Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

HHS Public Access
Author manuscript
Am Psychol. Author manuscript; available in PMC 2022 December 01.

Published in final edited form as:
Am Psychol. 2021 December ; 76(9): 1445–1453. doi:10.1037/amp0000937.

Visual metacognition: Measures, models and neural correlates

Dobromir Rahnev
School of Psychology, Georgia Institute of Technology, Atlanta, GA

Abstract

Visual metacognition is the ability to evaluate one's performance on visual perceptual tasks. The field of visual metacognition unites the long tradition of visual psychophysics with the younger field of metacognition research. This article traces the historical roots of the field and reviews progress in the areas of (i) constructing appropriate measures of metacognitive ability, (ii) developing computational models, and (iii) revealing the neural correlates of visual metacognition. First, I review the most popular measures of metacognitive ability with an emphasis on their psychophysical properties. Second, I examine the empirical targets for modeling, the dominant modeling frameworks and the assumed computations underlying visual metacognition. Third, I explore the progress on understanding the neural correlates of visual metacognition by focusing on anatomical and functional studies, as well as causal manipulations. What emerges is a picture of substantial progress on constructing measures, developing models and revealing the neural correlates of metacognition, but very little integration between these three areas of inquiry. I then explore the deep, intrinsic links between the three areas of research and argue that continued progress requires the recognition and exploitation of these links. Throughout, I discuss the implications of progress in visual metacognition for other areas of metacognition research, and pinpoint specific advancements that could be adopted by researchers working in other subfields of metacognition.

Keywords

visual metacognition; metacognitive ability; confidence generation; perceptual decision making

Historical overview of visual metacognition

Visual metacognition is the ability to evaluate one's performance on visual perceptual tasks. It is a critical skill in our everyday life that allow us, for example, to recognize our poor ability to see in foggy conditions and thus drive slower. More generally, our visual metacognition allows us to know whether to commit to a decision or either seek advice from others (Pescetelli & Yeung, 2021) or gather more information ourselves (Desender et al., 2018). The field has dual roots in research on visual perception and metacognition. Both fields can be tracked all the way antiquity and were topics already examined by Aristotle.

Correspondence Dobromir Rahnev, Georgia Institute of Technology, 654 Cherry Str. NW, Atlanta, GA 30332, rahnev@psych.gatech.edu.

Competing interests
None

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The modern research on visual perception dates back to the mid-19th century with pioneering work by Fechner, Helmholtz, and others who cast perception as a process of inference and laid the foundation of psychophysics research (Fechner, 1860; Helmholtz, 1856). Many of the early developments – such as the graded manipulation of visual stimuli and the careful measurement of individual performance – remain important to this day. However, the early research used metacognitive measures of evaluation, such as confidence ratings, as a means to understand vision rather than as a topic of research in their own right (Green & Swets, 1966; Peirce & Jastrow, 1884).

The modern research on metacognition can be traced back at least to the middle of the 20th century (Clarke et al., 1959; Hart, 1965). It was exactly 50 years ago when John H. Flavell coined the term “metamemory” (Flavell, 1971). Flavell later popularized the term “metacognition” (Flavell, 1979), which he defined as “knowledge and cognition about cognitive phenomena.” The early research on metacognition was mostly conducted within the fields of memory and development, but has since expanded to many other fields and boasts its own dedicated journals and conferences.

Within this context, the birth of the field of visual metacognition can be traced to the late 2000s. That was the point when the already growing interest in understanding confidence in perceptual tasks (Baranski & Petrusic, 1994; Fleet et al., 1987; Keren, 1988) converged with a series of studies on the neural correlates of confidence in visual tasks (Del Cul et al., 2009; Fleming et al., 2010; Lau & Passingham, 2006; Rounis et al., 2010) and the development of improved measures of metacognitive ability (Maniscalco & Lau, 2012; Rounis et al., 2010). The field was additionally fueled by the finding of numerous dissociations between confidence and accuracy (Rahnev et al., 2011; Rahnev, Bahdo, et al., 2012; Rahnev, Maniscalco, et al., 2012; Wilimzig et al., 2008; Zylberberg et al., 2012), which helped to firmly establish metacognitive judgments of confidence as a separate object of study in their own right.

Since then, the field of visual metacognition has made substantial progress in understanding how people evaluate their performance in perceptual tasks (for reviews, see Fleming & Dolan, 2012; Mamassian, 2016; Meyniel et al., 2015; Shekhar & Rahnev, 2021a). The field has grown and matured, which can be seen from such projects as the creation of the Confidence Database (the largest repository of open data in the behavioral sciences; Rahnev et al., 2020) and an ambitious experiment in field-wide collective goal-setting (Rahnev et al., 2021).

Research on visual metacognition typically involves asking participants to provide confidence judgments about the accuracy of their perceptual decisions. In this context, the term “confidence” refers to the ratings that participants provide, “metacognitive ability” to the extent to which these ratings predict one’s accuracy on a task, and “visual metacognition” to the general field of studying how people evaluate their own performance via confidence ratings. Practically speaking, much of the research on visual metacognition seeks to understand how people give confidence judgments in perceptual tasks. The field has mostly focused on studying trial-by-trial retrospective confidence judgments in low-level visual tasks with comparatively less attention paid to prospective confidence judgments (but

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

see Fleming et al., 2016), global judgments about one's perceptual ability (but see Rouault & Fleming, 2020), or more complex visual stimuli (but see Lapate et al., 2020).

Three topics have emerged as central to the whole field of visual metacognition: (1) constructing appropriate measures of metacognitive ability, (2) developing computational models of metacognition, and (3) revealing the neural correlates of visual metacognition. Here I review recent progress in all three of these areas with special attention on the implications of the progress in visual metacognition for research in other domains of metacognition research. This review demonstrates that these three components of visual metacognition have been studied largely in isolation. I argue that continued progress requires the recognition of the deep links between the three areas, and examine recent work that has begun to build bridges between them.

Measures of visual metacognitive ability

One of the basic requirements for understand visual metacognition is the accurate measurement of metacognitive ability. Metacognitive ability is one's capacity for providing confidence ratings that predict his or her accuracy. Many studies have shown that people vary substantially in their metacognitive ability (Faivre et al., 2018; Mazancieux et al., 2020). Demonstrating the importance of this construct, metacognitive ability has been shown to correlate with both brain volume (Allen et al., 2017; Fleming et al., 2010) and the severity of several psychiatric symptoms (Jia et al., 2019; Rouault et al., 2018). Accurate measurement of metacognitive ability is also critical for understanding whether metacognition is a domain-general process that generalizes across different tasks (Faivre et al., 2018; Mazancieux et al., 2020) and assessing the effects of metacognitive training (Carpenter et al., 2019; Haddara & Rahnev, 2021). Given the importance of the issue, it is no surprise that several competing measures have been developed.

Traditional measures

There are several traditional measures of metacognitive ability. One of the most widespread (especially in the field of metamemory) is the Goodman–Kruskall gamma coefficient (or just *gamma*), which is essentially a rank correlation between trial-by-trial confidence and accuracy (Nelson, 1984). A related measure is the Pearson correlation (known as *phi*) between the raw values of confidence and accuracy (Kornell et al., 2007). In contrast to these two correlation measures, another traditional measure – the area under the Type 2 ROC function (Clarke et al., 1959) – relies on signal detection analyses.

Despite their intuitive appeal and widespread usage, all of these traditional measures have a critical flaw in that they all strongly depend on the performance on the primary task (Fleming & Lau, 2014). In other words, easier tasks automatically lead not just to higher accuracy on the perceptual task but also to higher estimated metacognitive ability when assessed with these measures. This property means that the traditional measures confound metacognitive ability with task performance.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

meta-d' and *Mratio*

To address this limitation of traditional measures, Maniscalco & Lau (2012) developed a new approach to measuring metacognitive ability based on signal detection theory (SDT). In this approach, the informativeness of confidence ratings is expressed in the units of *d'*, the standard SDT measure of perceptual sensitivity. The resulting measure was therefore called *meta-d'*. Just as the traditional measures above, *meta-d'* depends on the difficulty of the primary task. However, this approach allows the construction of a new measure, commonly referred to as *Mratio*, which is derived as the ratio *meta-d'*/*d'*. Indeed, *Mratio* has been found to measure metacognitive ability independent of primary task difficulty in both simulations (Barrett et al., 2013) and empirical data (Shekhar & Rahnev, 2021b). This property has made *Mratio* the most popular measure of metacognitive ability within the field of visual metacognition.

Relationship to other domains of metacognition research

The development of *Mratio* has allowed the field of visual metacognition to study metacognitive ability independent from primary task performance. Unfortunately, this measure has been slow to penetrate other areas, which continue to use traditional measures such as *gamma*. The use of these traditional measures means that many findings of differences in metacognitive ability across a variety of fields are likely to be an artifact of imperfectly matched primary task performance. For example, a recent paper that carefully matched accuracy on a memory task among older and younger adults concluded that the finding of a metacognitive difference between the two groups is an epiphenomenon of the age differences in memory (Hertzog et al., 2021). The adoption of *Mratio* could help other subfields of metacognition research avoid such epiphenomenal findings in the first place and weed them out with much greater ease.

An important area that requires additional research is the fuller characterization of the psychometric properties of *Mratio* and other measures of metacognitive ability. For example, beyond their dependence on the primary task performance, it is important to empirically assess whether these measures are independent from both response bias and metacognitive bias (the tendency to give either high or low confidence ratings) with recent research suggesting that *Mratio* may be confounded with metacognitive bias (Shekhar & Rahnev, 2021b; Xue et al., 2021). In addition, fundamental features of these measures such as their test-retest reliability are just beginning to be examined (Guggenmos, 2021). Finally, it is important to examine whether visual metacognitive ability is related to performance in other aspects of visual cognition such as imagery (Rademaker & Pearson, 2012), mental rotation, etc. The next few years are likely to see much progress on these topics.

Models of visual metacognition

Given its deep roots in visual psychophysics, it is no surprise that since its inception, the field of visual metacognition has strongly emphasized computational modeling. This emphasis has resulted in a proliferation of theories with more than a dozen distinct competing models. Here I review the target effects that models of visual metacognition

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

have been trying to explain, the dominant frameworks adopted, and the two main hypotheses for the computations underlying visual metacognition.

Target effects for models of visual metacognition

Models of visual metacognition have typically been developed with the goal of explaining certain target effects, especially ones related to measured metacognitive ability. Perhaps the most common target effect is explaining "metacognitive inefficiency," that is, the observation that metacognitive judgments of confidence are typically less informative about the accuracy of a decision than they could be (Shekhar & Rahnev, 2021a). Such inefficiency is typically modeled by assuming the presence of "metacognitive noise," which is random noise that affects the confidence ratings but not the primary decision (J. W. Bang et al., 2019; Barrett et al., 2013; Fleming & Daw, 2017; Maniscalco & Lau, 2016). Pinpointing the exact sources of this metacognitive noise is an area of active research (Shekhar & Rahnev, 2021a).

Conversely, metacognitive confidence judgments sometimes appear to exhibit "super-efficiency" such that they seem more informative than the theoretical maximum (Fleming & Daw, 2017). This phenomenon is usually thought to involve post-decisional processes similar to the ones involved in error detection (Yeung & Summerfield, 2012) and changes of mind (van den Berg et al., 2016), but may also be at least partly a methodological artifact (Rahnev & Fleming, 2019). Models that try to explain metacognitive super-efficiency typically assume that the metacognitive system has access to evidence not available for the primary decision either because the evidence arrived after the decision was made or because of the existence of partially independent streams of processing of object-level and metacognitive information (Fleming & Daw, 2017; Mamassian, 2020; Pleskac & Busemeyer, 2010). The exact reasons for the phenomenon of metacognitive super-efficiency thus remain ill-understood.

Finally, other models have focused on explaining more specific target effects. Some examples include the bias of confidence ratings to ignore disconfirming information (Koizumi et al., 2015; Peters et al., 2017; Samaha et al., 2016; Zylberberg et al., 2012), the influence of the previous trial's confidence rating (Rahnev et al., 2015) and the use of the clarity of the stimulus as a cue for confidence (Rausch et al., 2018).

It is thus clear that there are many target effects that models of visual metacognition may need to explain. The challenge ahead is to determine which of these effects are most critical for understanding the mechanisms of visual metacognition and focusing the field's modeling efforts on them.

Dominant frameworks

By far the most popular framework for models of visual metacognition is SDT (Green & Swets, 1966). The essence of SDT is that sensory processing is noisy and thus the internal activation generated by a stimulus varies from trial to trial. Within the SDT framework, confidence ratings are given by comparing the strength of the internal activation to predefined criteria. The popularity of the framework stems from SDT's dominance in explaining performance in perceptual tasks more generally and its minimal assumptions that

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

allow flexible model construction. Consequently, the majority of models in the field are in some way based on SDT (Aitchison et al., 2015; J. W. Bang et al., 2019; Fleming & Daw, 2017; Hangya et al., 2016; Maniscalco & Lau, 2016; Peters et al., 2017; Rausch et al., 2018; Shekhar & Rahnev, 2021b).

However, one limitation of all SDT-based models is that they cannot account for reaction time because SDT does not specify how the internal activation evolves over time. This limitation has motivated the adoption of an alternative framework based on sequential sampling, which can be used to jointly model choice, reaction time, and confidence. Different models within this framework have postulated very different mechanisms for confidence: from post-decisional evidence accumulation (Pleskac & Busemeyer, 2010), to using the difference between two competing accumulators (Vickers, 1979), to employing reaction time as a cue for confidence (Kiani et al., 2014), to having different accumulators specifically linked to each confidence response (Ratcliff & Starns, 2009, 2013). There is thus little convergence thus far on how the sequential sampling framework should be used in the context of confidence, and the properties of most of these models have not been examined in detail.

The existence of two separate modeling frameworks is both exciting and disconcerting. Although there are deep links between SDT and sequential sampling (Griffith et al., 2021), the actual models of visual metacognition developed from one framework cannot always be directly translated to the other framework. Yet, there is of yet almost no work on directly comparing the performance of the two frameworks in explaining metacognitive judgments. This topic is thus a particularly exciting avenue for future research.

The computations underlying visual metacognition

Regardless of which target effects one tries to explain and which framework one adopts, a remaining critical issue for each model is to describe the assumed computations underlying visual metacognition. The majority of existing models postulate simple computations where the level of evidence favoring one choice over another is directly translated to a specific confidence rating via several criteria such that higher evidence directly translates to higher confidence. The appeal of this mechanism comes from its simplicity and biological plausibility. However, the nature of the relevant internal evidence can vary substantially between models. For example, some models assume that confidence is based only on the evidence for the chosen option and ignores the evidence for all other options (Koizumi et al., 2015; Peters et al., 2017; Samaha et al., 2016; Zylberberg et al., 2012), while many others assume that the evidence used for metacognitive judgments already reflects the difference between the evidence for each option (J. W. Bang et al., 2019; Green & Swets, 1966; Maniscalco & Lau, 2016; Pleskac & Busemeyer, 2010; Vickers, 1979).

However, basing confidence directly on signal strength is not optimal. Ideally, confidence should be based on a Bayesian computation of the posterior probability of being correct and multiple models assume that such computations underlie confidence judgments (Aitchison et al., 2015; Fleming & Daw, 2017; Meyniel et al., 2015). Nevertheless, it is still unclear whether people can actually perform the complex computations required to estimate the posterior probability of being correct and this is a topic of active research (Bertana et al., 2021).

The distinction between basing metacognitive judgments of confidence directly on internal activations versus on Bayesian computations has strong implications about how visual metacognition is conceptualized. Resolving the issue has been selected as one of the central short-term goals for the whole field (Rahnev et al., 2021) and research on this question is thus likely to proliferate in the coming years.

Relationship to other domains of metacognition research

Although early models of metacognition were developed primarily within the context of metamemory, visual metacognition has now become the dominant driver of model development in metacognition research. The issues related to the relevant target effects, frameworks and assumed computations are relevant to any subfield of metacognition research. Further, the field of visual metacognition has led the way in not just developing new models but also engaging in extensive model comparison and selection (Aitchison et al., 2015; Maniscalco & Lau, 2016; Shekhar & Rahnev, 2021b). Researchers from other domains of metacognition research would be wise to consider and build on the modeling progress made within visual metacognition.

Neural correlates of visual metacognition

Beyond its focus on models and measures, the field of visual metacognition has put a strong emphasis on revealing the neural correlates of metacognition. Research in this area has its roots in studies on metamemory conducted from the 1980s on. Early studies examined patients with Korsakoff's syndrome and found that they exhibited metamemory impairments likely due to accompanying frontal lobe dysfunction (Janowsky et al., 1989; Shimamura & Squire, 1986). Follow-up studies focusing on patients with lesions to frontal cortex (Pannu et al., 2005; Schnyer et al., 2004) and neuroimaging (Chua et al., 2006; Kikyo et al., 2002; Moritz et al., 2006) confirmed the critical role of various prefrontal regions in both feeling-of-knowing and retrospective confidence judgments of memory. These studies provided the groundwork for later studies that focused on visual metacognition.

One of the main challenges for determining the neural bases of visual metacognition is to ensure that the brain mechanisms identified are specific to metacognition. Indeed, metacognitive processes often co-exist (and can be easily confused) with other cognitive processes such as attention, working memory or cognitive effort (Fleming & Dolan, 2012). In addition, since metacognitive judgments of confidence depend so heavily on the difficulty of the first order task, it is critical to ensure that the identified neural mechanisms truly correspond to metacognition rather than simply to accuracy or reaction time. Researchers have adopted multiple strategies to successfully meet these challenges. Here I briefly review relevant research that revealed the anatomical correlates of metacognitive ability, applied creative manipulations in the context of functional neuroimaging or used causal manipulations to selectively target metacognitive judgments.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Anatomical correlates of metacognitive ability

An early line of research that sought to identify the brain structures selectively involved in metacognition relied on correlating different anatomical brain measures with subject-by-subject metacognitive ability. In an influential early study, Fleming et al. (2010) equated performance across participants on a perceptual task and then correlated their measured metacognitive ability with gray matter volume. They found that higher gray matter volume in the anterior prefrontal cortex (aPFC) was correlated with higher metacognitive ability. This finding has been replicated and extended by several other studies (Allen et al., 2017; McCurdy et al., 2013). In addition, a recent study further found that metacognitive ability in a perceptual task is related to the functional integrity of the superior longitudinal fasciculus – a major white matter tract that connects frontal and parietal lobes (Zheng et al., 2021). Together, these studies demonstrate that people's metacognitive ability can be related to various anatomical substrates, mostly connected to the frontal lobes. By equating the accuracy on the task for all participants, the experimenters could ensure that the anatomical correlates were independent of first-order performance.

Evidence from functional neuroimaging

Although studies correlating metacognitive ability with anatomical measures have been influential, they can only reveal individual anatomical structures. However, there is accumulating evidence that metacognition emerges via network interactions and not simply from the function of single regions (Yeon et al., 2020). Revealing the networks underlying visual metacognition has largely relied on neuroimaging research.

The most common approach has been to simply compare the brain activity for trials with low vs. high confidence (e.g., Fleck, 2005). However, given that low and high confidence trials also differ in accuracy and reaction time, this approach confounds the metacognitive processes with first-order performance. Subsequent studies have used a variety of approaches to address this confound. For example, an early study used a metacontrast masking paradigm that produced a difference in subjective visibility judgments for conditions with matched performance (Lau & Passingham, 2006). Other studies have compared brain activations in conditions where participants did or did not rate their confidence (Fleming et al., 2012; Morales et al., 2018; Yeon et al., 2020), employed decoded neurofeedback to induce a change in confidence (Cortese et al., 2016), or compared confidence in the accuracy of one's own vs. another agent's choice (Pereira et al., 2020). Another approach has been to go beyond trial-by-trial confidence and instead focus on the formation of global judgments about one's perceptual ability (Rouault & Fleming, 2020), an understudied topic that provides a distinct perspective on visual metacognition.

This research has implicated a number of mostly frontal, parietal and cingulate areas as integral to visual metacognition (Baird et al., 2013; Fleck, 2005; Fleming et al., 2012; Lau & Passingham, 2006; Morales et al., 2018; Yeon et al., 2020). Importantly, it has also shown the critical role of the network communications between these regions such as the observed increase in functional connectivity between frontal and visual areas associated with engaging in metacognitive evaluation (Fleming et al., 2012). Future work needs to explore these network interactions in more depth with a particular focus on how metacognition may relate to the interactions within and between large brain-wide communities.

Causal studies

Beyond the commonly used correlational techniques, several studies have employed techniques that afford a causal association between a brain area and metacognition. An early study delivered transcranial magnetic stimulation (TMS) to dorsolateral prefrontal cortex (DLPFC) and reported a selective influence on metacognitive scores with no change in the primary task performance (Rounis et al., 2010). Follow-up TMS studies have confirmed the causal role of both DLPFC and aPFC in visual metacognition (Lapate et al., 2020; Rahnev et al., 2016; Shekhar & Rahnev, 2018). The conclusions from these brain stimulation studies have received further support from a study that found selective impairments in visual but not memory metacognition for patients with anterior frontal lesions (Fleming et al., 2014). Overall, this research convincingly demonstrates that a number of areas in the prefrontal cortex are involved in visual metacognitive evaluation.

Relationship to other domains of metacognition research

One critical question with implications for all subfields of metacognition research is to what extent metacognition is domain-general or domain-specific. This question has been addressed at great length within the context of visual metacognition and metamemory (Faivre et al., 2018; Fleming et al., 2014; McCurdy et al., 2013; Morales et al., 2018). This research has revealed a widespread domain-general network of areas in frontal and posterior midline regions, but with several domain-specific areas such as aPFC for visual metacognition and precuneus for memory metacognition. These findings demonstrate the close link between the various subareas of metacognition research. They also raise exciting questions regarding the sources and implications of the partial domain-specificity observed, with these questions like to receive substantial attention in the coming years.

The need for synergy in research on measures, models and neural correlates

The brief review thus far makes it obvious that tremendous progress has been made in the last 15 years in constructing measures of metacognitive ability, developing models and revealing the neural correlates of visual metacognition. However, a limitation of research to date has been that these three areas have been investigated mostly in isolation. The problem with such lack of integration is that, as explained below, deep understanding of any of these areas requires synergy with the others.

The link between models and measures of metacognition

As already reviewed, there are more than a dozen models and over half a dozen measures of visual metacognition. Yet, none of the existing measures are derived from an existing model of metacognition, and none of the models are informed by the existing measures of metacognition. This lack of integration is problematic because it ignores the intimate relationship between models and measures of metacognition. Specifically, a psychophysical

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

measure is never truly model-free; in fact, any measure implies an underlying model even if the model has not been explicitly derived (Macmillan & Creelman, 2005). Conversely, every model of metacognition has implications about the measurement of metacognitive ability and could be used to derive a measure of metacognition.

Fortunately, the critical link between models and measures is becoming recognized. Recently, Shekhar & Rahnev (2021b) examined the dependence of *Mratio* on participants' metacognitive bias (that is, their bias towards being more or less confident; Fleming & Lau, 2014). Ideally, one's metacognitive bias should be independent of one's estimated metacognitive ability because metacognitive bias can be manipulated at will (i.e., one can easily increase or decrease their reported level of confidence). However, Shekhar & Rahnev found that higher confidence led to higher *Mratio* values (the finding was subsequently replicated and extended by Xue et al., 2021). This led them to propose a new model of metacognition with signal-dependent, lognormally-distributed metacognitive noise that could explain this relationship. This model was found to outperform more standard models with either no metacognitive noise or signal-independent metacognitive noise. Conversely, the estimated metacognitive noise (called σ_{meta}) was found not to depend on one's metacognitive bias (it was also independent of task difficulty), which could make it a better measure of metacognitive ability. Both the new model and the new measure of metacognition need to be further tested, but the approach clearly demonstrates how jointly considering models and measures of metacognition can lead to developments in both areas.

The link between models and neural mechanisms of metacognition

Similar to the link between models and measures, there is an intimate link between models and neural correlates of metacognition. On one hand, the identity of the brain regions implicated in metacognition constrain the possible models. For example, the critical role of the prefrontal cortex in metacognition is consistent with models where metacognitive evaluation is performed by a second-order system that monitors the first-order system that makes the primary decision (Nelson & Narens, 1990). On the other hand, models of metacognition constrain the possible computational roles that different brain areas could perform. However, as with models and measures, the intimate link between models and neural mechanisms has rarely been exploited to improve our understanding in either area.

Nevertheless, this link is also beginning to be exploited. For example, Bang & Fleming (2018) constructed a model that separates confidence and sensory reliability. The model allowed them to identify brain areas associated separately with each quantity and thus to elucidate the computational roles of different brain areas. Similarly, Shekhar & Rahnev (2018) found that TMS to DLPFC and aPFC had dissociable effects on average confidence and metacognition. These results supported a hierarchical model of visual metacognition (J. W. Bang et al., 2019; Maniscalco & Lau, 2016). The resulting model, in turn, suggested distinct functional roles for DLPFC and aPFC such that DLPFC reads out the strength of the sensory evidence and relays it to aPFC, which makes the confidence judgement by incorporating additional, non-perceptual information.

As the examples above demonstrate, an approach that jointly considers the measures, models and neural mechanisms of visual metacognition has the promise of driving progress in all

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

three areas of research. The coming years should see a proliferation of this approach in visual metacognition and beyond.

Conclusions

The field of visual metacognition is only about 15 years old but has already made great strides. Building on the solid foundation of psychophysics research, it has adopted a staunchly computational approach that is displayed in the development of numerous models. This computational focus has allowed it to construct improved measures of metacognitive ability that transcend the limitations of traditional measures. Finally, the ease of manipulating visual stimuli has allowed for unprecedented progress in revealing the neural correlates and mechanisms of visual metacognition. The next challenge is to better integrate the research on measures, models and neural correlates of metacognition, as well as to make stronger connections to other subfields of metacognition research.

Acknowledgements

This work was supported by the National Institute of Health (award: R01MH119189) and the Office of Naval Research (award: N00014–20-1–2622).

References

Aitchison L, Bang D, Bahrami B, & Latham PE (2015). Doubly Bayesian Analysis of Confidence in Perceptual Decision-Making. PLoS Computational Biology, 11(10), e1004519. 10.1371/journal.pcbi.1004519 [PubMed: 26517475]

Allen M, Glen JC, Müllensiefen D, Schwarzkopf DS, Fardo F, Frank D, Callaghan MF, & Rees G (2017). Metacognitive ability correlates with hippocampal and prefrontal microstructure. NeuroImage, 149, 415–423. 10.1016/j.neuroimage.2017.02.008 [PubMed: 28179164]

Baird B, Smallwood J, Gorgolewski KJ, & Margulies DS (2013). Medial and lateral networks in anterior prefrontal cortex support metacognitive ability for memory and perception. The Journal of Neuroscience, 33(42), 16657–16665. 10.1523/JNEUROSCI.0786-13.2013 [PubMed: 24133268]

Bang D, & Fleming SM (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. Proceedings of the National Academy of Sciences of the United States of America, 115(23), 6082–6087. 10.1073/pnas.1800795115 [PubMed: 29784814]

Bang JW, Shekhar M, & Rahnev D (2019). Sensory noise increases metacognitive efficiency. Journal of Experimental Psychology: General, 148(3), 437–452. 10.1037/xge0000511 [PubMed: 30382720]

Baranski JV, & Petrusic WM (1994). The calibration and resolution of confidence in perceptual judgments. Perception & Psychophysics, 55(4), 412–428. http://www.ncbi.nlm.nih.gov/pubmed/8036121 [PubMed: 8036121]

Barrett AB, Dienes Z, & Seth AK (2013). Measures of metacognition on signal-detection theoretic models. Psychological Methods, 18(4), 535–552. 10.1037/a0033268 [PubMed: 24079931]

Bertana A, Chetverikov A, van Bergen RS, Ling S, & Jehee JFM (2021). Dual strategies in human confidence judgments. Journal of Vision, 21(5), 21. 10.1167/jov.21.5.21

Carpenter J, Sherman MT, Kievit RA, Seth AK, Lau H, & Fleming SM (2019). Domain-general enhancements of metacognitive ability through adaptive training. Journal of Experimental Psychology: General, 148(1), 51–64. 10.1037/xge0000505 [PubMed: 30596440]

Chua EF, Schacter DL, Rand-Giovannetti E, & Sperling RA (2006). Understanding metamemory: neural correlates of the cognitive process and subjective level of confidence in recognition memory. NeuroImage, 29(4), 1150–1160. 10.1016/j.neuroimage.2005.09.058 [PubMed: 16303318]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Clarke FR, Birdsall TG, & Tanner WP (1959). Two Types of ROC Curves and Definitions of Parameters. The Journal of the Acoustical Society of America, 31(5), 629–630. 10.1121/1.1907764

Cortese A, Amano K, Koizumi A, Kawato M, & Lau H (2016). Multivoxel neurofeedback selectively modulates confidence without changing perceptual performance. Nature Communications, 7, 13669. 10.1038/ncomms13669

Cul Del, a, Dehaene S, Reyes P, Bravo E, & Slachevsky A (2009). Causal role of prefrontal cortex in the threshold for access to consciousness. Brain, 132(Pt 9), 2531–2540. 10.1093/brain/awp111 [PubMed: 19433438]

Desender K, Boldt A, & Yeung N (2018). Subjective Confidence Predicts Information Seeking in Decision Making. Psychological Science, 29(5), 761–778. 10.1177/0956797617744771 [PubMed: 29608411]

Faivre N, Filevich E, Solovey G, Kühn S, & Blanke O (2018). Behavioral, Modeling, and Electrophysiological Evidence for Supramodality in Human Metacognition. The Journal of Neuroscience, 38(2), 263–277. 10.1523/JNEUROSCI.0322-17.2017 [PubMed: 28916521]

Fechner GT (1860). Elemente der Psychophysik. Breitkopf und Härtel.

Flavell JH (1971). First Discussant’s Comments: What is Memory Development the Development of? Human Development, 14(4), 272–278. 10.1159/000271221

Flavell JH (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. American Psychologist, 34(10), 906–911. 10.1037/0003-066X.34.10.906

Fleck MS (2005). Role of Prefrontal and Anterior Cingulate Regions in Decision-Making Processes Shared by Memory and Nonmemory Tasks. Cerebral Cortex, 16(11), 1623–1630. 10.1093/cercor/bhj097 [PubMed: 16400154]

Fleet MML, Brigham JC, & Bothwell RK (1987). The Confidence-Accuracy Relationship: The Effects of Confidence Assessment and Choosing. Journal of Applied Social Psychology, 17(2), 171–187. 10.1111/j.1559-1816.1987.tb00308.x

Fleming SM, & Daw ND (2017). Self-evaluation of decision performance: A general Bayesian framework for metacognitive computation. Psychological Review, 124(1), 91–114. 10.1037/rev0000045 [PubMed: 28004960]

Fleming SM, & Dolan RJ (2012). The neural basis of metacognitive ability. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 367(1594), 1338–1349. 10.1098/rstb.2011.0417 [PubMed: 22492751]

Fleming SM, Huijgen J, & Dolan RJ (2012). Prefrontal contributions to metacognition in perceptual decision making. The Journal of Neuroscience, 32(18), 6117–6125. 10.1523/JNEUROSCI.6489-11.2012 [PubMed: 22553018]

Fleming SM, & Lau HC (2014). How to measure metacognition. Frontiers in Human Neuroscience, 8. 10.3389/fnhum.2014.00443

Fleming SM, Massoni S, Gajdos T, & Vergnaud J-C (2016). Metacognition about the past and future: quantifying common and distinct influences on prospective and retrospective judgments of self-performance. Neuroscience of Consciousness, 2016(1), niw018. 10.1093/nc/niw018 [PubMed: 30356936]

Fleming SM, Ryu J, Golfinos JG, & Blackmon KE (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. Brain, 137(10), 2811–2822. 10.1093/brain/awu221 [PubMed: 25100039]

Fleming SM, Weil RS, Nagy Z, Dolan RJ, & Rees G (2010). Relating introspective accuracy to individual differences in brain structure. Science, 329(5998), 1541–1543. 10.1126/science.1191883 [PubMed: 20847276]

Green DM, & Swets JA (1966). Signal detection theory and psychophysics. John Wiley & Sons Ltd.

Griffith T, Baker SA, & Lepora NF (2021). The statistics of optimal decision making: Exploring the relationship between signal detection theory and sequential analysis. Journal of Mathematical Psychology, 103, 102544. 10.1016/j.jmp.2021.102544

Guggenmos M (2021). Validity and reliability of metacognitive performance measures. PsyArXiv. 10.31234/osf.io/jrkzm

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Haddara N, & Rahnev D (2021). The impact of feedback on perceptual decision making and metacognition: Reduction in bias but no change in sensitivity. Psychological Science.

Hangya B, Sanders JL, & Kepecs A (2016). A Mathematical Framework for Statistical Decision Confidence. Neural Computation, 28(9), 1840–1858. 10.1162/NECO_a_00864 [PubMed: 27391683]

Hart JT (1965). Memory and the feeling-of-knowing experience. Journal of Educational Psychology, 56(4), 208–216. 10.1037/h0022263 [PubMed: 5825050]

Helmholtz HLF (1856). Treatise on physiological optics. Thoemmes Continuum.

Hertzog C, Curley T, & Dunlosky J (2021). Are age differences in recognition-based retrieval monitoring an epiphenomenon of age differences in memory? Psychology and Aging, 36(2), 186–199. 10.1037/pag0000595 [PubMed: 33793261]

Janowsky JS, Shimamura AP, Kritchevsky M, & Squire LR (1989). Cognitive impairment following frontal lobe damage and its relevance to human amnesia. Behavioral Neuroscience, 103(3), 548–560. [PubMed: 2736069]

Jia W, Zhu H, Ni Y, Su J, Xu R, Jia H, & Wan X (2019). Disruptions of frontoparietal control network and default mode network linking the metacognitive deficits with clinical symptoms in schizophrenia. Human Brain Mapping, hbm.24887. 10.1002/hbm.24887

Keren G (1988). On the ability of monitoring non-veridical perceptions and uncertain knowledge: Some calibration studies. Acta Psychologica, 67(2), 95–119. 10.1016/0001-6918(88)90007-8 [PubMed: 3407444]

Kiani R, Corthell L, & Shadlen MN (2014). Choice Certainty Is Informed by Both Evidence and Decision Time. Neuron, 84(6), 1329–1342. 10.1016/j.neuron.2014.12.015 [PubMed: 25521381]

Kikyo H, Ohki K, & Miyashita Y (2002). Neural correlates for feeling-of-knowing: An fMRI parametric analysis. Neuron, 36(1), 177–186. 10.1016/S0896-6273(02)00939-X [PubMed: 12367516]

Koizumi A, Maniscalco B, & Lau H (2015). Does perceptual confidence facilitate cognitive control? Attention, Perception, & Psychophysics, 77(4), 1295–1306. 10.3758/s13414-015-0843-3

Kornell N, Son LK, & Terrace HS (2007). Transfer of metacognitive skills and hint seeking in monkeys. Psychological Science, 18(1), 64–71. 10.1111/j.1467-9280.2007.01850.x [PubMed: 17362380]

Lapate RC, Samaha J, Rokers B, Postle BR, & Davidson RJ (2020). Perceptual metacognition of human faces is causally supported by function of the lateral prefrontal cortex. Communications Biology, 3(1), 360. 10.1038/s42003-020-1049-3 [PubMed: 32647260]

Lau H, & Passingham RE (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. Proceedings of the National Academy of Sciences of the United States of America, 103(49), 18763–18768. 10.1073/pnas.0607716103 [PubMed: 17124173]

Macmillan NA, & Creelman CD (2005). Detection Theory: A User's Guide (2nd ed.). Erlbaum.

Mamassian P (2016). Visual Confidence. Annual Review of Vision Science, 2(1), annurev-vision-111815–114630. 10.1146/annurev-vision-111815-114630

Mamassian P (2020). Confidence Forced-Choice and Other Metaperceptual Tasks. Perception, 49(6), 616–635. 10.1177/0301006620928010 [PubMed: 32552488]

Maniscalco B, & Lau H (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Consciousness and Cognition, 21(1), 422–430. 10.1016/j.concog.2011.09.021 [PubMed: 22071269]

Maniscalco B, & Lau H (2016). The signal processing architecture underlying subjective reports of sensory awareness. Neuroscience of Consciousness, 2016(1), 1–17. 10.1093/nc/niw002

Mazancieux A, Fleming SM, Souchay C, & Moulin CJA (2020). Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks. Journal of Experimental Psychology: General, 149(9), 1788–1799. 10.1037/xge0000746 [PubMed: 32191079]

McCurdy LY, Maniscalco B, Metcalfe J, Liu KY, de Lange FP, & Lau H (2013). Anatomical Coupling between Distinct Metacognitive Systems for Memory and Visual Perception. The Journal of Neuroscience, 33(5), 1897–1906. 10.1523/JNEUROSCI.1890-12.2013 [PubMed: 23365229]

Meyniel F, Sigman M, & Mainen ZF (2015). Confidence as Bayesian Probability: From Neural Origins to Behavior. Neuron, 88(1), 78–92. 10.1016/j.neuron.2015.09.039 [PubMed: 26447574]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Morales J, Lau H, & Fleming SM (2018). Domain-General and Domain-Specific Patterns of Activity Supporting Metacognition in Human Prefrontal Cortex. The Journal of Neuroscience, 38(14), 3534–3546. 10.1523/JNEUROSCI.2360-17.2018 [PubMed: 29519851]

Moritz S, Gläscher J, Sommer T, Büchel C, & Braus DF (2006). Neural correlates of memory confidence. NeuroImage, 33(4), 1188–1193. 10.1016/j.neuroimage.2006.08.003 [PubMed: 17029986]

Nelson TO (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychological Bulletin, 95(1), 109–133. http://www.ncbi.nlm.nih.gov/pubmed/6544431 [PubMed: 6544431]

Nelson TO, & Narens L (1990). Metamemory: A theoretical framework and some new findings. In Bower G (Ed.), The Psychology of Learning and Motivation (pp. 125–141). Academic Press.

Pannu JK, Kaszniak AW, & Rapcsak SZ (2005). Metamemory for faces following frontal lobe damage. Journal of the International Neuropsychological Society : JINS, 11(6), 668–676. 10.1017/S1355617705050873 [PubMed: 16248902]

Peirce CS, & Jastrow J (1884). On Small Differences in Sensation. Memoirs of the National Academy of Sciences, 3, 75–83.

Pereira M, Faivre N, Iturrate I, Wirthlin M, Serafini L, Martin S, Desvachez A, Blanke O, Van De Ville D, & Millán JDR (2020). Disentangling the origins of confidence in speeded perceptual judgments through multimodal imaging. Proceedings of the National Academy of Sciences, 117(15), 8382–8390. 10.1073/pnas.1918335117

Pescetelli N, & Yeung N (2021). The role of decision confidence in advice-taking and trust formation. Journal of Experimental Psychology: General, 150(3), 507–526. 10.1037/xge0000960 [PubMed: 33001684]

Peters MAK, Thesen T, Ko YD, Maniscalco B, Carlson C, Davidson M, Doyle W, Kuzniecky R, Devinsky O, Halgren E, & Lau H (2017). Perceptual confidence neglects decision-incongruent evidence in the brain. Nature Human Behaviour, 1(7), 0139. 10.1038/s41562-017-0139

Pleskac TJ, & Busemeyer JR (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. Psychological Review, 117(3), 864–901. 10.1037/a0019737 [PubMed: 20658856]

Rademaker RL, & Pearson J (2012). Training Visual Imagery: Improvements of Metacognition, but not Imagery Strength. Frontiers in Psychology, 3, 224. 10.3389/fpsyg.2012.00224 [PubMed: 22787452]

Rahnev D, Bahdo L, de Lange FP, & Lau H (2012). Prestimulus hemodynamic activity in dorsal attention network is negatively associated with decision confidence in visual perception. Journal of Neurophysiology, 108(5), 1529–1536. 10.1152/jn.00184.2012 [PubMed: 22723670]

Rahnev D, Balsdon T, Charles L, Gardelle V. de, Denison RN, Desender K, Faivre N, Filevich E, Fleming S, Jehee J, Lau H, Lee ALF, Locke SM, Mamassian P, Odegaard B, Peters MAK, Reyes G, Rouault M, Sackur J, … Zylberberg A (2021). Consensus goals for the field of visual metacognition. PsyArXiv. 10.31234/OSF.IO/Z8V5X

Rahnev D, Desender K, Lee ALF, Adler WT, Aguilar-Lleyda D, Akdoğan B, Arbuzova P, Atlas LY, Balcı F, Bang JW, Bègue I, Birney DP, Brady TF, Calder-Travis J, Chetverikov A, Clark TK, Davranche K, Denison RN, Dildine TC, … Zylberberg A (2020). The Confidence Database. Nature Human Behaviour, 4(3), 317–325. 10.1038/s41562-019-0813-1

Rahnev D, & Fleming SM (2019). How experimental procedures influence estimates of metacognitive ability. Neuroscience of Consciousness, 2019(1), niz009. 10.1093/nc/niz009 [PubMed: 31198586]

Rahnev D, Koizumi A, McCurdy LY, D'Esposito M, & Lau H (2015). Confidence Leak in Perceptual Decision Making. Psychological Science, 26(11), 1664–1680. 10.1177/0956797615595037 [PubMed: 26408037]

Rahnev D, Maniscalco B, Graves T, Huang E, De Lange FP, & Lau H (2011). Attention induces conservative subjective biases in visual perception. Nature Neuroscience, 14(12), 1513–1515. 10.1038/nn.2948 [PubMed: 22019729]

Rahnev D, Maniscalco B, Luber B, Lau H, & Lisanby SH (2012). Direct injection of noise to the visual cortex decreases accuracy but increases decision confidence. Journal of Neurophysiology, 107(6), 1556–1563. 10.1152/jn.00985.2011 [PubMed: 22170965]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Rahnev D, Nee DE, Riddle J, Larson AS, & D'Esposito M (2016). Causal evidence for frontal cortex organization for perceptual decision making. Proceedings of the National Academy of Sciences, 113(20), 6059–6064. 10.1073/pnas.1522551113

Ratcliff R, & Starns JJ (2009). Modeling confidence and response time in recognition memory. Psychological Review, 116(1), 59–83. 10.1037/a0014086 [PubMed: 19159148]

Ratcliff R, & Starns JJ (2013). Modeling confidence judgments, response times, and multiple choices in decision making: recognition memory and motion discrimination. Psychological Review, 120(3), 697–719. 10.1037/a0033152 [PubMed: 23915088]

Rausch M, Hellmann S, & Zehetleitner M (2018). Confidence in masked orientation judgments is informed by both evidence and visibility. Attention, Perception, and Psychophysics, 80, 134–154. 10.3758/s13414-017-1431-5

Rouault M, & Fleming SM (2020). Formation of global self-beliefs in the human brain. Proceedings of the National Academy of Sciences of the United States of America, 117(44), 27268–27276. 10.1073/pnas.2003094117 [PubMed: 33060292]

Rouault M, Seow T, Gillan CM, & Fleming SM (2018). Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. Biological Psychiatry, 84(6), 443–451. 10.1016/j.biopsych.2017.12.017 [PubMed: 29458997]

Rounis E, Maniscalco B, Rothwell JC, Passingham RE, & Lau H (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. Cognitive Neuroscience, 1(3), 165–175. 10.1080/17588921003632529 [PubMed: 24168333]

Samaha J, Barrett JJ, Sheldon AD, LaRocque JJ, & Postle BR (2016). Dissociating Perceptual Confidence from Discrimination Accuracy Reveals No Influence of Metacognitive Awareness on Working Memory. Frontiers in Psychology, 7, 851. 10.3389/fpsyg.2016.00851 [PubMed: 27375529]

Schnyer DM, Verfaellie M, Alexander MP, LaFleche G, Nicholls L, & Kaszniak AW (2004). A role for right medial prefontal cortex in accurate feeling-of-knowing judgements: evidence from patients with lesions to frontal cortex. Neuropsychologia, 42(7), 957–966. 10.1016/j.neuropsychologia.2003.11.020 [PubMed: 14998710]

Shekhar M, & Rahnev D (2018). Distinguishing the roles of dorsolateral and anterior PFC in visual metacognition. Journal of Neuroscience, 38(22), 5078–5087. 10.1523/JNEUROSCI.3484-17.2018 [PubMed: 29720553]

Shekhar M, & Rahnev D (2021a). Sources of Metacognitive Inefficiency. Trends in Cognitive Sciences, 25(1), 12–23. 10.1016/j.tics.2020.10.007 [PubMed: 33214066]

Shekhar M, & Rahnev D (2021b). The nature of metacognitive inefficiency in perceptual decision making. Psychological Review, 128(1), 45–70. 10.1037/rev0000249 [PubMed: 32673034]

Shimamura AP, & Squire LR (1986). Memory and metamemory: a study of the feeling-of-knowing phenomenon in amnesic patients. Journal of Experimental Psychology. Learning, Memory, and Cognition, 12(3), 452–460. http://www.ncbi.nlm.nih.gov/pubmed/2942629 [PubMed: 2942629]

van den Berg R, Anandalingam K, Zylberberg A, Kiani R, Shadlen MN, & Wolpert DM (2016). A common mechanism underlies changes of mind about decisions and confidence. ELife, 5, e12192. 10.7554/eLife.12192 [PubMed: 26829590]

Vickers D (1979). Decision Processes in Visual Perception. Academic Press.

Wilimzig C, Tsuchiya N, Fahle M, Einhäuser W, & Koch C (2008). Spatial attention increases performance but not subjective confidence in a discrimination task. Journal of Vision, 8(5), 1–10. 10.1167/8.5.7

Xue K, Shekhar M, & Rahnev D (2021). The nature of metacognitive noise confounds metacognitive sensitivity and metacognitive bias. PsyArXiv. 10.31234/osf.io/buahk

Yeon J, Shekhar M, & Rahnev D (2020). Overlapping and unique neural circuits are activated during perceptual decision making and confidence. Scientific Reports, 10(1), 20761. 10.1038/s41598-020-77820-6 [PubMed: 33247212]

Yeung N, & Summerfield C (2012). Metacognition in human decision-making: confidence and error monitoring. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 367(1594), 1310–1321. 10.1098/rstb.2011.0416 [PubMed: 22492749]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Zheng Y, Wang D, Ye Q, Zou F, Li Y, & Kwok SC (2021). Diffusion property and functional connectivity of superior longitudinal fasciculus underpin human metacognition. Neuropsychologia, 156, 107847. 10.1016/j.neuropsychologia.2021.107847 [PubMed: 33812946]

Zylberberg A, Barttfeld P, & Sigman M (2012). The construction of confidence in a perceptual decision. Frontiers in Integrative Neuroscience, 6(September), 79. 10.3389/fnint.2012.00079 [PubMed: 23049504]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Public Significance Statement

The paper provides an overview of the field of visual metacognition. It discusses progress on constructing measures of metacognitive ability, developing computational models and revealing the neural correlates of visual metacognition. It is argued that continued progress requires synergy between these areas of study, and that progress in visual metacognition has direct implications for research in all subfields of metacognition research.