Jorge Morales[1]

Confidence in Consciousness Science

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

Abstract: *In her book* Introspection*, Maja Spener urges us to split the study of introspection in two: as a mental capacity and as a scientific method. In this article, I focus on the latter and argue that metacognition in general, and confidence ratings in particular, offer a more promising methodological bridge between subjective experience and objective measurement in consciousness science than Spener allows. While Spener provides a nuanced taxonomy of introspective modes and emphasizes the importance of calibration, she remains sceptical that confidence ratings, unlike introspection as a method, reveal the conscious character of perceptual states. I contend that this scepticism is unwarranted. Confidence ratings possess objective anchors, are easy to operationalize, and correlate tightly with both subjective visibility judgments — a hallmark of introspection — and known neural markers of awareness. Empirical and theoretical considerations show that confidence and consciousness are deeply intertwined, making confidence ratings a tractable and mechanistically informed tool for consciousness and introspection research.*

The ability to know one's own mind is front and centre of philosophical and psychological theory and practice. It is no exaggeration to say that, despite behaviourists' worst instincts, there would be no modern psychological science without introspection. Philosophy of mind, philosophy of action, and ethics would lose a most precious resource. Studying introspection, however, is hard. A millennia-old

Correspondence:
Email: j.morales@northeastern.edu

[1] Northeastern University, Boston, MA, USA.

Journal of Consciousness Studies, **32**, No. 11–12, 2025, pp. 206–20
DOI: 10.53765/20512201.32.11.206

challenge for scholars of self-knowledge — from the inception of the Delphic 'know thyself' maxim to the establishment of psychophysics — is how to determine the exact value of introspection. One of the most important lessons from Maja Spener's (2024) latest book, *Introspection: First-Person Access in Science and Agency*, is that to answer the question about introspection's value (e.g. its accuracy, reliability, trustworthiness, etc.), we need to split the question in two. On one hand, there is introspection as a mental capacity, and on the other, there is introspection as a scientific method. Spener spends the first half of her book on the latter, and the second on the former. Here I will focus on introspection as a method in consciousness science and, in particular, on its relationship with confidence.[2]

Spener's approach to introspection has several features that make it rather attractive. First, as just mentioned, she offers a complex taxonomy of introspection as a capacity and as a method that interacts with three different types of modes of access. Complementing her previous historical dive into introspection in the twentieth century (Spener, 2018), Spener recovers in her book three modes of introspective access from nineteenth-century introspectionist psychology: inner attention, inner apprehension, and retrospection (2024, chapter 1). In turn, she highlights the importance of calibrating introspection and in doing so she introduces a third useful distinction between calibration of operation (the conditions under which introspection operates) and calibration of content (the mental states over which it operates). This 2 × 3 × 2 matrix paints a sophisticated and complex picture of introspection that Spener leverages throughout the book. That Spener takes seriously the complexity of the states that we introspect both in daily life and in laboratory tasks, as well as the methods involved in their study, is a refreshing contribution that adds to an existing literature devoted to a pluralistic understanding of introspective capacities (Franco, 2024; Renero, 2019). The contrast with philosophers' dear but too-simple examples about seeing red patches is notable and very welcome.

A second distinctive ingredient in Spener's book is its solid historical foundation. With a profound understanding of the history of psychology, she takes readers through the details of the methods of nineteenth- and twentieth-century titans in psychological science:

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

---

[2] For a review focusing on Spener's second half of the book, specifically on the role of introspective access in self-management capacities, see Michel (forthcoming).

Mill, Wundt, James, and the underappreciated Müller. By rediscovering lessons from the past and reinterpreting them with her novel taxonomy, we can better disentangle concepts and better calibrate introspective methods. One of these lessons, specifically from Wundt's 'planned observation' approach, is that we can subject introspective methodology to 'Baconian' adequacy conditions of experimental evaluation:

(1) The observer should bring about at will the onset of the target phenomenon.
(2) The observer must be able to focus their attention on the target phenomenon.
(3) Results must be confirmed by repeated observations under the same conditions.
(4) Onset conditions must be determinable by varying accompanying conditions.

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

Spener emphasizes the problems with (2): attending a mental state as a target of introspection can change or even destroy it. This issue appears to make condition (2) — the observer must be able to focus their attention on the target phenomenon — impossible to fulfil: one cannot introspect without attending the target state and one alters the target state if one introspects. Despite a long tradition concerned about introspection's contamination risk, I think we should not worry too much about this issue. Both the size of the problem and its uniqueness to introspection have been largely exaggerated. Auguste Comte, for example, has strong words: '…this pretended direct contemplation of the mind by itself is a pure illusion… It is in fact evident that, by an invincible necessity, the human mind can observe directly all phenomena except its own proper states' (cited in James, 1890, p. 188). Comte's anguish stems from a simplistic model of the mind/brain: 'The thinker cannot divide himself into two, of whom one reasons whilst the other observes him reason. The organ observed and the organ observing being, in this case, identical, how could observation take place? This pretended psychological method is then radically null and void' (*ibid.*). The mind/brain is, we should not need to say this, more complex than a simple observational device.

The idea that the target of investigation — be this in perception, in scientific inquiry, or in introspection — must remain immutable during and after the observation is at best a caricature and at worse a misconception. As I have argued elsewhere (Morales, 2024), any detection mechanism such as our perceptual systems, a measuring

device, or introspection is bound to alter its target even if just minimally. Thermometers, for instance, will inevitably alter the temperature of the system they are tasked to measure. We may choose to ignore this because of the negligible size of the effect or precisely because it is unavoidable. But turning a blind eye is different from thinking that the act of measuring does not alter the target.

Perhaps more importantly, it is not true that we do not alter the targets of perception (to which introspection is often compared). It is probably true that eyes do not alter external objects. But we do not see with our eyes, we see with our minds and brains. And to see we need to attend, and attention most definitely alters perception (e.g. Carrasco, Ling and Read, 2004; Wright and Ward, 2008). Since psychologists and vision scientists usually do not think of this attentive modification as an unsurmountable problem, but rather as an interesting phenomenon to study, we should feel the same way about how mental states are altered during introspection. Retrospection — the more respected version of introspection in which the focusing of the mind on a mental state takes place *after* the experience as opposed to concurrently with it — is not significantly different. If introspection can alter a current mental state, retrospection can certainly alter a recently experienced state (or at least the way in which we remember it). So it is true that we probably cannot know what the unfiltered state of a mental state is like (on this, I am in agreement with Spener), but this is true across different types of measurement. Introspection, in this respect, is not special.

I find the third Baconian principle, in contrast to the second one, more interesting: 'Results must be confirmed by repeated observations under the same conditions.' How are experimenters supposed to confirm that repeated introspective observations are in fact the same introspective *act* as opposed to merely the same introspective *report*? Something that makes studying subjective experiences in general, and introspection in particular, so hard is that the scholar has no independent access to the object of study. The ground truth about mental states and their surrounding conditions are hidden from philosophers and scientists alike. By definition, subjective conscious states such as visual or painful experiences are only accessible to the subject of those experiences (and as discussed above, not even to them if what we have in mind is a pure, concurrent version of the mental state). Without third-person access to the ground truth, and without a trustworthy calibration about the reliability and accuracy of introspection, many philosophical and scientific endeavours are in peril.

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

How can a science of consciousness satisfy Bacon's third condition? How can the outcome of introspective reports be confirmed across repeated observations if we do not have an independent way of knowing the target states?

Some pessimistic researchers have indeed rejected the idea that introspective or subjective measures, as they are often called, can be validly used in research about the mind and, specifically, about consciousness (Irvine, 2012; 2013). As Spener puts it, 'critical rejection of this type of methodological approach turns out to be a foundational challenge to the field of research' (Spener, 2024, p. 70). The main reason is that 'the primary source of data about consciousness does seem to involve subjective access of some sort, hence subjective measures of consciousness seem necessary to a science of consciousness' (*ibid.*). But this poses 'a methodological conundrum: subjective measures of consciousness seem indispensable, yet too unreliable for the science of consciousness' (*ibid.*, p. 72). What to do?

Spener's proposal, as laid out in Chapter 6, suggests a middle ground position between the pessimistic outlook outlined above and a naïvely positive approach that takes introspective reports at face value. Her proposal includes specifying the limitations of each model of introspective access as well as the distinctive experimental procedures that they require. Moreover, she argues, we should avoid thinking that every mode is suitable for every mental phenomenon and that introspective data give evidence for general theoretical claims (*ibid.*, p. 117). Specific introspective access modes and specific experimental procedures are, at best, informative about specific experiences.

So far, so good. However, what the proposal has in breadth and insight, it suffers from practical guidance about how to implement it. I am most sympathetic to a middle ground approach: introspection is not infallible but it is not irredeemably flawed either (Chirimuuta, 2014; Morales, 2024; Spener, 2015). Limiting its functions to modes, states, and experimental practices is truly useful. What Spener does not tell us — not really — are the details of how we are meant to obtain these positive outcomes in a real experimental setting. In Chapter 5, she explores but ultimately rejects two common ways of getting subjects to produce introspective reports about their conscious experiences: via the Perceptual Awareness Scale or PAS (Overgaard and Sandberg, 2021; Ramsøy and Overgaard, 2004) and via confidence ratings (Cheesman and Merikle, 1984; 1986). I will leave the

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

Perceptual Awareness Scale to the side and focus on confidence ratings instead.[3]

Current consciousness research often uses confidence ratings as an index of consciousness and 'confidence ratings have long been held to be good indicators of consciousness episodes' (Spener, 2024, p. 89). Participants in psychophysics experiments typically report their own confidence in the correctness of their response in a task in which the experimenter knows the ground truth (e.g. perceptual and memory tasks). This, according to promoters of confidence ratings, makes them unlike traditional subjective (i.e. introspective) methods in that they do not aim to report experience directly. By eliciting confidence ratings there is hope to bypass the methodological conundrum highlighted above: the judgments are subjective, but experimenters have access to the ground truth (i.e. the stimuli presented to the subject and the correctness of their responses).

This hope notwithstanding, Spener's main objection against the use of confidence ratings as a way of infusing some degree of objectivity into subjectivity research is that 'there is a surprising lack of detail on how confidence ratings are revealing of any perceptual consciousness involved in type 1 performances' (*ibid.*, p. 91). Crucially, she thinks that 'the heart of the [confidence ratings] program seems to **rest on a brute intuition** about an epistemic connection between confidence ratings and conscious character of target perceptions' (*ibid.*, p. 92, my emphasis). Other than taking confidence ratings as a general indicator of the absence or presence of conscious perception, Spener continues, 'not much more is said about this crucial basing relation providing the nexus between confidence judgments and conscious perception' (*ibid.*, p. 92). In general, I do not disagree with Spener's frustration with scientists' underdeveloped views about the epistemic relationship between consciousness and confidence. However, I think we can ultimately overcome the scepticism about the usefulness of confidence ratings in consciousness research. It is possible that the reliance on confidence ratings was initially based on a 'brute intuition', but this does not reflect the current state of affairs anymore. There are practical and empirical reasons to support confidence ratings as a legitimate way of indexing conscious states.[4]

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

---

[3] For other criticisms about the Perceptual Awareness Scale, see Michel (2019).

[4] Here, I rehearse and expand some of the arguments laid out in Morales and Lau (2022). For further discussion, see Michel (2023).

One of the main advantages of using confidence ratings over subjective ratings of visibility is how easy it is to explain them to subjects. Unlike subjective reports where one has to appeal to visibility, the subjective character of experience, or — God forbid — qualia, confidence is straightforwardly defined as the subjective probability of being correct in the primary task (Morales and Lau, 2022; Norman and Price, 2015). While visibility reports are about subjective experiences that are inaccessible to the experimenter, when using confidence ratings there is a truth of the matter: confidence ratings predict (or not) the correctness of participants' responses, thus objectively tying confidence ratings to task performance and the external stimuli controlled by the experimenter. This link has been productively used in modern computational models that estimate metacognitive sensitivity (Fleming, 2017; Fleming and Lau, 2014; Galvin *et al.*, 2003; Maniscalco and Lau, 2012). This alone is already a remarkable advantage over subjective ratings, one that has produced successful results not just in consciousness studies, but also in fields assessing memory and perceptual decision-making (Rahnev *et al.*, 2020), political beliefs (Rollwage, Dolan and Fleming, 2018), and misinformation (Dobbs *et al.*, 2023), among many others.

The ease of instructing subjects to use confidence ratings provides an opportunity to avoid a significant amount of measurement noise. For example, when using subjective reports, participants may interpret quite differently what 'seeing something' or 'not seeing anything' means. This is the so-called 'criterion content problem' (Kahneman, 1968). Subjective reports about a stimulus necessarily are the result of certain unspecified criteria of what to focus on when reporting the experience. Put simply, what counts as 'seeing something' or 'not seeing anything' is often not explained to subjects and therefore not clear to them. And even if explained, the criterion of what counts as 'absent' or 'present' may still change across subjects or even for the same subject across different trials or tasks. This problem of criterion content, and more generally of response criterion, arises most notably in commonly used yes/no detection paradigms (Irvine, 2013; Nartker *et al.*, 2025; Phillips, 2021). This problem also emerges with measures such as the Perceptual Awareness Scale mentioned above, which asks subjects to characterize their experiences on a 1–4 scale with the labels 'no experience', 'weak experience', 'almost a clear experience', and 'a clear experience' (Michel, 2019; Overgaard and Sandberg, 2021; Ramsøy and Overgaard, 2004). Spener correctly highlights these thorny issues in her book.

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

Relying on confidence ratings is not without complications. Rating confidence may carry over or 'leak' from one trial to another (Rahnev *et al.*, 2015), and it may also prime confidence (Double and Birney, 2019a,b). Moreover, subjects are often not metacognitively perfect which renders them over- or under-confident for their level of performance (Maniscalco *et al.*, 2025; Rahnev, 2025). People may also base their confidence ratings on non-experiential evidence such as context, risk-aversion, pay-off schemes, or prior knowledge about the structure of the task or the stimulus properties (Hautus, Macmillan and Creelman, 2021; Locke *et al.*, 2020; Weise, 2025; Zylberberg, Barttfeld and Sigman, 2012).

Notwithstanding these issues, and in stark contrast with subjective ratings of awareness, confidence can be expressed on a well-calibrated and meaningful scale. This is true even if mapping subjective probabilities onto a specific scale introduces some noise. For example, subjects may never use the end points of the scale because they never feel they are just guessing but they also never feel they are absolutely sure about their answers. Even in a case like this, their intermediate ratings can be meaningfully re-scaled to just mean 'low' vs. 'high' confidence. In contrast, this cannot be meaningfully done with subjective ratings: if someone uses 'a weak experience' or 'an almost clear experience' throughout the experiment, these cannot be re-scaled to mean 'not seen' vs. 'seen' or 'no experience' vs. 'a clear experience'. Confidence ratings afford flexibility in data analysis to experimenters.

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

Another advantage of using confidence ratings in consciousness science is that probability in terms of percentage likelihood can be applied broadly and across domains. Probability of being correct is comparable and equally applicable across different tasks (e.g. detection, discrimination, recognition, etc.), modalities with different phenomenology (Faivre *et al.*, 2017; e.g. visual and auditory perception, memory, etc.; Gardelle, Corre and Mamassian, 2016; Mazancieux *et al.*, 2023), and different cognitive domains (e.g. perception and memory; Fitzgerald, Arvaneh and Dockree, 2017; Lund *et al.*, 2025; Mazancieux *et al.*, 2020; McCurdy *et al.*, 2013; Morales, Lau and Fleming, 2018; Rouault *et al.*, 2018; Rouault, Lebreton and Pessiglione, 2022). In contrast, subjective ratings are most applicable to single stimulus detection (and perhaps discrimination), and the response options obviously need to be modified depending on the nature of the task and the stimuli involved (e.g.

comparing two sets of different stimuli, detecting if something has changed or is missing, etc.).

Besides these practical considerations, empirically speaking, subjective ratings of consciousness and confidence ratings produce very similar behavioural results and the neural bases of both are intricately connected. This is perhaps the most crucial point against the idea that using confidence ratings is nothing but 'a brute intuition'. Behaviourally, visibility and confidence ratings have similar profiles. Peters and Lau (2015) obtained qualitatively identical results in an experiment where stimuli were rendered unconscious to different extents across trials by using a technique called metacontrast masking. Regardless of whether they asked subjects to rate their confidence on having seen the stimuli or to judge their visibility, the behavioural results looked effectively the same. Even researchers who have found (small) differences between subjective reports and confidence ratings in special conditions (e.g. very low contrast) admit that 'there was a considerable association between the two ratings that were required after each trial, indicating that the patterns of the ratings are quite similar' (Zehetleitner and Rausch, 2013, p. 1423).

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

From a neuroscientific perspective, the differences between these two types of subjective appraisals are not easily detectable. The underlying neural dynamics and neural mechanisms supporting different types of subjective reflection about one's experience largely overlap. First, certain features of the brain's global dynamics affect visibility and confidence ratings in a similar way. Spontaneous low frequency brain oscillations (< 30 Hz) reflect (or perhaps affect) neuronal excitability and, with it, performance and type II ratings during psychophysical tasks (Samaha *et al.*, 2020). In particular, low frequency oscillations with lower pre-stimulus power (i.e. oscillations of lower magnitude right before the presentation of a stimulus) bias observers to report both higher confidence and higher subjective visibility (Benwell *et al.*, 2017; Samaha, Iemi and Postle, 2017). Second, many studies using different types of neuroimaging techniques across different species have consistently found astonishingly similar neural correlates of consciousness. The prefrontal cortex (PFC), specifically certain areas in the dorsolateral and orbitofrontal PFC, has been found to support subjective reports of awareness (Cul *et al.*, 2009; Lau and Passingham, 2006), visibility ratings (Rounis *et al.*, 2010), and confidence ratings both in animals (Miyamoto *et al.*, 2017; 2018) and humans (Cortese *et al.*, 2016; Fleming, Huijgen and Dolan, 2012; Morales, Lau and Fleming, 2018). Importantly, these findings likely

reflect activity primarily associated with the underlying perceptual experience rather than with the mere act of reporting it (Hatamimajoumerd *et al.*, 2022; Michel and Morales, 2020).

Subjective assessments of confidence and subjective (i.e. introspective) reports of one's awareness share a behavioural and neural profile. This is probably more than a mere coincidence, of course. Explicit reports of confidence heavily rely on conscious experiences. A subject may be confident in what they saw *because* they saw it clearly, or they may have low confidence in the accuracy of their performance *because* the stimulus was not clear or they were not conscious of it. This is not only an intuitive connection, but one that has support in experimental results. Vlassova, Donkin and Pearson (2014), for example, found that adding unconscious information in a visual task improves performance but it does not alter confidence or metacognitive sensitivity. In other words, to be more confident in your improved performance, the reason for the improvement must enter your conscious awareness.

This is part of a more general point: the strength of our conscious experiences influences the performance of our self-reflective mechanisms. Whether it is confidence or visibility, a clearer, stronger experience is more likely than not to produce (or at least be associated with) an accurate, reliable judgment. What I call 'mental strength' — the intensity of conscious experiences — can be thought of as an input to the introspective mechanism (Morales, 2023; 2024). The stronger the experience, the more likely it is that an accurate introspective judgment will take place and the more likely that it will yield a high-confidence judgment too. An incredibly valuable contribution to the field by Spener is the idea that we need to calibrate introspection (Spener, 2015; 2024). I think, however, that a triangulation between the intensity of experience, subjects' introspective accuracy, and the confidence in their performance is a crucial part of the solution to the 'methodological conundrum' of consciousness science. First, the intensity of experience is heavily modulated by the intensity of external stimuli, which is under experimenters' control and therefore relatively 'objective'. Second, confidence ratings have objective properties that allow for external validation (the ground truth is known during an experiment and correctness conditions can be unambiguously established). Third, confidence ratings are tightly linked to our introspective access to conscious experiences. It is this last bridge that makes confidence ratings a valuable resource. Subjective judgments and confidence reports can come apart, in principle,

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

but in practice they go in tandem. This is why a science of consciousness can harmoniously use confidence ratings as proxies for subjective visibility ratings.

Let us take stock. In her book *Introspection*, Spener has offered us a fantastic recipe for dissecting and taxonomizing introspection (both the capacity and the method). She also points us in what I think is exactly the right direction: a middle-ground conception of introspection. Introspection is neither infallible, nor completely untrustworthy. This is a crucial lesson that both historical and contemporary philosophical approaches often overlook. Notwithstanding this and many other virtues of the book, one point where Spener seemed to move too fast was in dismissing confidence as a promising bridging approach for the science of consciousness. With their subjective character and their objective properties, confidence ratings can really offer a tractable approach with specific mechanistic and computational details to achieve the middle ground Spener's book promotes. The science of confidence and metacognition is buoying, with special issues, workshops, and entire labs dedicated to studying it. Surely philosophers can and should help flesh out better the connections between confidence reports and subjective measures of consciousness. The work is never done, but the theoretical foundations are laid out and the extant empirical evidence points towards their deeply entangled nature.

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

References

Benwell, C.S.Y., Tagliabue, C.F., Veniero, D., Cecere, R., Savazzi, S. & Thut, G. (2017) Prestimulus EEG Power predicts conscious awareness but not objective visual performance, *eNeuro*, **4** (6). doi: 10.1523/eneuro.0182-17.2017

Carrasco, M., Ling, S. & Read, S. (2004) Attention alters appearance, *Nature Neuroscience*, **7** (3), pp. 308–313. doi: 10.1038/nn1194

Cheesman, J. & Merikle, P.M. (1984) Priming with and without awareness, *Perception & Psychophysics*, **36** (4), pp. 387–395. doi: 10.3758/bf03202793

Cheesman, J. & Merikle, P.M. (1986) Distinguishing conscious from unconscious perceptual processes, *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, **40** (4), pp. 343–367. doi: 10.1037/h0080103

Chirimuuta, M. (2014) Psychophysical methods and the evasion of introspection, *Philosophy of Science*, **81** (5), pp. 914–926. doi: 10.1086/677890

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

Cul, A.D., Dehaene, S., Reyes, P., Bravo, E. & Slachevsky, A. (2009) Causal role of prefrontal cortex in the threshold for access to consciousness, *Brain*, **132** (9), pp. 2531–2540. doi: 10.1093/brain/awp111

Dobbs, M., DeGutis, J., Morales, J., Joseph, K. & Swire-Thompson, B. (2023) Democrats are better than Republicans at discerning true and false news but do not have better metacognitive awareness, *Communications Psychology*, **1** (1), art. 46. doi: 10.1038/s44271-023-00040-x

Double, K.S. & Birney, D.P. (2019a) Do confidence ratings prime confidence?, *Psychonomic Bulletin & Review*, **26** (3), pp. 1035–1042. doi: 10.3758/s13423-018-1553-3

Double, K.S. & Birney, D.P. (2019b) Reactivity to measures of metacognition, *Frontiers in Psychology*, **10**, art. 2755. doi: 10.3389/fpsyg.2019.02755

Faivre, N., Filevich, E., Solovey, G., Kuhn, S. & Blanke, O. (2017) Behavioural, modeling, and electrophysiological evidence for supramodality in human metacognition, *Journal of Neuroscience*, **38** (2), pp. 263–277. doi: 10.1523/jneurosci.0322-17.2017

Fitzgerald, L.M., Arvaneh, M. & Dockree, P.M. (2017) Domain-specific and domain-general processes underlying metacognitive judgments, *Consciousness and Cognition*, **49**, pp. 264–277. doi: 10.1016/j.concog.2017.01.011

Fleming, S.M. (2017) HMeta-d: Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings, *Neuroscience of Consciousness*, **2017** (1), nix007. doi: 10.1093/nc/nix007

Fleming, S.M., Huijgen, J. & Dolan, R.J. (2012) Prefrontal contributions to metacognition in perceptual decision making, *Journal of Neuroscience*, **32** (18), pp. 6117–6125. doi: 10.1523/jneurosci.6489-11.2012

Fleming, S.M. & Lau, H. (2014) How to measure metacognition, *Frontiers in Human Neuroscience*, **8**, art. 443. doi: 10.3389/fnhum.2014.00443

Franco, K.S. (2024) Pluralism about introspection, *Mind & Language*, **39** (2), pp. 293–309. doi: 10.1111/mila.12483

Galvin, S.J., Podd, J.V., Drga, V. & Whitmore, J. (2003) Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions, *Psychonomic Bulletin & Review*, **10** (4), pp. 843–876, [Online], http://link.springer.com/article/10.3758/BF03196546.

Gardelle, V. de, Corre, F.L. & Mamassian, P. (2016) Confidence as a common currency between vision and audition, *PLoS ONE*, **11** (1), e0147901-11. doi: 10.1371/journal.pone.0147901

Hatamimajoumerd, E., Murty, N.A.R., Pitts, M. & Cohen, M.A. (2022) Decoding perceptual awareness across the brain with a no-report fMRI masking paradigm, *Current Biology*, **32** (19), pp. 4139–4149. doi: 10.1016/j.cub.2022.07.068

Hautus, M.J., Macmillan, N.A. & Creelman, C.D. (2021) *Detection Theory: A User's Guide*, 3rd ed., London: Routledge. doi: 10.4324/9781003203636-17

Irvine, E. (2012) *Consciousness as a Scientific Concept: A Philosophy of Science Perspective*, Dordrecht: Springer Netherlands. doi: 10.1007/978-94-007-5173-6

Irvine, E. (2013) Measures of consciousness, *Philosophy Compass*, **8** (3), pp. 285–297. doi: 10.1111/phc3.12016

James, W. (1890) *The Principles of Psychology*, New York: Dover.

Kahneman, D. (1968) Method, findings, and theory in studies of visual masking, *Psychological Bulletin*, **70** (6), pp. 404–425. doi: 10.1037/h0026731

Lau, H. & Passingham, R.E. (2006) Relative blindsight in normal observers and the neural correlate of visual consciousness, *Proceedings of the National Academy of Sciences*, **103** (49), pp. 18763–18768. doi: 10.1073/pnas.0607716103

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

Locke, S.M., Gaffin-Cahn, E., Hosseinizaveh, N., Mamassian, P. & Landy, M.S. (2020) Priors and payoffs in confidence judgments, *Attention, Perception, & Psychophysics*, **82** (6), pp. 3158–3175. doi: 10.3758/s13414-020-02018-x

Lund, A.E., Correa, C., Fardo, F., Fleming, S.M. & Allen, M. (2025) Domain generality in metacognitive ability: A confirmatory study across visual perception, memory, and general knowledge, *Journal of Experimental Psychology: Learning, Memory and Cognition*. doi: 10.31234/osf.io/p97s5

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

Maniscalco, B., Castaneda, O.G., Odegaard, B., Morales, J., Rajananda, S., Denison, R. & Peters, M.A.K. (2025) The relative psychometric function: A general analysis framework for relating psychological processes, *PsyArXiv*. doi: 10.31234/osf.io/5qrjn_v2

Mazancieux, A., Fleming, S.M., Souchay, C. & Moulin, C.J.A. (2020) Is there a G factor for metacognition? Correlations in retrospective metacognitive sensitivity across tasks, *Journal of Experimental Psychology: General*, **149** (9), pp. 1788–1799. doi: 10.1037/xge0000746

Mazancieux, A., Pereira, M., Faivre, N., Mamassian, P., Moulin, C.J.A. & Souchay, C. (2023) Towards a common conceptual space for metacognition in perception and memory, *Nature Reviews Psychology*, **2** (12), pp. 751–766. doi: 10.1038/s44159-023-00245-1

McCurdy, L.Y., Maniscalco, B., Metcalfe, J., Liu, K.Y., Lange, F.P. de & Lau, H. (2013) Anatomical coupling between distinct metacognitive systems for memory and visual perception, *Journal of Neuroscience*, **33** (5), pp. 1897–1906. doi: 10.1523/jneurosci.1890-12.2013

Michel, M. (2019) The mismeasure of consciousness: A problem of coordination for the Perceptual Awareness Scale, *Philosophy of Science*, **86** (5), pp. 1239–1249.

Michel, M. (2023) Confidence in consciousness research, *Wiley Interdisciplinary Reviews: Cognitive Science*, **14** (2), e1628. doi: 10.1002/wcs.1628

Michel, M. (forthcoming) Review of *Introspection: First-Person Access in Science and Agency*, by Maja Spener, *Mind*.

Michel, M. & Morales, J. (2020) Minority reports: Consciousness and the prefrontal cortex, *Mind & Language*, **35** (4), pp. 493–513. doi: 10.1111/mila.12264

Miyamoto, K., Osada, T., Setsuie, R., Takeda, M., Tamura, K., Adachi, Y. & Miyashita, Y. (2017) Causal neural network of metamemory for retrospection in primates, *Science*, **355** (6321), pp. 188–193. doi: 10.1126/science.aal0162

Miyamoto, K., Setsuie, R., Osada, T. & Miyashita, Y. (2018) Reversible silencing of the frontopolar cortex selectively impairs metacognitive judgment on non-experience in primates, *Neuron*, **97** (4), pp. 980–989. doi: 10.1016/j.neuron.2017.12.040

Morales, J. (2023) Mental strength: A theory of experience intensity, *Philosophical Perspectives*, **37** (1), pp. 248–268. doi: 10.1111/phpe.12189

Morales, J. (2024) Introspection is signal detection, *The British Journal for the Philosophy of Science*, **75** (1), pp. 99–126. doi: 10.1086/715184

Morales, J. & Lau, H. (2022) Confidence tracks consciousness, in Weisberg, J. (ed.) *Qualitative Consciousness: Themes from the Philosophy of David Rosenthal*, pp. 91–108, Cambridge: Cambridge University Press.

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

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

Nartker, M., Firestone, C., Egeth, H. & Phillips, I. (2025) Sensitivity to visual features in inattentional blindness, *eLife*, **13**, RP100337. doi: 10.7554/elife.100337.3

Norman, E. & Price, M.C. (2015) Measuring consciousness with confidence ratings, in Overgaard, M. (ed.) *Behavioral Methods in Consciousness Research*, pp. 159–180, Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199688890.003.0010

Overgaard, M. & Sandberg, K. (2021) The Perceptual Awareness Scale — recent controversies and debates, *Neuroscience of Consciousness*, **1**, niab044. doi: 10.1093/nc/niab044

Peters, M.A.K. & Lau, H. (2015) Human observers have optimal introspective access to perceptual processes even for visually masked stimuli, *eLife*, **4**, e09651. doi: 10.7554/elife.09651

Phillips, I. (2021) Blindsight is qualitatively degraded conscious vision, *Psychological Review*, **3** (128), pp. 558–584. doi: 10.1037/rev0000254

Rahnev, D. (2025) A comprehensive assessment of current methods for measuring metacognition, *Nature Communications*, **16** (1), art. 701. doi: 10.1038/s41467-025-56117-0

Rahnev, D., Koizumi, A., McCurdy, L.Y., D'Esposito, M. & Lau, H. (2015) Confidence leak in perceptual decision making, *Psychological Science*, **26** (11), pp. 1664–1680. doi: 10.1177/0956797615595037

Rahnev, D., *et al.* (2020) The confidence database, *Nature Human Behaviour*, **4** (3), pp. 317–325. doi: 10.1038/s41562-019-0813-1

Ramsøy, T.Z. & Overgaard, M. (2004) Introspection and subliminal perception, *Phenomenology and the Cognitive Sciences*, **3** (1), pp. 1–23. doi: 10.1023/b:phen.0000041900.30172.e8

Renero, A. (2019) Modes of introspective access: A pluralist approach, *Philosophia*, **47** (3), pp. 823–844. doi: 10.1007/s11406-018-9989-2

Rollwage, M., Dolan, R.J. & Fleming, S.M. (2018) Metacognitive failure as a feature of those holding radical beliefs, *Current Biology*, **28** (24), pp. 4014–4021. doi: 10.1016/j.cub.2018.10.053

Rouault, M., McWilliams, A., Allen, M.G. & Fleming, S.M. (2018) Human metacognition across domains: Insights from individual differences and neuroimaging, *Personality Neuroscience*, **1**, e17. doi: 10.1017/pen.2018.16

Rouault, M., Lebreton, M. & Pessiglione, M. (2022) A shared brain system forming confidence judgment across cognitive domains, *Cerebral Cortex*, **33** (4), pp. 1426–1439. doi: 10.1093/cercor/bhac146

Rounis, E., Maniscalco, B., Rothwell, J.C., Passingham, R.E. & Lau, H. (2010) Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness, *Cognitive Neuroscience*, **1** (3), pp. 165–175. doi: 10.1080/17588921003632529

Samaha, J., Iemi, L. & Postle, B.R. (2017) Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy, *Consciousness and Cognition*, **54** (12), pp. 47–55. doi: 10.1016/j.concog.2017.02.005

Copyright (c) Imprint Academic
For personal use only -- not for reproduction

Samaha, J., Iemi, L., Haegens, S. & Busch, N.A. (2020) Spontaneous brain oscillations and perceptual decision-making, *Trends in Cognitive Sciences*, **24** (8), pp. 639–653. doi: 10.1016/j.tics.2020.05.004

Spener, M. (2015) Calibrating introspection, *Philosophical Issues*, **25** (1), pp. 300–321. doi: 10.1111/phis.12062

Spener, M. (2018) Introspecting in the 20th century, in Kind, A. (ed.) *Philosophy of Mind in the Twentieth and Twenty-First Centuries*, pp. 148–174, London: Routledge.

Spener, M. (2024) *Introspection: First-Person Access in Science and Agency*, Oxford: Oxford University Press. doi: 10.1093/oso/9780198867449.003.0007

Vlassova, A., Donkin, C. & Pearson, J. (2014) Unconscious information changes decision accuracy but not confidence, *Proceedings of the National Academy of Sciences*, **111** (45), pp. 16214–16218. doi: 10.1073/pnas.1403619111

Weise, L. (2025) Payoff designs in post-decision wagering: A systematic review, *Metacognition and Learning*, **20** (1), art. 6. doi: 10.1007/s11409-024-09411-z

Wright, R.D. & Ward, L.M. (2008) *Orienting of Attention*, Oxford: Oxford University Press.

Zehetleitner, M. & Rausch, M. (2013) Being confident without seeing: What subjective measures of visual consciousness are about, *Attention, Perception, & Psychophysics*, **75** (7), pp. 1406–1426. doi: 10.3758/s13414-013-0505-2

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

Copyright (c) Imprint Academic
For personal use only -- not for reproduction