli than unseen stimuli. Concerns were raised that this measure was affected by loss aversion[207] and pertained more to metacognitive access than to subjective visibility[208]. Consequently, the measure is now used to incentivize optimal metacognitive judgements irrespective of loss aversion[209,210]. Opt-out paradigms provide a proxy for low confidence, allowing participants to opt out of a decision if their putative confidence in the choice is low. This procedure has the advantage of being applicable to non-verbal species such as non-human primates[81], rodents[76] or preverbal infants[211]. In some versions of the paradigm, an opt-out response option is provided[81], whereas in others a delay is imposed between the response and the reward, during which participants can opt out and restart a new trial without reward[76,211]. In the latter case, the waiting time can be used as a continuous proxy for confidence. Evidence for domain-specific metacognition has been recently found using an opt-out paradigm with non-human primates[212]. The main criticism of this measure is that participants could correctly opt out of a decision by simple reinforcement learning without relying on a second-order monitoring mechanism[213] (for a more general overview of animal metacognition, see refs. 214,215).

Specific indirect measures have been developed for metamemory. When two tests using the same items are performed, one can measure the time a participant re-studies an item (study time allocation[216]) or the decision to re-study an item or not (re-study choice[217]). In this context, participants allocate more time to re-study an item that they did not previously recall compared with recalled items, implying that they have accurate knowledge about previous failures[195].

Although the presence of between-subject correlations across domains supports domain-general metacognitive efficiency, there is large variability in the magnitude of correlations. Most of the above-mentioned studies that reported correlations at the group level used hierarchical estimations of metacognitive efficiency[44], which takes into account both within-subject and between-subject variability to directly estimate covariance in metacognitive efficiencies across domains but can also inflate correlation estimates[45] (see refs. 40,46 for evidence for domain-general metacognition based on non-hierarchical estimates). Furthermore, measures of metacognitive efficiency have low half-split reliability[47] which makes it difficult to estimate metacognitive abilities as a trait.

Another issue with research in this area is that correlations across domains might actually reflect the influence of domain-general

Box 3

Motor metacognition

Although most studies on metacognition have focused on the perceptual and memory domains, the motor domain has seen a surge of interest in the past 10 years. An early attempt to characterize motor metacognition found that when given distorting visual feedback while reaching, participants appropriately adjusted their hand trajectories but misjudged the effects of the distortion[218]. This result suggests that people have limited access to the details of movements while achieving a goal[219,220]. Similar metacognitive inefficiencies are found when participants are asked to follow an unpredictably moving dot cloud with a mouse cursor[221].

Other studies reported that participants appropriately adjusted their confidence when detecting variable amounts of visuomotor distortions, which seems to contradict the notion of limited access to motor performance[222]. This contradiction could be resolved by considering that participants optimally calibrate their confidence in detecting distortions — even when they fail to report them — based on a summary statistic of the visual feedback[223]. In other words, one can automatically correct distortions that remain undetected, and still have a subjective feeling of confidence that tracks performance. Although this heuristic might be useful for automatically monitoring motor actions, it does not specify whether people can explicitly monitor low-level movement parameters when prompted. When instructed to do so, participants can monitor their performance in throwing a virtual ball based on the position of their arm as well as the resulting trajectory of the ball thrown[87]. Generally, participants might be better at monitoring some aspects of their motor actions (such as the movement duration) than others (such as when they initiated the movement[224]).

Finally, studying motor metacognition is also of interest regarding the distinction between the monitoring of internal (such as mnesic) and external (such as sensory) signals, as both efferent and reafferent signals might serve as objects for meta-representations. In one study using a robotic device to passively move participants' fingers, there was only a change in confidence bias for active versus passive finger movement, suggesting a contribution of efferent signals on confidence but not metacognitive efficiency[225]. In the same vein, actions paired to visual stimuli were found to increase confidence ratings but leave metacognitive performance unchanged compared with a condition with no paired actions[96]. These results suggest that confidence ratings do not improve based on efferent information. Overall, these results indicate that motor processes are the subject of metacognitive evaluations, and that efferent signals can a priori modulate confidence judgements without impacting the quality of metacognitive monitoring. It will be important to refine the contribution of motor signals to metacognitive monitoring in the future to further establish the domain-general nature of the metacognitive architecture.

metacognitive bias. In theory, metacognitive judgements can be seen as accurate (sensitivity) when metacognitive bias is high (overconfidence) or low (underconfidence) and measures of bias and sensitivity should be independent. However, in practice, correlations between metacognitive efficiency and bias are found[48]. Thus, as metacognitive bias is mainly domain general[8,10–13,41], a spurious domain-general metacognitive sensitivity could arise when bias and sensitivity correlate. In sum, most of the evidence supporting the domain generality of confidence is relatively weak because it is mainly based on between-subject correlations.

Beyond the between-subject correlation approach, more direct methods have been used to address whether metacognition involves domain-general mechanisms. One method is to determine whether confidence is encoded with a 'common currency' across different tasks. For instance, in the confidence forced-choice paradigm, participants were able to compare confidence across visual and auditory decisions with the same precision as for the comparison of two trials within the same sensory modality[49]. Another study used an audiovisual discrimination paradigm in which participants had to judge whether the most salient visual and auditory stimuli were on the same side[9]. Computational models reproducing confidence estimates about such audiovisual decisions use supramodal formats of confidence in which auditory and visual confidence signals are either integrated or compared with one another, also suggesting a common currency. Research using this approach remains limited to comparisons between sensory modalities, and evidence supporting a common currency of confidence across the domains of perception and memory is lacking.

Another approach involves assessing the domain generality of confidence in individuals with neuropsychological impairments. One study compared metacognitive efficiency across three groups of participants (individuals with lesions in the anterior prefrontal cortex, individuals with temporal lobe lesions and neurologically intact individuals) and found a deficit in perceptual (but not memory) metacognitive efficiency in the group with anterior prefrontal cortex lesions relative to individuals with lesions in the temporal lobe and neurologically intact individuals, and no difference in first-order performance or metacognitive bias[50]. Another study also found a deficit in perceptual (but not memory) metacognitive efficiency with substance-dependent individuals, supporting domain-specific impairments of metacognition[51].

However, such dissociations do not occur in other contexts. No specific differences were found between metacognitive efficiency for visual perception and semantic memory for subclinical psychiatric symptom dimensions (anxious depression, compulsive behaviour and intrusive thought, and social withdrawal)[40]. In participants who are neurotypical, the disruption of the precuneus through transcranial magnetic stimulation has been shown to selectively alter metacognitive efficiency in a memory task but not a visual perception task[52,53]. In the memory domain, neuropsychological studies have consistently shown inaccurate episodic feelings of knowing with preserved semantic feelings of knowing in different populations such as individuals with multiple sclerosis and autism spectrum disorder[54,55], although this pattern might be driven by first-order performance differences and therefore uninformative regarding metacognition. Overall, despite a growing interest in the neuropsychology of metacognition, the evidence supporting domain-general or domain-specific architecture is mixed, and systematic investigations of bias-free metacognitive performance indices across various populations[56] and cognitive domains are needed.

Models of decisional judgements

Here we review models of decisional judgements that can account for processes driven by first-order or second-order evidence and as such

Box 4

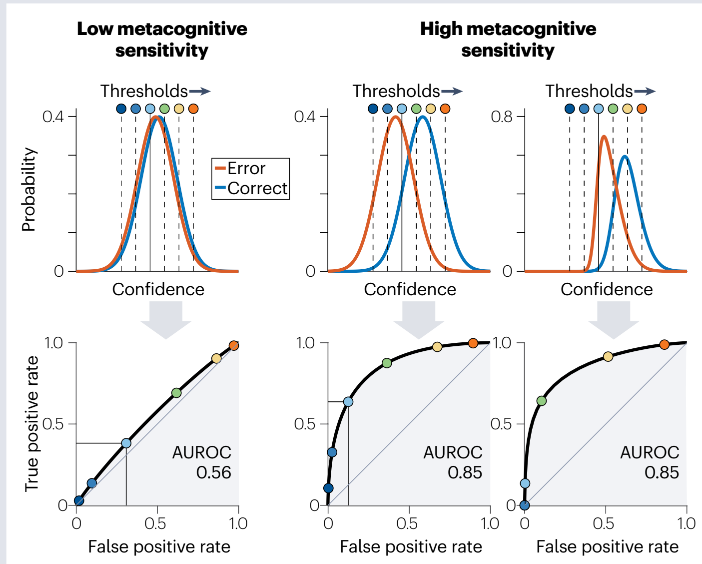
Quantifying metacognition

Metacognitive sensitivity relates to the ability of an individual to adjust a decisional judgement (typically confidence) to the performance of the first-order task. Thus, measures of metacognition are not directly informative of metacognitive sensitivity as they have to be compared with actual task performance.

A first measure of metacognitive sensitivity is the confidence gap, which compares average confidence judgements after correct versus incorrect first-order responses (see the figure, upper graphs) but overlooks the variance of the two distributions. For instance, a participant with metacognitive bias who only uses the upper part of the confidence scale would have a lower confidence gap than a participant who uses the entire confidence scale. Another approach is to assess the relationship between confidence judgements and first-order performance using Pearson's or Goodman–Kruskal γ-correlations[226]. However, these methods cannot isolate metacognitive sensitivity from metacognitive bias[22,37].

The issue of contamination of metacognitive sensitivity by metacognitive bias can be avoided by computing the area under the receiving operating characteristic (AUROC) curve (see the figure, shaded areas in lower graphs). A two-dimensional curve is constructed by computing the percentage of correct responses classified as such (true positive rate; vertical axis in lower graphs) and the percentage of incorrect responses classified as correct (false negative rate; horizontal axis in lower graphs) by setting different thresholds on the confidence (coloured dots on upper and lower graphs). The area under this curve ranges from 0.5 (chance level) to 1.0 (a threshold that perfectly classifies correct and incorrect responses). AUROC curves can dissociate metacognitive sensitivity from bias[77,227].

Metacognitive efficiency refers to the metacognitive sensitivity given the information available for the first-order decision. To compare metacognitive efficiency using AUROC curves, task performance should be equalized across conditions and participants. Another possibility that does not require this equalization is to use a model-based approach and compare first-order sensitivity (*d*′) and an estimated-*d*′ value or meta-*d*′ value given an 'ideal observer' model of confidence without metacognitive noise[77]. Many newer studies use the ratio between meta-*d*′ and *d*′ values, termed the *M*-ratio[44], which enables control of task performance differences and metacognitive bias. Finally, new approaches are being proposed that attempt to fit metacognitive noise with a generative model of confidence judgements[228], including in situations in which the model includes other parameters that can be easily confounded with metacognitive noise[60].



could point to domain-general mechanisms. Computational models attempt to explain how decisional judgements originate from sensory evidence. They mostly differ in the way that first-order evidence is reused (or not) by putative second-order metacognitive processes[57]. Most models are hierarchical: they assume a common source of evidence for first-order decisions and second-order judgements, but with possibly different readouts (Fig. 2). Other models assume either that evidence for decisions or recall and metacognition is not identical (albeit correlated)[58–60] or that evidence for metacognition is processed separately from the first-order process[61,62].

Hierarchical models define confidence as a readout of the first-order decisional process, which is modelled either as a static snapshot of perceptual evidence (based on signal detection theory or an extension thereof), or as a dynamic process that accounts for the way first-order decisions unfold over time. For perceptual decision-making, dynamic process models assume that noisy sensory evidence is accrued over time up to a boundary[63], leading to a decision. Similar models have been developed for recognition memory, assuming that the match between a test item and memory produces evidence that is also accrued over time[64,65]. Notably, these dynamic models can also be extended to be closer to a neural implementation using neuronal networks[66–69]. A second distinction of first-order models is whether evidence is encoded in a relative way (positive for one choice, negative for the other choice such as in signal detection theory or for the drift diffusion model[70]) (Fig. 2a,b) or whether absolute evidence is encoded for each choice, enabling multiple-choice decisions (such as in two-dimensional signal detection theory[71] or accumulator models[72]) (Fig. 2c–f).

In most hierarchical models, confidence depends on the level of perceptual evidence, whatever the underlying first-order model (static or dynamic, absolute or relative evidence). The third orthogonal distinction between models is the way in which perceptual evidence is transformed into a confidence value[14], which varies largely depending on the underlying first-order model and whether or not it considers post-decisional evidence.

The simplest hierarchical models of confidence formation are solely based on the strength of first-order evidence (whether static

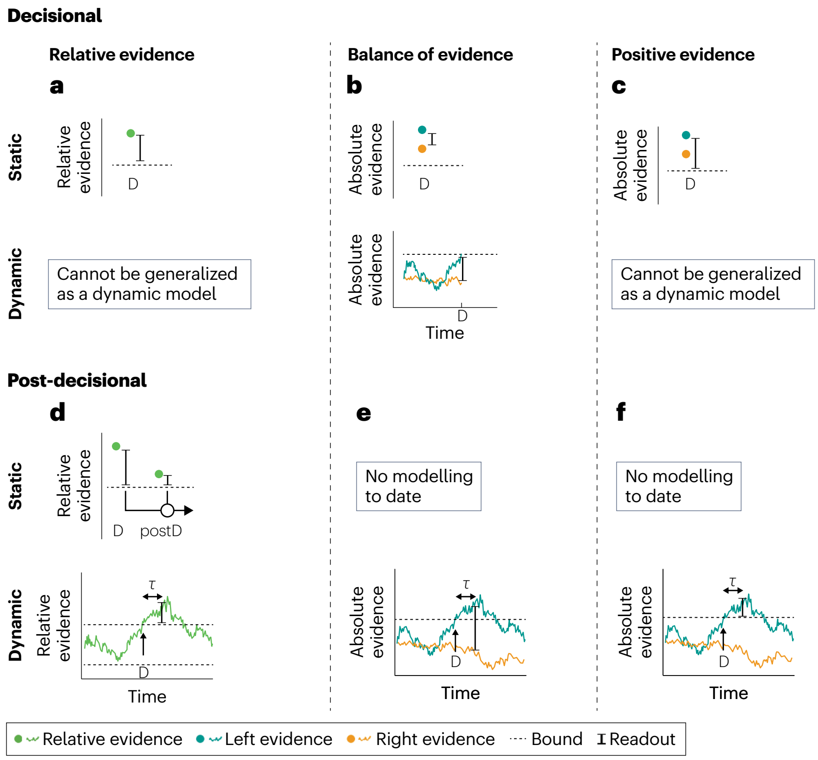


Fig. 2 | Confidence readouts for hierarchical models in a typical left–right discrimination task. a–f, Confidence can be read out from evidence (vertical axis) at the time of the decision ('decisional'), or later ('post-decisional'). Static models assume that only snapshots of evidence are available, and dynamic models capture how evidence evolves over time (horizontal axis). Perceptual evidence can be either relative (left minus right; **a**,**d**) or absolute, encoding evidence for left and right separately (**b**,**c**,**e**,**f**). Confidence is read out from the distance between (static) relative evidence (green full circle) and the decision bound (horizontal dashed line) at the time of the decision D[75,76]. This readout cannot be generalized to a dynamic model as accumulated relative evidence is constant at the time of the decision, unless a collapsing bound is assumed and confidence would then be equivalent to the response time (panel **a**). Static models of absolute evidence can read out confidence as the balance of evidence for left and right choices at the time of decision D[184]. In dynamic models, this is equivalent to reading out the state of the losing accumulator (right – incongruent with the choice[64,81]) (panel **b**). Static models can also read out confidence from positive evidence only (congruent to the choice[57]). This readout cannot be produced for dynamic models as positive evidence is constant at the time of decision (see panel **a**) (panel **c**). The static model in panel **a** can integrate decisional (D) and post-decisional (postD) evidence either as a weighted sum or through a second-order mechanism[59]. In dynamic models, relative evidence accumulated post-decisionally can be read out at a fixed time τ[185,186] (panel **d**). Other post-decisional dynamic models of confidence delay the confidence readout by a fixed time τ[187,188]. Confidence can be based on a balance of evidence[189,190] (panel **e**) or on positive evidence (panel **f**) only, congruent to the choice[102].

or dynamic) and assume a readout of the distance between noisy sensory information and a decision boundary. These models are found in recognition memory[73,74] and perception[75,76], with decisional judgements possibly degraded by an additional source of (metacognitive) noise[77–79]. Other confidence readouts have been developed within a probabilistic framework. Examples include defining confidence as the probability of a choice being correct knowing the sensory evidence[80], as a log-probability ratio of two possible choices[81], or as the precision (inverse variance) of the underlying sensory distribution[82]. Newer models have formalized the idea that decisional judgements are based on both prior beliefs about memory (information-based metacognition) and processing experience during the memory process (experience-based metacognition)[18]. Such Bayesian inference models of confidence are found in perception[83,84] and memory[85].

In sum, the aforementioned models describe how the strength of first-order evidence (possibly augmented by post-decisional evidence) relates to confidence. Although commonalities have been found across domains in the way that evidence is read out to build confidence, cross-domain modelling studies are needed to better isolate a common mechanism. As most published works consider first-order evidence as domain-specific, the domain generality of metacognition could either arise from a common source of metacognitive noise[9,49] or from (possibly domain-general) factors that differentially influence second-order metacognitive processes.

Isolating second-order processes

Hierarchical models imply that any factor influencing first-order processes will partially be reflected in decisional confidence judgements, but this overlap does not imply that the factor directly influences metacognition. For example, different types of visuospatial attention impact discrimination performance at the first-order level but not metacognitive sensitivity when controlling for first-order performance[86]. Similarly, visual information improves first-order performance in motor tasks but not metacognition per se[87]. Thus, to isolate a second-order process, one needs to find differences in confidence that cannot be explained by differences in first-order processes, either by titrating task difficulty[88] or by comparing with the ideal confidence observer[77]. Consequently, it is crucial to account for first-order performance to be able to test for the domain generality of metacognition (Fig. 1a,b), or else one risks being fooled by spurious domain generality (Fig. 1d) that is driven by factors that commonly influence first-order performance, such as attention. In the following sections we review dissociations between first-order and second-order processes in adults who are neurotypical, individuals with clinical conditions and children. Studies with adults who are neurotypical often used decisional measures, whereas developmental studies and patient studies use paradigms that are suitable for these specific populations, mainly underpinned by information-based metacognition and using adecisional judgements.

Dissociations in adults who are neurotypical

In the perception domain, one procedure to find dissociations at the second-order level is to manipulate the level of positive evidence for a decision while maintaining the same signal-to-noise ratio, which leads to changes in confidence independent from first-order performance[89,90]. Confidence can also be manipulated while keeping first-order performance constant by increasing the volatility of sensory evidence through time-varying noise[91], which suggests that this volatility impacts second-order processes (metacognition) specifically. Various other factors have been found to influence second-order processes specifically. Metacognitive sensitivity is higher for perceptual target detection when decisions are congruent with prior expectations of target presence[83] or for unexpected action outcomes[92]. Furthermore, participants make better use of prior expectations (such as internal expectations or external cues that a stimulus will be of one sort) at the metacognitive level than for first-order decisions[93].

Metaperception can be influenced by other factors such as the presence or absence of a motor action to enact the decision. Confidence is specifically modulated by the presence of electroencephalographic and electromyographic activity occurring before participants provide a first-order response[94]. The link between confidence and motor activity is also supported at the behavioural level, notably by studies that manipulated or characterized motor commands associated with first-order responses[9,94–97], as well as sensorimotor activity leading to the first-order response[95–97]. Conversely, metacognitive performance decreases when first-order responses are given after confidence ratings[98], perturbed using transcranial magnetic stimulation[99] and sensorimotor conflicts[95], or made by an external agent[100–102]. In sum, numerous findings have reported dissociations in metacognition by showing that confidence varies between two conditions in a way that cannot be explained by differences in first-order performance alone.

Studies in the memory domain have used manipulations that differentially influence first-order performance and confidence. For instance, varying the strength of evidence for lures modified memory performance without changing the magnitude of confidence judgements[103]. Many studies have also investigated such dissociations using judgements of learning. For instance, perceptual manipulation of to-be-learnt words such as larger font[104] and louder volume[105] increase the magnitude of judgements without influencing recall performance. Such effects, known as metacognitive illusions[106], have also been observed when words are blurred[107] or for auditory degraded words[108] and result in a decrease in the magnitude of judgements.

So far, the same manipulations have not been studied in a memory task and a perception task in the same study. Thus, metacognitive illusions within a domain do not provide direct evidence for domain generality. However, the pattern of findings observed in metacognitive illusions might point to weak domain generality by identifying factors that are specific to second-order judgements and that might or might not be shared between domains.

Some factors clearly generate dissociations between the metacognitive level and first-order performance. These dissociations are further supported by transcranial magnetic stimulation studies that show that metacognitive efficiency can be specifically influenced with no effect on first-order performance. Metacognitive efficiency increases when disrupting brain activity in the occipital cortex[109] and decreases when disrupting brain activity in the premotor[99] or dorsolateral prefrontal cortices[110,111] (but see ref. 112) in perceptual tasks. In the memory domain, metacognitive efficiency also increases when disrupting the dorsolateral prefrontal cortex for temporal working memory[113] but decreases when the perturbation occurs on the precuneus for episodic memory[52].

However, it is unclear how factors such as motor processes are integrated into decisional judgements. For example, reaction time positively correlates with confidence and influences these judgements[114,115] but this effect can be explained by a metacognitive process integrating reaction times into confidence ratings[116] or by first-order processes such as the level of correlated noise between two accumulators[117]. Likewise, response caution – how participants balance response speed and accuracy – influences decisional judgements by changing the amount of post-decisional information available: positive correlations have been found between reaction time and metacognitive efficiency[118]. It is therefore unclear whether post-decisional evidence accumulation is a first-order or second-order process. These considerations are important as they could imply spurious domain generality; domain-specific first-order processes could inform metacognitive processes differently from first-order decisions and common metacognitive measures would fail to diagnose such cases[118].

Another possibility is that people might have access to cues to specifically inform their decisional judgements beyond first-order processes[119]. When making a metamemory decision, people might infer the strength of their memory trace using cues, which can be diagnostic if pertinent to the memory retrieval itself or non-diagnostic if they have no influence on retrieval success. Mnemonic cues include internal indicators or signals that can be used to evaluate one's level of memory performance. Several mnemonic cues, such as the familiarity of the probe[120] or its fluency[121], influence various decisional judgements. Fluency refers to the subjective experience of processing information easily[122]. In particular, answer fluency (the ease with which information comes to mind[123]) influences confidence judgements across multiple domains. Responses that are easily retrieved (both correct and incorrect) are judged with higher confidence in semantic memory tasks[124]. Within the fluency framework, response times could also be used as a cue to inform confidence[116]. In metaperception, people can be fooled in their confidence judgement by some visibility cues[125]. For instance, a stimulus with high contrast or a long presentation duration might increase confidence, neglecting the fact that these positive properties can be annihilated by less obvious cues such as pixel noise (and there is some evidence for an alternative interpretation)[89,90]. Overall, decisional judgements across domains are influenced by cues even when these cues are not diagnostic of performance, suggesting a dissociation between first-order and second-order processes.

In sum, strong evidence supports a dissociation between the information available to first-order and second-order processes. This conclusion implies that some factors – such as sensorimotor activity when reporting a decision – might influence metacognition similarly for different domains and therefore result in domain generality. However, other factors are inherently specific to a domain, such as sensory noise in metaperception[91] or familiarity in metamemory[120], and this specificity weakens the conclusion of domain generality. Alternatively, one could posit a domain-general metacognitive process that adaptively weighs different factors (some domain-general and some domain-specific) according to the task at hand, thereby incorporating these factors into strong domain generality. For instance, some researchers have proposed that cross-task correlations for metacognitive sensitivity might be driven by common sources of metacognitive inefficiency (that is, different types of noise) for different tasks[126]. Consequently, cross-task correlations can be observed under a strong domain-general architecture in which a common metacognitive module is altered by different types of metacognitive noise and first-order factors impact distinct metacognitive modules. Future work is required to better understand the underlying mechanisms of these dissociations between first-order and second-order processes. Even when accounting for first-order performance, it is not always straightforward to ascribe the influence of a factor to first-order or second-order processes, which could lead to spurious domain generality being overlooked.

Clinical dissociations

Dissociations between first-order and second-order processes have also been observed in clinical populations. Individuals with frontal lobe lesions were impaired on an episodic memory and not on a semantic memory feeling of knowing task. However, the episodic feeling of knowing deficit was present only when the memory task was completed 3 days after encoding versus 5 min after encoding. These results therefore showed that the feeling of knowing deficit was linked to the memory deficit[127]. However, other individuals with impaired memory and frontal and temporal lobe lesions showed preserved decisional metamemory in episodic feeling of knowing judgements, judgements of learning, or confidence judgements[128–130].

Decisional judgements have also been explored in Alzheimer's disease. In individuals with Alzheimer's disease, judgements of learning, semantic feelings of knowing[131–133] and confidence judgements[134,135] were accurate, despite impaired memory performance. However, episodic feelings of knowing were found to be inaccurate[136]. A dissociation between episodic and semantic feeling of knowing was also found in individuals with multiple sclerosis: Episodic feeling of knowing deficits were more prevalent in individuals with low memory performance[54] but no deficits were present when first-order performance was controlled[137]. Feeling of knowing judgements have also been explored in Parkinson's disease, where first-order memory impairment is associated with second-order deficits in these judgements[138,139]. In functional cognitive disorder, individuals showed no impairment in metamemory or metaperception when controlling for first-order performance, but differences were found in global reports of subjective performance[140].

Decisional judgements have also been explored in psychiatric disorders such as bipolar disorder. Most studies with individuals with bipolar disorder report limited correspondence between memory performance and metacognitive judgements[141], mainly observing underestimation of memory performance rather than a deficit in metacognitive sensitivity. Underconfidence in both memory and perception was also observed in individuals with obsessive-compulsive disorder[16], but metacognitive sensitivity deficits were only found in a subclinical sample[142,143] and not in a clinical sample[144]. Finally, deficits in decisional judgements have been observed in vision[145,146], audition[147] and memory[148,149] in individuals with schizophrenia, but a meta-analysis found no evidence for a metacognitive deficit among studies that controlled for first-order performance[150].

Global judgements, which are simple and have low attentional demands, are particularly suited to studies with clinical patients. These studies operationalize accuracy as the unsigned difference between predictions and performance. Patients with Alzheimer's disease, when comparisons between control groups and patient groups have been made, overestimate their memory performance for word lists[151,152], flashbulb memories[153], and perceptual performance in visuospatial tasks[154]. However, metacognitive accuracy improves when predictions are made after participants have experienced the task[151,155,156].

In sum, most research with clinical populations outlines metacognitive impairments that could simply be a consequence of a first-order deficit. Particularly in metamemory, the majority of studies have used measures of metacognitive sensitivity that do not control for differences in first-order performance (such as the γ-correlation). It is therefore unclear whether such deficits are purely metacognitive or the consequence of a first-order deficit. To conclude in favour of a pure metacognitive deficit, a more convincing datum would be a population of individuals with impaired second-order performance in the context of preserved first-order performance. The case of blindsight might correspond to such a dissociation between perception and metaperception, although the specific level at which vision is impaired remains debated[157,158]. In blindsight, individuals who have cortical blindness due to a damage to the visual cortex have residual abilities to discriminate or detect certain types of stimuli without being aware of this ability.

In metamemory, the analogous dissociation between preserved first-level and impaired second-level processing has also been found. In a study of children with autism spectrum disorder, the researchers found inaccurate episodic feeling of knowing (but not semantic feeling of knowing) along with no deficit in either recognition or recall[55]. However, there are equivocal findings from other studies: some individuals with autism spectrum disorder show impaired episodic feeling of knowing sensitivity in the context of impaired memory performance[159] and others show no difference in first-order performance alongside impaired metacognitive sensitivity in confidence judgements[160]. A meta-analytic review[161] found a reduction in metacognitive accuracy across tasks but performance was not diminished overall. It has been proposed that a specific impairment in metacognition might result from an inability to cast the self into the past, related to autonoetic consciousness[162], a dysfunction related to self-representation (Box 5). Future research should further investigate access to metacognitive information in cases in which there is no first-order deficit.

Developmental studies

Finally, metacognition and the relationship between first-order and second-order processes have also been explored in development. The potential of domain transferability of metacognition has been of particular interest in relation to school achievement. Development of metacognition is thought to begin as domain-specific and then generalize across domains as children mature[163–165]. For example, children as young as 5 years old were metacognitively accurate on non-verbal emotion and numerical discrimination tasks but metacognition in the two domains was unrelated, lending support to domain-specific metacognition[166]. Similar findings were also reported for two academic domains (arithmetic and spelling); children who were 7–8 years old could make metacognitive judgements in both domains but no correlation was found between the domains. However, a positive cross-task correlation was found at age 8–9 years[167]. Future longitudinal studies could examine how metacognition develops with age considering the relationship between self-concept defined as domain-specific[168] and metacognitive monitoring[169].

To summarize, both clinical and developmental studies present heterogeneous findings, pointing to the importance of control for first-order performance to understand further domain generality in these populations. However, many studies point to the impact of frontal lobe functioning on metacognitive behaviours.

Beliefs and adecisional judgements

So far we have focused on metacognition in decisional judgements. However, metacognitive processes also extend to more generalized beliefs and evaluations across items within a task, and across tasks with domains. Here we review work based on measures of metacognition that belong to beliefs, expertise and information-based metacognition (Table 1). These adecisional judgements are based on metaknowledge rather than experiences because participants have not experienced the task at the moment of the judgement. Adecisional metacognition is not strictly second-order: the evaluation is not derivative of any process in hand, but an estimation of task factors. Global predictions are a notable form of adecisional judgement, especially when an evaluation is made before performing a task. In a typical paradigm, participants judge how many items from an entire study list they will subsequently recall[155,170,171]. However, adecisional judgements have also been extended to other first-order tasks such as short-term memory[172], processing speed and verbal fluency[137] or perceptual tasks[143].

Most studies have used between-subject correlations to examine population-level accuracy of global predictions in different groups[173]. When global predictions are made before and after a study phase, the between-subject correlation between prediction and performance is higher after studying the items, indicating an effect of monitoring[170,173]. However, the modal value of individual-level predictions is found to be tethered to the midpoint of the scale (for instance, predicting recall of six items for a list of twelve[174]), which suggests that to a large extent these judgements are based on generalized beliefs and rules of thumb. It is also possible that the judgement values represent a Gaussian distribution around a central point, but the same participants predict recall of 5 items from a 10-item list and 10 items from a 20-item list[156], suggesting that the judgements are rule-based and not distribution-based.

When item by item and global predictions have been carried out in the same task, mean values of item by item judgements (that is, metacognitive bias) usually correlate with global predictions[170]. Similarly, some authors have used the mean value of item by item judgements to examine between-subject patterns of accuracy. For instance, mean retrospective confidence correlates with performance but, seemingly, more so for general knowledge than episodic memory[175,176]. Some researchers have considered global metacognition as a 'self-rated ability' scale, in which participants rate their performance in comparison with their peers (with options such as '0–9% of people would be worse than me')[177]. This paradigm yields similar findings, with significant correlations between self-rated ability and mean levels of retrospective confidence on general knowledge (sport) questions but not for an episodic task (face recognition)[177]. In applied fields, there was much enthusiasm for the idea that metacognitive accuracy (and self-rated ability) in one domain might help to interpret evaluations in another domain, such as the comparison between global evaluations of general knowledge and eyewitness testimony. However, early research did not find robust relationships across domains and therefore interest in this approach has diminished.

Global metacognition, defined as constructs such as self-efficacy and self-beliefs, has also received attention as a theoretical entity potentially related to daily functioning[178]. In daily life, individuals might draw on a set of beliefs and evaluations that could influence their

Box 5

Anosognosia

Metacognition has long been studied in motor, sensory and cognitive deficits through the lens of self-awareness. For example, motor awareness has been explored since the first proposal of the term anosognosia[229] to describe patients unaware of the existence of their paralysis. Anosognosia is complex: some individuals fail to acknowledge one deficit but recognize another (such as upper but not lower limb paralysis[230]), or fail to adapt their behaviour according to their knowledge (such as admitting their deficits but attempting to walk, or denying deficits but remaining in bed[231,232]). In the sensory domain, Anton–Babinski syndrome is also called visual anosognosia. Individuals with this condition have binocular visual loss but denial of blindness and relatively well-preserved cognition.

Hundreds of scientific papers have been published on anosognosia[233] and the concept has also been used to describe unawareness of cognitive functions or lack of cognitive insight in neurological and psychiatric disorders such as frontal lobe lesions[127], Alzheimer's disease[234], Parkinson's disease[235,236], and schizophrenia[237]. Those studies have revealed a multifaceted view of unawareness across different cognitive domains. For example, individuals with frontal lobe lesions often demonstrate what is called 'utilization behaviour'[238]. These individuals will engage in a stereotypical action in the sight of an object, despite not being explicitly asked to (such as starting to use a stapler that is on a desk), which could be explained by a lack of awareness of goals and intentions[219].

Domain generality or domain specificity of anosognosia has been explored in several theories. The cognitive awareness model[239,240] has been proposed for anosognosia for memory disorders. This model posits the existence of a separate 'metacognitive awareness system' that provides conscious awareness of ability or error. An updated version of this model suggests that domain-specific monitoring processes are situated at a lower level and refer to 'cognitive comparator mechanisms'[241]. The role of the cognitive comparator mechanisms is to compare recent errors with previous experiences in each domain, leading to a global self-representation of one's own abilities. Thus, these theories seem to predict a general awareness system or central supervisory system that would lead to anosognosia across domains if disconnected. By contrast, an impairment of the comparator would lead to domain-specific anosognosia.

mental health[178] or learning[171]. One description of global metacognition as self-performance estimates proposes that global metacognition is 'aggregated' over time from local confidence, emphasizing the role of external feedback[179] (see also ref. 180). Global estimates of confidence and metacognitive bias in local confidence can interact during a task and are correlated[13]. Indeed, some models of self-esteem have proposed a two-way relationship between multiple levels of beliefs[181].

On a theoretical level, the notion of a generalized global belief about function based on an agglomeration of local judgements seems as though it should operate at a domain-general level in metacognition. However, there are too few studies to date to support this view. The field would benefit more from experiments that test the idea that local judgements for one task are extrapolated up into a form of global and domain-general metacognitive awareness for a different kind of task. However, metacognitive bias and self-efficacy might produce a spurious domain generality for adecisional judgements. For instance, if a participant finds one task easy, this assessment might spill over into their belief about other tasks, thinking that the other tasks will be similarly easy. Similarly, a general arousal level across tasks might produce a spurious form of domain generality for adecisional judgements.

For global evaluations, the idea that a predisposition or a rule of thumb guides judgements would imply a form of domain specificity based on knowledge of the task. For example, one could make a domain-specific prediction based on recalling about half the items from a list of words. However, it is hard to see how cross-domain comparisons (such as our opening example of knowing you are better at detecting your brother's face in a crowd than remembering his birthday) could be made without some form of domain-general comparison common currency. This possibility needs more empirical investigation. A further point of interest is to consider the process of aggregation of decisional metacognitive evaluations into a global evaluation that does not pertain to any one decision (and hence is 'adecisional'). Future work should consider whether people can extrapolate across different tasks in different domains to form one generalized belief or sense of expertise.

Summary and future directions

Establishing whether metacognition is domain general is of importance to understand its architecture and to establish its effects in real-world decision-making and clinical conditions. If metacognition were domain general, this would mean that training of metacognition on one domain might mitigate difficulties in metacognition in another cognitive domain.

To create a common conceptual space in which to consider metacognitive processes, we detailed possible domain-specific and domain-general configurations of the metacognitive architecture and emphasized the division into decisional and adecisional forms of metacognition. Perhaps the clearest evidence for domain generality comes from decisional metacognition. We report cross-domain correlations for both metacognitive bias and metacognitive efficiency for such decisional processes, even though the magnitude of the latter seems low. By contrast, neuropsychological studies posit that domain-specific processes are at play in decisional metacognition, pointing towards a weak domain generality or, perhaps, spurious domain generality.

Thus far, the main approach to address domain generality has been to index second-order processes by using retrospective confidence judgements: cross-task correlations yield between-subject correlations for confidence judgements. This approach is an appropriate first step in accessing domain generality, but conclusions are limited by the fact that first-order and adecisional factors can produce spurious correlations across tasks with similar structures and task demands. The use of between-subject correlations can suggest domain-general patterns but is not suitable for hypothesis testing about the architecture of metacognition. If a strong version of a domain-general metacognition is to be found, it will be in mechanisms that apply across tasks of diverse types and structures, such as when reaction times are estimated to build confidence. Because reaction times are present in decision-making across various tasks, finding that confidence is estimated from reaction times might support domain generality, although first-order strategies influencing reaction time can also result in spurious domain generality[118]. In this way, a domain-general architecture would involve a final common pathway for all types of decisional metacognitive evaluations, and the possibility of a common process computed from processing times is an exciting development for the field.

To go beyond correlations between domains, more complex experimental paradigms are needed. One way to assess the domain specificity of a process is to use functional independence. From this perspective, processes are assumed to be independent if a variable has an effect on one process and no effect or the opposite effect on another process. For instance, if X influences Y in the perception domain but does not influence Z in the memory domain, one would assume that Y and Z are independent. Conversely, if participants behave similarly not just across domains but also across all conditions, this would be converging evidence for domain generality. For instance, one study has shown that confidence is shaped by participants' behavioural goal similarly in visual perception and value-based decisions[182]. Another approach is what we can call 'metacognitive transfer' (by analogy to Bayesian transfer, which has been proposed as a test of whether Bayesian decision theory is a good model of human visual perception[183]). Testing metacognitive transfer would consist of identifying some key metacognitive features in two domains (such as perception and memory). If this metacognitive feature generalizes across domains, that would be evidence in favour of domain generality. For instance, a metacognitive feature might be the change in efficiency in evaluating global confidence judgements for two different set sizes. If there is domain generality, measuring metacognitive performance in one modality (such as efficiency change between two set sizes in perception) would predict metacognitive performance in the other modality (such as efficiency change between two set sizes in memory). For both of these methodological advancements, it is clear that metacognitive models that work across domains would be highly beneficial.

A final suggestion for future neuropsychological studies is to test metacognition systematically across domains (not just for the impaired first-order function) and to test metacognition on tasks in which the individuals do not show a first-order deficit (either by manipulating difficulty or by extending the scope of the study to other domains). If metacognitive proficiency relies on a final common pathway, it should be possible to find an individual with impaired metacognitive access across all domains.

In conclusion, we have discussed a common conceptual space for metacognition with a decisional process based on the outputs of first-order processing, which can be common across tasks and domains. Thus far, the evidence favours a weak version of domain-general metacognition, with some processes shared between tasks across different domains. It now remains to test the architectures discussed here across a set of more diverse metacognitive tasks and also at mechanistic, implementational and neural levels.

Published online: 17 November 2023

References

1. Flavell, J. H. Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry. *Am. Psychol.* **34**, 906–911 (1979).
2. Fleur, D. S., Bredeweg, B. & van den Bos, W. Metacognition: ideas and insights from neuro- and educational sciences. *NPJ Sci. Learn.* **6**, 13 (2021).
3. Evans, J. R. & Fisher, R. P. Eyewitness memory: balancing the accuracy, precision and quantity of information through metacognitive monitoring and control. *Appl. Cogn. Psychol.* **25**, 501–508 (2011).
4. Pannu, J. K. & Kaszniak, A. W. Metamemory experiments in neurological populations: a review. *Neuropsychol. Rev.* **15**, 105–130 (2005).
5. Rosenthal, D. M. Consciousness, content, and metacognitive judgments. *Conscious. Cogn.* **9**, 203–214 (2000).
6. Brown, R., Lau, H. & LeDoux, J. E. Understanding the higher-order approach to consciousness. *Trends Cogn. Sci.* **23**, 754–768 (2019).
7. Rouault, M., McWilliams, A., Allen, M. G. & Fleming, S. M. Human metacognition across domains: insights from individual differences and neuroimaging. *Personal. Neurosci.* **1**, e17 (2018).
8. Ais, J., Zylberberg, A., Barttfeld, P. & Sigman, M. Individual consistency in the accuracy and distribution of confidence judgments. *Cognition* **146**, 377–386 (2016).
9. Faivre, N., Filevich, E., Solovey, G., Kühn, S. & Blanke, O. Behavioral, modeling, and electrophysiological evidence for supramodality in human metacognition. *J. Neurosci.* **38**, 263–277 (2018).
10. Lee, A. L. F., Ruby, E., Giles, N. & Lau, H. Cross-domain association in metacognitive efficiency depends on first-order task types. *Front. Psychol.* **9**, 2464 (2018).
11. Mazancieux, A., Fleming, S. M., Souchay, C. & Moulin, C. J. A. Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks. *J. Exp. Psychol. Gen.* **149**, 1788–1799 (2020).
12. Mazancieux, A., Dinze, C., Souchay, C. & Moulin, C. J. A. Metacognitive domain specificity in feeling-of-knowing but not retrospective confidence. *Neurosci. Conscious.* **2020**, niaa001 (2020).
13. McWilliams, A., Bibbey, H., Steinbeis, N., David, A. S. & Fleming, S. M. Age-related decreases in global metacognition are independent of local metacognition and task performance. *Cognition* **235**, 105389 (2022).
14. Mamassian, P. Visual confidence. *Annu. Rev. Vis. Sci.* **2**, 459–481 (2016).
15. Grimaldi, P., Lau, H. & Basso, M. A. There are things that we know that we know, and there are things that we do not know we do not know: confidence in decision-making. *Neurosci. Biobehav. Rev.* **55**, 88–97 (2015).
16. Hoven, M. et al. Abnormalities of confidence in psychiatry: an overview and future perspectives. *Transl. Psychiatry* **9**, 268 (2019).
17. Koriat, A. The feeling of knowing: some metatheoretical implications for consciousness and control. *Conscious. Cogn.* **9**, 149–171 (2000).
18. Koriat, A. in *The Cambridge Handbook of Consciousness* Vol. 981 (ed. Zelazo, P. D.) 289–325 (Cambridge Univ. Press, 2007).
19. Nelson, T. O. & Narens, L. in *Metacognition: Knowing About Knowing* (eds Metcalfe, J. & Shimamura, A. P.) 1–25 (MIT Press, 1994).
20. Fleming, S. M., Dolan, R. J. & Frith, C. D. Metacognition: computation, biology and function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1280–1286 (2012).
21. Koriat, A. The self-consistency model of subjective confidence. *Psychol. Rev.* **119**, 80–113 (2012).
22. Fleming, S. M. & Lau, H. C. How to measure metacognition. *Front. Hum. Neurosci.* **8**, 443 (2014).
23. Marr, D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (W.H. Freeman & Co., 1982).
24. Vaccaro, A. G. & Fleming, S. M. Thinking about thinking: a coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain Neurosci. Adv.* **2**, 2398212818810591 (2018).
25. Fu, Z. et al. The geometry of domain-general performance monitoring in the human medial frontal cortex. *Science* **376**, eabm9922 (2022).
26. Dunlosky, J. & Tauber, S. U. K. *The Oxford Handbook of Metamemory* (Oxford Univ. Press, 2016).
27. Arbuckle, T. Y. & Cuddy, L. L. Discrimination of item strength at time of presentation. *J. Exp. Psychol.* **81**, 126–131 (1969).
28. Rhodes, M. G. & Tauber, S. K. The influence of delaying judgments of learning on metacognitive accuracy: a meta-analytic review. *Psychol. Bull.* **137**, 131–148 (2011).
29. Schacter, D. L. Feeling of knowing in episodic memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **9**, 39 (1983).
30. Hart, J. T. Memory and the feeling-of-knowing experience. *J. Educ. Psychol.* **56**, 208–216 (1965).
31. Koriat, A. How do we know that we know? The accessibility model of the feeling of knowing. *Psychol. Rev.* **100**, 609–639 (1993).
32. Norman, E., Blakstad, O., Johnsen, Ø., Martinsen, S. K. & Price, M. C. The relationship between feelings-of-knowing and partial knowledge for general knowledge questions. *Front. Psychol.* **7**, 996 (2016).
33. Fleming, S. M., Massoni, S., Gajdos, T. & Vergnaud, J.-C. Metacognition about the past and future: quantifying common and distinct influences on prospective and retrospective judgments of self-performance. *Neurosci. Conscious.* **2016**, niw018 (2016).
34. Mei, N., Rankine, S., Olafsson, E. & Soto, D. Similar history biases for distinct prospective decisions of self-performance. *Sci. Rep.* **10**, 5854 (2020).
35. Barrett, A. B., Dienes, Z. & Seth, A. K. Measures of metacognition on signal-detection theoretic models. *Psychol. Methods* **18**, 535–552 (2013).
36. Benjamin, A. S. & Diaz, M. in *Handbook of Memory and Metamemory* (eds Dunlosky, J. & Bjork, R. A.) 73–94 (Psychology Press, 2008).
37. Masson, M. E. J. & Rotello, C. M. Sources of bias in the Goodman–Kruskal γ coefficient measure of association: implications for studies of metacognitive processes. *J. Exp. Psychol. Learn. Mem. Cogn.* **35**, 509–527 (2009).
38. Sherman, M., Barrett, A. B. & Kanai, R. in *Behavioral Methods in Consciousness Research* (ed. Overgaard, M.) 87–106 (Oxford Univ. Press, 2015).
39. Morales, J., Lau, H. & Fleming, S. M. Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *J. Neurosci.* **38**, 3534–3546 (2018).
40. Benwell, C. S. Y., Mohr, G., Wallberg, J., Kouadio, A. & Ince, R. A. A. Psychiatrically relevant signatures of domain-general decision-making and metacognition in the general population. *NPJ Ment. Health Res.* **1**, 1–17 (2022).
41. Lund, A. E., Correa, C., Fardo, F., Fleming, S. & Allen, M. Domain generality in metacognitive ability: a confirmatory study across visual perception, memory, and general knowledge. Preprint at *PsyArXiv* https://doi.org/10.31234/osf.io/p97s5 (2023).
42. Lehmann, M., Hagen, J. & Ettinger, U. Unity and diversity of metacognition. *J. Exp. Psychol. Gen.* **151**, 2396–2417 (2022).
43. Mazor, M., Friston, K. J. & Fleming, S. M. Distinct neural contributions to metacognition for detecting, but not discriminating visual stimuli. *eLife* **9**, e53900 (2020).
44. Fleming, S. M. HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neurosci. Conscious.* **2017**, nix007 (2017).
45. Paulewicz, B. & Blaut, A. The bhsdtr package: a general-purpose method of Bayesian inference for signal detection theory models. *Behav. Res. Methods* **52**, 2122–2141 (2020).
46. Hu, X., Yang, C. & Luo, L. Are the contributions of processing experience and prior beliefs to confidence ratings domain-general or domain-specific? *J. Exp. Psychol. Gen.* **152**, 28–44 (2022).
47. Guggenmos, M. Measuring metacognitive performance: type 1 performance dependence and test–retest reliability. *Neurosci. Conscious.* **2021**, niab040 (2021).
48. Xue, K., Shekhar, M. & Rahnev, D. Examining the robustness of the relationship between metacognitive efficiency and metacognitive bias. *Conscious. Cogn.* **95**, 103196 (2021).
49. de Gardelle, V., Le Corre, F. & Mamassian, P. Confidence as a common currency between vision and audition. *PLoS ONE* **11**, e0147901 (2016).
50. Fleming, S. M., Ryu, J., Golfinos, J. G. & Blackmon, K. E. Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain* **137**, 2811–2822 (2014).
51. Sadeghi, S., Ekhtiari, H., Bahrami, B. & Ahmadabadi, M. N. Metacognitive deficiency in a perceptual but not a memory task in methadone maintenance patients. *Sci. Rep.* **7**, 7052 (2017).
52. Ye, Q., Zou, F., Lau, H., Hu, Y. & Kwok, S. C. Causal evidence for mnemonic metacognition in human precuneus. *J. Neurosci.* **38**, 6379–6387 (2018).
53. Ye, Q. et al. Individual susceptibility to TMS affirms the precuneal role in meta-memory upon recollection. *Brain Struct. Funct.* **224**, 2407–2419 (2019).
54. Beatty, W. W. & Monson, N. Metamemory in multiple sclerosis. *J. Clin. Exp. Neuropsychol.* **13**, 309–327 (1991).
55. Wojcik, D. Z., Moulin, C. J. A. & Souchay, C. Metamemory in children with autism: exploring 'feeling-of-knowing' in episodic and semantic memory. *Neuropsychology* **27**, 19–27 (2013).
56. Gottesman, I. I. & Gould, T. D. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am. J. Psychiatry* **160**, 636–645 (2003).
57. Maniscalco, B., Peters, M. A. K. & Lau, H. Heuristic use of perceptual evidence leads to dissociation between performance and metacognitive sensitivity. *Atten. Percept. Psychophys.* **78**, 923–937 (2016).
58. Jang, Y., Wallsten, T. S. & Huber, D. E. A stochastic detection and retrieval model for the study of metacognition. *Psychol. Rev.* **119**, 186–200 (2012).
59. Fleming, S. M. & Daw, N. D. Self-evaluation of decision-making: a general Bayesian framework for metacognitive computation. *Psychol. Rev.* **124**, 91–114 (2017).
60. Mamassian, P. & de Gardelle, V. Modeling perceptual confidence and the confidence forced-choice paradigm. *Psychol. Rev.* **129**, 976–998 (2022).
61. Balsdon, T., Wyart, V. & Mamassian, P. Confidence controls perceptual evidence accumulation. *Nat. Commun.* **11**, 1753 (2020).
62. Balsdon, T., Mamassian, P. & Wyart, V. Separable neural signatures of confidence during perceptual decisions. *eLife* **10**, e68491 (2021).
63. Shadlen, M. N. & Kiani, R. Decision making as a window on cognition. *Neuron* **80**, 791–806 (2013).
64. Van Zandt, T. ROC curves and confidence judgments in recognition memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **26**, 582–600 (2000).
65. Ratcliff, R. & Starns, J. J. Modeling confidence and response time in recognition memory. *Psychol. Rev.* **116**, 59–83 (2009).
66. Insabato, A., Pannunzi, M., Rolls, E. T. & Deco, G. Confidence-related decision making. *J. Neurophysiol.* **104**, 539–547 (2010).
67. Wei, Z. & Wang, X.-J. Confidence estimation as a stochastic process in a neurodynamical system of decision making. *J. Neurophysiol.* **114**, 99–113 (2015).
68. Maniscalco, B. et al. Tuned inhibition in perceptual decision-making circuits can explain seemingly suboptimal confidence behavior. *PLoS Comput. Biol.* **17**, e1008779 (2021).

69. Atiya, N. A. A., Huys, Q. J. M., Dolan, R. J. & Fleming, S. M. Explaining distortions in metacognition with an attractor network model of decision uncertainty. *PLoS Comput. Biol.* **17**, e1009201 (2021).
70. Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. Diffusion decision model: current issues and history. *Trends Cogn. Sci.* **20**, 260–281 (2016).
71. Peters, M. A. & Lau, H. Human observers have optimal introspective access to perceptual processes even for visually masked stimuli. *eLife* **5**, e09651 (2015).
72. Churchland, A. K., Kiani, R. & Shadlen, M. N. Decision-making with multiple alternatives. *Nat. Neurosci.* **11**, 693–702 (2008).
73. Yonelinas, A. P. The nature of recollection and familiarity: a review of 30 years of research. *J. Mem. Lang.* **46**, 441–517 (2002).
74. Wixted, J. T. Dual-process theory and signal-detection theory of recognition memory. *Psychol. Rev.* **114**, 152–176 (2007).
75. Treisman, M. & Faulkner, A. The setting and maintenance of criteria representing levels of confidence. *J. Exp. Psychol. Hum. Percept. Perform.* **10**, 119–139 (1984).
76. Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231 (2008).
77. Maniscalco, B. & Lau, H. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Conscious. Cogn.* **21**, 422–430 (2012).
78. Maniscalco, B. & Lau, H. The signal processing architecture underlying subjective reports of sensory awareness. *Neurosci. Conscious.* **2016**, niw002 (2016).
79. Shekhar, M. & Rahnev, D. The nature of metacognitive inefficiency in perceptual decision making. *Psychol. Rev.* **128**, 45–70 (2021).
80. Pouget, A., Drugowitsch, J. & Kepecs, A. Confidence and certainty: distinct probabilistic quantities for different goals. *Nat. Neurosci.* **19**, 366–374 (2016).
81. Kiani, R. & Shadlen, M. N. Representation of confidence associated with a decision by neurons in the parietal cortex. *Science* **324**, 759–764 (2009).
82. Meyniel, F., Sigman, M. & Mainen, Z. F. Confidence as Bayesian probability: from neural origins to behavior. *Neuron* **88**, 78–92 (2015).
83. Sherman, M. T., Seth, A. K., Barrett, A. B. & Kanai, R. Prior expectations facilitate metacognition for perceptual decision. *Conscious. Cogn.* **35**, 53–65 (2015).
84. Constant, M., Salomon, R. & Filevich, E. Judgments of agency are affected by sensory noise without recruiting metacognitive processing. *eLife* **11**, e72356 (2022).
85. Hu, X. et al. A Bayesian inference model for metamemory. *Psychol. Rev.* **128**, 824–855 (2021).
86. Landry, M., Da Silva Castanheira, J., Sackur, J. & Raz, A. Investigating how the modularity of visuospatial attention shapes conscious perception using type I and type II signal detection theory. *J. Exp. Psychol. Hum. Percept. Perform.* **47**, 402–422 (2021).
87. Arbuzova, P. et al. Measuring metacognition of direct and indirect parameters of voluntary movement. *J. Exp. Psychol. Gen.* **150**, 2208–2229 (2021).
88. Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J. & Rees, G. Relating introspective accuracy to individual differences in brain structure. *Science* **329**, 1541–1543 (2010).
89. Koizumi, A., Maniscalco, B. & Lau, H. Does perceptual confidence facilitate cognitive control? *Atten. Percept. Psychophys.* **77**, 1295–1306 (2015).
90. Samaha, J., Barrett, J. J., Sheldon, A. D., LaRocque, J. J. & Postle, B. R. Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. *Front. Psychol.* **7**, 851 (2016).
91. Zylberberg, A., Fetsch, C. R. & Shadlen, M. N. The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *eLife* **5**, e17688 (2016).
92. Yon, D., Zainzinger, V., de Lange, F. P., Eimer, M. & Press, C. Action biases perceptual decisions toward expected outcomes. *J. Exp. Psychol. Gen.* **150**, 1225–1236 (2021).
93. Constant, M., Pereira, M., Faivre, N. & Filevich, E. Prior information differentially affects discrimination decisions and subjective confidence reports. *Nat. Commun.* **14**, 5473 (2023).
94. Gajdos, T., Fleming, S. M., Saez Garcia, M., Weindel, G. & Davranche, K. Revealing subthreshold motor contributions to perceptual confidence. *Neurosci. Conscious.* **2019**, niz001 (2019).
95. Faivre, N. et al. Sensorimotor conflicts alter metacognitive and action monitoring. *Cortex* **124**, 224–234 (2020).
96. Filevich, E., Koß, C. & Faivre, N. Response-related signals increase confidence but not metacognitive performance. *eNeuro* **7**, https://doi.org/10.1523/ENEURO.0326-19.2020 (2020).
97. Dotan, D., Pinheiro-Chagas, P., Al Roumi, F. & Dehaene, S. Track it to crack it: dissecting processing stages with finger tracking. *Trends Cogn. Sci.* **23**, 1058–1070 (2019).
98. Siedlecka, M., Paulewicz, B. & Wierzchoń, M. But I was so sure! Metacognitive judgments are less accurate given prospectively than retrospectively. *Front. Psychol.* **7**, 218 (2016).
99. Fleming, S. M. et al. Action-specific disruption of perceptual confidence. *Psychol. Sci.* **26**, 89–98 (2015).
100. Patel, D., Fleming, S. M. & Kilner, J. M. Inferring subjective states through the observation of actions. *Proc. R. Soc. Lond. B: Biol. Sci.* **279**, 4853–4860 (2012).
101. Vuillaume, L., Martin, J.-R., Sackur, J. & Cleeremans, A. Comparing self- and hetero-metacognition in the absence of verbal communication. *PLoS ONE* **15**, e0231530 (2020).
102. Pereira, M. et al. Disentangling the origins of confidence in speeded perceptual judgments through multimodal imaging. *Proc. Natl Acad. Sci. USA* **117**, 8382–8390 (2020).
103. Hanczakowski, M., Butowska, E., Philip Beaman, C., Jones, D. M. & Zawadzka, K. The dissociations of confidence from accuracy in forced-choice recognition judgments. *J. Mem. Lang.* **117**, 104189 (2021).
104. Rhodes, M. G. & Castel, A. D. Memory predictions are influenced by perceptual information: evidence for metacognitive illusions. *J. Exp. Psychol. Gen.* **137**, 615–625 (2008).
105. Rhodes, M. G. & Castel, A. D. Metacognitive illusions for auditory information: effects on monitoring and control. *Psychon. Bull. Rev.* **16**, 550–554 (2009).
106. Koriat, A. & Bjork, R. A. Mending metacognitive illusions: a comparison of mnemonic-based and theory-based procedures. *J. Exp. Psychol. Learn. Mem. Cogn.* **32**, 1133–1145 (2006).
107. Yue, C. L., Castel, A. D. & Bjork, R. A. When disfluency is—and is not—a desirable difficulty: the influence of typeface clarity on metacognitive judgments and memory. *Mem. Cognit.* **41**, 229–241 (2013).
108. Besken, M. & Mulligan, N. W. Perceptual fluency, auditory generation, and metamemory: analyzing the perceptual fluency hypothesis in the auditory modality. *J. Exp. Psychol. Learn. Mem. Cogn.* **40**, 429–440 (2014).
109. Rahnev, D. A., Maniscalco, B., Luber, B., Lau, H. & Lisanby, S. H. Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence. *J. Neurophysiol.* **107**, 1556–1563 (2012).
110. Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E. & Lau, H. Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cogn. Neurosci.* **1**, 165–175 (2010).
111. Chiang, T.-C., Lu, R.-B., Hsieh, S., Chang, Y.-H. & Yang, Y.-K. Stimulation in the dorsolateral prefrontal cortex changes subjective evaluation of percepts. *PLoS ONE* **9**, e106943 (2014).
112. Bor, D., Schwartzman, D. J., Barrett, A. B. & Seth, A. K. Theta-burst transcranial magnetic stimulation to the prefrontal or parietal cortex does not impair metacognitive visual awareness. *PLoS ONE* **12**, e0171793 (2017).
113. Gogulski, J., Zetter, R., Nyrhinen, M., Pertovaara, A. & Carlson, S. Neural substrate for metacognitive accuracy of tactile working memory. *Cereb. Cortex* **27**, 5343–5352 (2017).
114. Baranski, J. V. & Petrusic, W. M. Probing the locus of confidence judgments: experiments on the time to determine confidence. *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 929–945 (1998).
115. Rahnev, D. et al. The confidence database. *Nat. Hum. Behav.* **4**, 317–325 (2020).
116. Fleming, S. M., van der Putten, E. J. & Daw, N. D. Neural mediators of changes of mind about perceptual decisions. *Nat. Neurosci.* **21**, 617–624 (2018).
117. Kiani, R., Corthell, L. & Shadlen, M. N. Choice certainty is informed by both evidence and decision time. *Neuron* **84**, 1329–1342 (2014).
118. Desender, K., Vermeylen, L. & Verguts, T. Dynamic influences on static measures of metacognition. *Nat. Commun.* **13**, 4208 (2022).
119. Koriat, A. Monitoring one's own knowledge during study: a cue-utilization approach to judgments of learning. *J. Exp. Psychol. Gen.* **126**, 349–370 (1997).
120. Metcalfe, J., Schwartz, B. L. & Joaquim, S. G. The cue-familiarity heuristic in metacognition. *J. Exp. Psychol. Learn. Mem. Cogn.* **19**, 851–861 (1993).
121. Alter, A. L. & Oppenheimer, D. M. Uniting the tribes of fluency to form a metacognitive nation. *Pers. Soc. Psychol. Rev.* **13**, 219–235 (2009).
122. Oppenheimer, D. M. The secret life of fluency. *Trends Cogn. Sci.* **12**, 237–241 (2008).
123. Benjamin, A. S. & Bjork, R. A. in *Implicit Memory and Metacognition* (ed. Reder, L. M.) 321–350 (Psychology Press, 2014).
124. Kelley, C. M. & Lindsay, D. S. Remembering mistaken for knowing: ease of retrieval as a basis for confidence in answers to general knowledge questions. *J. Mem. Lang.* **32**, 1–24 (1993).
125. Rausch, M. & Zehetleitner, M. Visibility is not equivalent to confidence in a low contrast orientation discrimination task. *Front. Psychol.* **7**, 591 (2016).
126. Shekhar, M. & Rahnev, D. Sources of metacognitive inefficiency. *Trends Cogn. Sci.* **25**, 12–23 (2021).
127. Janowsky, J. S., Shimamura, A. P. & Squire, L. R. Memory and metamemory: comparisons between patients with frontal lobe lesions and amnesic patients. *Psychobiology* **17**, 3–11 (1989).
128. Shimamura, A. P. & Squire, L. R. Memory and metamemory: a study of the feeling-of-knowing phenomenon in amnesic patients. *J. Exp. Psychol. Learn. Mem. Cogn.* **12**, 452–460 (1986).
129. Schnyer, D. M. et al. A role for right medial prefrontal cortex in accurate feeling-of-knowing judgments: evidence from patients with lesions to frontal cortex. *Neuropsychologia* **42**, 957–966 (2004).
130. Pinon, K., Allain, P., Kefi, M. Z., Dubas, F. & Le Gall, D. Monitoring processes and metamemory experience in patients with dysexecutive syndrome. *Brain Cogn.* **57**, 185–188 (2005).
131. Bäckman, L. & Lipinska, B. Monitoring of general knowledge: evidence for preservation in early Alzheimer's disease. *Neuropsychologia* **31**, 335–345 (1993).
132. Lipinska, B. & Bäckman, L. Feeling-of-knowing in fact retrieval: further evidence for preservation in early Alzheimer's disease. *J. Int. Neuropsychol. Soc.* **2**, 350–358 (1996).
133. Moulin, C. J., Perfect, T. J. & Jones, R. W. Evidence for intact memory monitoring in Alzheimer's disease: metamemory sensitivity at encoding. *Neuropsychologia* **38**, 1242–1250 (2000).
134. Moulin, C. J. A., James, N., Perfect, T. J. & Jones, R. W. Knowing what you cannot recognise: further evidence for intact metacognition in Alzheimer's disease. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* **10**, 74–82 (2003).

135. Pappas, B. A. et al. Alzheimer's disease and feeling-of-knowing for knowledge and episodic memory. *J. Gerontol.* **47**, P159–P164 (1992).
136. Souchay, C., Isingrini, M., Pillon, B. & Gil, R. Metamemory accuracy in Alzheimer's disease and frontotemporal lobe dementia. *Neurocase* **9**, 482–492 (2003).
137. Mazancieux, A., Moulin, C. J. A., Casez, O. & Souchay, C. A multidimensional assessment of metacognition across domains in multiple sclerosis. *J. Int. Neuropsychol. Soc.* **27**, 124–135 (2021).
138. Souchay, C., Isingrini, M. & Gil, R. Metamemory monitoring and Parkinson's disease. *J. Clin. Exp. Neuropsychol.* **28**, 618–630 (2006).
139. Souchay, C. & Smith, S. J. Subjective states associated with retrieval failures in Parkinson's disease. *Conscious. Cogn.* **22**, 795–805 (2013).
140. Bhome, R. et al. Metacognition in functional cognitive disorder. *Brain Commun.* **4**, fcac041 (2022).
141. Demant, K. M., Vinberg, M., Kessing, L. V. & Miskowiak, K. W. Assessment of subjective and objective cognitive function in bipolar disorder: correlations, predictors and the relation to psychosocial function. *Psychiatry Res.* **229**, 565–571 (2015).
142. Hauser, T. U., Allen, M., NSPN Consortium, Rees, G. & Dolan, R. J. Publisher Correction: Metacognitive impairments extend perceptual decision making weaknesses in compulsivity. *Sci. Rep.* **8**, 6046 (2018).
143. Rouault, M., Seow, T., Gillan, C. M. & Fleming, S. M. Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. *Biol. Psychiatry* **84**, 443–451 (2018).
144. Hoven, M. et al. Metacognition and the effect of incentive motivation in two compulsive disorders: gambling disorder and obsessive-compulsive disorder. *Psychiatry Clin. Neurosci.* **76**, 437–449 (2022).
145. Moritz, S. et al. Sowing the seeds of doubt: a narrative review on metacognitive training in schizophrenia. *Clin. Psychol. Rev.* **34**, 358–366 (2014).
146. Dietrichkeit, M., Grzella, K., Nagel, M. & Moritz, S. Using virtual reality to explore differences in memory biases and cognitive insight in people with psychosis and healthy controls. *Psychiatry Res.* **285**, 112787 (2020).
147. Gawęda, Ł. & Moritz, S. The role of expectancies and emotional load in false auditory perceptions among patients with schizophrenia spectrum disorders. *Eur. Arch. Psychiatry Clin. Neurosci.* **271**, 713–722 (2021).
148. Moritz, S., Woodward, T. S. & Rodriguez-Raecke, R. Patients with schizophrenia do not produce more false memories than controls but are more confident in them. *Psychol. Med.* **36**, 659–667 (2006).
149. Berna, F., Zou, F., Danion, J.-M. & Kwok, S. C. Overconfidence in false autobiographical memories in patients with schizophrenia. *Psychiatry Res.* **279**, 374–375 (2019).
150. Rouy, M. et al. Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders. *Neurosci. Biobehav. Rev.* **126**, 329–337 (2021).
151. Ansell, E. L. & Bucks, R. S. Mnemonic anosognosia in Alzheimer's disease: a test of Agnew and Morris (1998). *Neuropsychologia* **44**, 1095–1102 (2006).
152. Moulin, C. J. A. et al. Retrieval-induced forgetting in Alzheimer's disease. *Neuropsychologia* **40**, 862–867 (2002).
153. Budson, A. E. et al. Memory and emotions for the September 11, 2001, terrorist attacks in patients with Alzheimer's disease, patients with mild cognitive impairment, and healthy older adults. *Neuropsychology* **18**, 315–327 (2004).
154. Barrett, A. M., Eslinger, P. J., Ballentine, N. H. & Heilman, K. M. Unawareness of cognitive deficit (cognitive anosognosia) in probable AD and control subjects. *Neurology* **64**, 693–699 (2005).
155. Moulin, C. J. A., Perfect, T. J. & Jones, R. W. Global predictions of memory in Alzheimer's disease: evidence for preserved metamemory monitoring. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* **7**, 230–244 (2000).
156. Moulin, C. & de La Rochefoucauld, F. D. in *Applied Metacognition* (eds. Perfect, T. J. & Schwartz, B. L.) 197–223 (Cambridge Univ. Press, 2002).
157. Ko, Y. & Lau, H. A detection theoretic explanation of blindsight suggests a link between conscious perception and metacognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1401–1411 (2012).
158. Balsdon, T. & Azzopardi, P. Absolute and relative blindsight. *Conscious. Cogn.* **32**, 79–91 (2015).
159. Grainger, C., Williams, D. M. & Lind, S. E. Metacognition, metamemory, and mindreading in high-functioning adults with autism spectrum disorder. *J. Abnorm. Psychol.* **123**, 650–659 (2014).
160. Grainger, C., Williams, D. M. & Lind, S. E. Metacognitive monitoring and control processes in children with autism spectrum disorder: diminished judgement of confidence accuracy. *Conscious. Cogn.* **42**, 65–74 (2016).
161. Carpenter, K. L. & Williams, D. M. A meta-analysis and critical review of metacognitive accuracy in autism. *Autism* **27**, 512–525 (2022).
162. Souchay, C., Guillery-Girard, B., Pauly-Takacs, K., Wojcik, D. Z. & Eustache, F. Subjective experience of episodic memory and metacognition: a neurodevelopmental approach. *Front. Behav. Neurosci.* **7**, 212 (2013).
163. Lyons, K. E. & Ghetti, S. in *Trends and Prospects in Metacognition Research* (eds Efklides, A. & Misailidi, P.) 259–278 (Springer, 2010).
164. Veenman, M. V. J. & Spaans, M. A. Relation between intellectual and metacognitive skills: age and task differences. *Learn. Individ. Differ.* **15**, 159–176 (2005).
165. Geurten, M., Meulemans, T. & Lemaire, P. From domain-specific to domain-general? The developmental path of metacognition for strategy selection. *Cogn. Dev.* **48**, 62–81 (2018).
166. Vo, V. A., Li, R., Kornell, N., Pouget, A. & Cantlon, J. F. Young children bet on their numerical skills: metacognition in the numerical domain. *Psychol. Sci.* **25**, 1712–1721 (2014).
167. Bellon, E., Fias, W. & De Smedt, B. Metacognition across domains: is the association between arithmetic and metacognitive monitoring domain-specific? *PLoS ONE* **15**, e0229932 (2020).
168. Shavelson, R. J. & Hubner, J. J. Self-concept: validation of construct interpretations. *Rev. Educ. Res.* **46**, 407–441 (1976).
169. Dapp, L. C. & Roebers, C. M. Metacognition and self-concept: elaborating on a construct relation in first-grade children. *PLoS ONE* **16**, e0250845 (2021).
170. Connor, L. T., Dunlosky, J. & Hertzog, C. Age-related differences in absolute but not relative metamemory accuracy. *Psychol. Aging* **12**, 50–71 (1997).
171. Händel, M., de Bruin, A. B. H. & Dresel, M. Individual differences in local and global metacognitive judgments. *Metacognition Learn.* **15**, 51–75 (2020).
172. Bertrand, J. M. et al. In the here and now: short term memory predictions are preserved in Alzheimer's disease. *Cortex* **119**, 158–164 (2019).
173. Silva, A. R., Pinho, M. S., Macedo, L., Souchay, C. & Moulin, C. Mnemonic anosognosia in Alzheimer's disease is caused by a failure to transfer online evaluations of performance: evidence from memory training programs. *J. Clin. Exp. Neuropsychol.* **39**, 419–433 (2017).
174. Hertzog, C., Saylor, L. L., Fleece, A. M. & Dixon, R. A. Metamemory and aging: relations between predicted, actual and perceived memory task performance. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* **1**, 203–237 (1994).
175. Luna, K. & Martín-Luengo, B. Confidence–accuracy calibration with general knowledge and eyewitness memory cued recall questions. *Appl. Cogn. Psychol.* **26**, 289–295 (2012).
176. Perfect, T. J. & Hollins, T. S. Predictive feeling of knowing judgements and postdictive confidence judgements in eyewitness memory and general knowledge. *Appl. Cogn. Psychol.* **10**, 371–382 (1996).
177. Perfect, T. J. The role of self-rated ability in the accuracy of confidence judgements in eyewitness memory and general knowledge. *Appl. Cogn. Psychol.* **18**, 157–168 (2004).
178. Seow, T. X. F., Rouault, M., Gillan, C. M. & Fleming, S. M. How local and global metacognition shape mental health. *Biol. Psychiatry* **90**, 436–446 (2021).
179. Rouault, M., Dayan, P. & Fleming, S. M. Forming global estimates of self-performance from local confidence. *Nat. Commun.* **10**, 1141 (2019).
180. Lee, A. L. F., de Gardelle, V. & Mamassian, P. Global visual confidence. *Psychon. Bull. Rev.* **28**, 1233–1242 (2021).
181. De Ruiter, N. M. P., Van Geert, P. L. C. & Kunnen, E. S. Explaining the 'How' of self-esteem development: the self-organizing self-esteem model. *Rev. Gen. Psychol.* **21**, 49–68 (2017).
182. Sepulveda, P. et al. Visual attention modulates the integration of goal-relevant evidence and not value. *eLife* **9**, e60705 (2020).
183. Maloney, L. T. & Mamassian, P. Bayesian decision theory as a model of human visual perception: testing Bayesian transfer. *Vis. Neurosci.* **26**, 147–155 (2009).
184. King, J.-R. & Dehaene, S. A model of subjective report and objective discrimination as categorical decisions in a vast representational space. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **369**, 20130204 (2014).
185. Pleskac, T. J. & Busemeyer, J. R. Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychol. Rev.* **117**, 864–901 (2010).
186. Desender, K., Donner, T. H. & Verguts, T. Dynamic expressions of confidence within an evidence accumulation framework. *Cognition* **207**, 104522 (2021).
187. Vickers, D. Uncertainty, choice, and the marginal efficiencies. *J. Post. Keynes. Econ.* **2**, 240–254 (1979).
188. van den Berg, R. et al. A common mechanism underlies changes of mind about decisions and confidence. *eLife* **5**, e12192 (2016).
189. Zylberberg, A., Barttfeld, P. & Sigman, M. The construction of confidence in a perceptual decision. *Front. Integr. Neurosci.* **6**, 79 (2012).
190. Rahnev, D., Nee, D. E., Riddle, J., Larson, A. S. & D'Esposito, M. Causal evidence for frontal cortex organization for perceptual decision making. *Proc. Natl Acad. Sci. USA* **113**, 6059–6064 (2016).
191. Morphew, J. W. Changes in metacognitive monitoring accuracy in an introductory physics course. *Metacogn. Learn.* **16**, 89–111 (2021).
192. Nelson, T. O. in *Psychology of Learning and Motivation* Vol. 26 (ed. Bower, G. H.) 125–173 (Academic, 1990).
193. Mazzoni, G., Cornoldi, C., Tomat, L. & Vecchi, T. Remembering the grocery shopping list: a study on metacognitive biases. *Appl. Cogn. Psychol.* **11**, 253–267 (1997).
194. McCarley, J. S. & Gosney, J. Metacognitive judgments in a simulated luggage screening task. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.* **49**, 1620–1624 (2005).
195. Son, L. K. & Metcalfe, J. Metacognitive and control strategies in study-time allocation. *J. Exp. Psychol. Learn. Mem. Cogn.* **26**, 204–221 (2000).
196. Jemstedt, A., Kubik, V. & Jönsson, F. U. What moderates the accuracy of ease of learning judgments? *Metacognition Learn.* **12**, 337–355 (2017).
197. Luna, K., Martín-Luengo, B. & Albuquerque, P. B. Do delayed judgements of learning reduce metamemory illusions? A meta-analysis. *Q. J. Exp. Psychol.* **71**, 1626–1636 (2018).
198. Metcalfe, J. & Finn, B. Evidence that judgments of learning are causally related to study choice. *Psychon. Bull. Rev.* **15**, 174–179 (2008).
199. Chang, M. & Brainerd, C. J. Association and dissociation between judgments of learning and memory: a meta-analysis of the font size effect. *Metacogn Learn.* **17**, 443–476 (2022).
200. Souchay, C., Isingrini, M. & Espagnet, L. Aging, episodic memory feeling-of-knowing, and frontal functioning. *Neuropsychology* **14**, 299–309 (2000).

201. Gruneberg, M. M. & Monks, J. 'Feeling of knowing' and cued recall. *Acta Psychol.* **38**, 257–265 (1974).
202. Kelemen, W. L., Frost, P. J. & Weaver, C. A. 3rd Individual differences in metacognition: evidence against a general metacognitive ability. *Mem. Cognit.* **28**, 92–107 (2000).
203. Meyniel, F., Schlunegger, D. & Dehaene, S. The sense of confidence during probabilistic learning: a normative account. *PLoS Comput. Biol.* **11**, e1004305 (2015).
204. Barthelmé, S. & Mamassian, P. Evaluation of objective uncertainty in the visual system. *PLoS Comput. Biol.* **5**, e1000504 (2009).
205. de Gardelle, V. & Mamassian, P. Does confidence use a common currency across two visual tasks? *Psychol. Sci.* **25**, 1286–1288 (2014).
206. Persaud, N., McLeod, P. & Cowey, A. Post-decision wagering objectively measures awareness. *Nat. Neurosci.* **10**, 257–261 (2007).
207. Schurger, A. & Sher, S. Awareness, loss aversion, and post-decision wagering. *Trends Cognit. Sci.* **12**, 209–210; author reply 210 (2008).
208. Seth, A. K. Post-decision wagering measures metacognitive content, not sensory consciousness. *Conscious. Cogn.* **17**, 981–983 (2008).
209. Dienes, Z. & Seth, A. Gambling on the unconscious: a comparison of wagering and confidence ratings as measures of awareness in an artificial grammar task. *Conscious. Cogn.* **19**, 674–681 (2010).
210. Massoni, S., Gajdos, T. & Vergnaud, J.-C. Confidence measurement in the light of signal detection theory. *Front. Psychol.* **5**, 1455 (2014).
211. Goupil, L. & Kouider, S. Behavioral and neural indices of metacognitive sensitivity in preverbal infants. *Curr. Biol.* **26**, 3038–3045 (2016).
212. Cai, Y. et al. Time-sensitive prefrontal involvement in associating confidence with task performance illustrates metacognitive introspection in monkeys. *Commun. Biol.* **5**, 799 (2022).
213. Le Pelley, M. E. Metacognitive monkeys or associative animals? Simple reinforcement learning explains uncertainty in nonhuman animals. *J. Exp. Psychol. Learn. Mem. Cogn.* **38**, 686–708 (2012).
214. Beran, M. Animal metacognition: a decade of progress, problems, and the development of new prospects. *Anim. Behav. Cogn.* **6**, 223–229 (2019).
215. Proust, J. P. From comparative studies to interdisciplinary research on metacognition. *Anim. Behav. Cogn.* **6**, 309–328 (2019).
216. Dunlosky, J. & Ariel, R. in *Advances in Research and Theory* Vol. 54 (ed. Ross, B. H.) 103–140 (Elsevier, 2011).
217. Finley, J. R., Tullis, J. G. & Benjamin, A. S. in *New Science of Learning: Cognition, Computers and Collaboration in Education* (eds Khine, M. S. & Saleh, I. M.) 109–131 (Springer, 2010).
218. Fourneret, P. & Jeannerod, M. Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologia* **36**, 1133–1140 (1998).
219. Blakemore, S. J., Wolpert, D. M. & Frith, C. D. Abnormalities in the awareness of action. *Trends Cogn. Sci.* **6**, 237–242 (2002).
220. Custers, R. & Aarts, H. The unconscious will: how the pursuit of goals operates outside of conscious awareness. *Science* **329**, 47–50 (2010).
221. Locke, S. M., Mamassian, P. & Landy, M. S. Performance monitoring for sensorimotor confidence: a visuomotor tracking study. *Cognition* **205**, 104396 (2020).
222. Sinanaj, I., Cojan, Y. & Vuilleumier, P. Inter-individual variability in metacognitive ability for visuomotor performance and underlying brain structures. *Conscious. Cogn.* **36**, 327–337 (2015).
223. Pereira, M. et al. Preserved metacognition for undetected visuomotor deviations. *J. Neurosci.* **43**, 6176–6184 (2023).
224. Jovanovic, L. & López-Moliner, J. Contrasting contributions of movement onset and duration to self-evaluation of sensorimotor timing performance. *Eur. J. NeuroSci.* **54**, 5092–5111 (2021).
225. Charles, L., Chardin, C. & Haggard, P. Evidence for metacognitive bias in perception of voluntary action. *Cognition* **194**, 104041 (2020).
226. Nelson, T. O. A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychol. Bull.* **95**, 109–133 (1984).
227. Maniscalco, B. & Lau, H. in *The Cognitive Neuroscience of Metacognition* (eds Fleming, S. M. & Frith, C. D.) 25–66 (Springer, 2014).
228. Guggenmos, M. Reverse engineering of metacognition. *eLife* **11**, e75420 (2022).
229. Babinski, J. Contribution a l'étude des troubles mentaux dans l'hémiplégie organique (anosognosie) [French]. *Rev. Neurol.* **27**, 845–848 (1914).
230. Berti, A., Làdavas, E. & Della Corte, M. Anosognosia for hemiplegia, neglect dyslexia, and drawing neglect: clinical findings and theoretical considerations. *J. Int. Neuropsychol. Soc.* **2**, 426–440 (1996).
231. Jehkonen, M., Laihosalo, M. & Kettunen, J. E. Impact of neglect on functional outcome after stroke—a review of methodological issues and recent research findings. *Restor. Neurol. Neurosci.* **24**, 209–215 (2006).
232. Nimmo-Smith, I., Marcel, A. J. & Tegnér, R. A diagnostic test of unawareness of bilateral motor task abilities in anosognosia for hemiplegia. *J. Neurol. Neurosurg. Psychiatry* **76**, 1167–1169 (2005).
233. Jenkinson, P. M. & Fotopoulou, A. Understanding Babinski's anosognosia: 100 years later. *Cortex* **61**, 1–4 (2014).
234. Souchay, C. Metamemory in Alzheimer's disease. *Cortex* **43**, 987–1003 (2007).
235. Amanzio, M. et al. Impaired awareness of movement disorders in Parkinson's disease. *Brain Cogn.* **72**, 337–346 (2010).
236. Jenkinson, P. M., Edelstyn, N. M. J., Stephens, R. & Ellis, S. J. Why are some Parkinson disease patients unaware of their dyskinesias? *Cogn. Behav. Neurol.* **22**, 117–121 (2009).
237. Riggs, S. E., Grant, P. M., Perivoliotis, D. & Beck, A. T. Assessment of cognitive insight: a qualitative review. *Schizophr. Bull.* **38**, 338–350 (2012).
238. Lhermitte, F. 'Utilization behaviour'and its relation to lesions of the frontal lobes. *Brain* **106**, 237–255 (1983).
239. Agnew, S. K. & Morris, R. G. The heterogeneity of anosognosia for memory impairment in Alzheimer's disease: a review of the literature and a proposed model. *Aging Ment. Health* **2**, 7–19 (1998).
240. Morris, R. G. & Hannesdottir, K. in *Cognitive Neuropsychology of Alzheimer's Disease* (eds Morris, R. & Becker, J.) 275–296 (Oxford Univ. Press, 2004).
241. Clare, L., Marková, I. S., Roth, I. & Morris, R. G. Awareness in Alzheimer's disease and associated dementias: theoretical framework and clinical implications. *Aging Ment. Health* **15**, 936–944 (2011).

Acknowledgements

The evidence for this Review began to accumulate at a workshop, supported by the Institut Universitaire de France fellowship awarded to C.J.A.M., held in Grenoble in July 2019 that all of the authors attended.

Author contributions

All authors contributed substantially to discussion of the content and the review of the literature. A.M. and M.P. wrote the main body of the article. The remaining authors wrote various subsections and boxes. All authors reviewed and edited the manuscript before submission.

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature Reviews Psychology* thanks Minyu Chang, Dobromir Rahnev and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.