

infected rhesus macaques. *bioRxiv* <https://doi.org/10.1101/2020.03.13.990226>.

17. McCray Jr., P.B., Pewe, L., Wohlford-Lenane, C., Hickey, M., Manzel, L., Shi, L., Netland, J., Jia, H.P., Halabi, C., Sigmund, C.D., *et al.* (2007). Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. *J. Virol.* **81**, 813–821.
18. Docherty, A.B., Harrison, E.M., Green, C.A., Hardwick, H.E., Pius, R., Norman, L., Holden, K.A., Read, J.M., Dondelinger, F., Carson, G., *et al.* (2020). Features of 20,133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* **369**, m1985.
19. Mao, L., Jin, H., Wang, M., Hu, Y., Chen, S., He, Q., Chang, J., Hong, C., Zhou, Y., Wang, D., *et al.* (2020). Neurologic manifestations of hospitalized patients with coronavirus disease 2019 in Wuhan, China. *JAMA Neurol.* **77**, 1–9.
20. Poyiadji, N., Shahin, G., Noujaim, D., Stone, M., Patel, S., and Griffith, B. (2020). COVID-19-associated acute hemorrhagic necrotizing encephalopathy: CT and MRI features. *296*, E119–E120.
21. Klok, F.A., Kruip, M., van der Meer, N.J.M., Arbous, M.S., Gommers, D., Kant, K.M., Kaptein, F.H.J., van Paassen, J., Stals, M.A.M., Huisman, M.V., *et al.* (2020). Incidence of thrombotic complications in critically ill ICU patients with COVID-19. *Thromb. Res.* **197**, 145–147.

¹Radboud University, 6525 XZ Nijmegen, The Netherlands. ²University Medical Center, Utrecht Brain Center, Utrecht University, 3584 CG Utrecht, The Netherlands.

³Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands. ⁴Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands. ⁵Maastricht University, 6211 LK Maastricht, The Netherlands. ⁶University of Groningen, 9712 CP Groningen, The Netherlands. ⁷University of Groningen, University Medical Center, 9713 GZ Groningen, The Netherlands.

⁸Biomedical Primate Research Center, 2288 GJ Rijswijk, The Netherlands. ⁹Metris BV, 2132 NG Hoofddorp, The Netherlands.

¹⁰University Medical Center, 3584 CX Utrecht, The Netherlands. ¹¹Leiden University Medical Center, 2333 ZA Leiden, The Netherlands.

¹²Amsterdam University Medical Centers, 1105 AZ Amsterdam, The Netherlands.

¹³Erasmus University Medical Center, 3015 GD Rotterdam, The Netherlands.

¹⁴Netherlands Institute for Neuroscience, Royal Netherlands Academy of Arts and Sciences, 1105 BA Amsterdam, The Netherlands. ¹⁵Swammerdam Institute for Life Sciences, University of Amsterdam, 1098 XH Amsterdam, The Netherlands. ¹⁶Charles River, 5231 DD Den Bosch, The Netherlands.

¹⁷Onco Institute, 3521 AL Utrecht, The Netherlands. ¹⁸Wageningen University, 6708 PB Wageningen, The Netherlands. ¹⁹Utrecht University, 3512 JE Utrecht, The Netherlands.

²⁰Life and Medical Sciences Institute, University of Bonn, 53115 Bonn, Germany.

²¹Netherlands Institute of Ecology, 6708 PB Wageningen, The Netherlands. ²²Viroclinics Xplore, 5374 RE Schaijk, The Netherlands.

²³Wageningen Bioveterinary Research, 8221 RA Lelystad, The Netherlands.

*E-mail: I.genzel@donders.ru.nl; Judith.Homberg@radboudumc.nl

My Word

Seeing consciousness through the lens of memory

Joseph E. LeDoux^{1,2,3,*}
and Hakwan Lau^{4,5}

We humans have long thought of ourselves in terms of bodily and mental spheres of existence. These days many of us understand that the mental aspect of who we are is embedded in the part of the body known as the brain, and is therefore also part of our physical, bodily self. Although most of us accept this scientific conclusion about the physicality of the mind, many nevertheless feel as though their mind possesses some quality lacking in other physical systems within their body, and even within their brain — we have first-hand knowledge of our thoughts and feelings, but not of the neural processes that control body processes related to digestion, respiration, heart rhythm, and so on, nor of the brain's control of much of our behavior.

In this My Word, we explore the question of what is different about a brain state about which you are conscious from one that you are not? We will argue that all conscious experiences involve elements of memory and meta-representation. This framework may help us better understand the subjective qualitative character of conscious experiences, as well as why meta-representations may be involved in generating even the simplest of such experiences.

Phenomenal ‘feel’ and conscious content

William James famously referred to the special property of conscious states in terms of ‘warmth and intimacy’. These days it is common to describe this as a phenomenal ‘feeling’ — the quality that makes red seem red, an apple seem like an apple, and fear feel fearful [1]. There is considerable disagreement about what this quality is and how it may come about in the brain in relation to the actual content;

what the experience is about, for example, an apple or a snake.

First-order theories, such as recurrent processing theory [2,3], posit that consciousness originates in brain regions specialized in the processing of a given kind of information. For perceptual states of consciousness, these include, for instance, visual or auditory cortices. For emotions like fear, subcortical areas such as the amygdala have been proposed to be a first-order locus [4]. In first-order theories, the phenomenal feel and content of the experience are fully accounted for by some specific pattern of neural activity within these first-order areas.

Higher-order theories, on the other hand, suggest that first-order representations may not be sufficient to account for either the phenomenal feel or the content of the conscious experience [5–7]. They posit that some higher-order cognitive mechanism, possibly involving circuits in prefrontal cortex, is needed in order to monitor or meta-represent the first-order information. As such, the information represented in first-order structures remains non-conscious (or pre-conscious) if no higher-order re-representation is involved. With respect to fear, this view posits that the amygdala controls non-conscious defensive responses, such as behavioral and physiological responses to threats, but that higher-order re-representation, possibly involving prefrontal cortex, is required in order to generate the subjective experience of fear in response to the threatening stimulus [7–9]. For higher-order theories, the phenomenal quality and conscious content both require the higher-order state.

Memory and consciousness

The pioneering 19th century psychologist Hermann von Helmholtz proposed that conscious perception is an ‘unconscious inference’ based in part on memory. Consistent with von Helmholtz’s logic, we propose that all states of consciousness depend on memory [7–9], and specifically on mechanisms that integrate sensory and memory information unconsciously (i.e. pre-consciously).

A fundamental distinction in memory research is between



Table 1. Anoetic, noetic, and auto-noetic consciousness.

Consciousness	Knowledge	Memory	Meta-representation	Self	Example thought
Anoetic	Non-knowing	Procedural	Non-cognitive	Implicit	Thought not involved
Noetic	Fact-knowing	Semantic	Cognitive	Assumed	"There is an apple"
Auto-noetic	Self-knowing	Episodic	Cognitive	Explicit	"I see an apple"

memories that are formed and stored in a way that can be consciously retrieved, and memories that are stored implicitly, and hence are not consciously experienced when retrieved [10]. People can have conscious introspections and talk about explicit memories, but they can only show the effects of past implicit procedural learning through nonverbal behavior. One consequence is that a person with damage to the circuits underlying explicit memories can procedurally acquire and perform behaviors which they have no conscious memory of having acquired; conversely, damage to one of the many procedural memory circuits prevents behavioral learning by that system, but spares the ability to remember the episode in which they tried to learn the behavioral response.

An influential view of the relation between memory and consciousness was developed by Endel Tulving [11]. He identified three kinds of conscious states that are each associated with a distinct kind of memory (Table 1). Although Tulving's taxonomy was proposed to account for how memory contributes to consciousness, we believe it is also useful as a way of using memory to partition conscious states and thereby better understand consciousness itself.

Auto-noetic consciousness is explicitly about your self in the present moment in relation to your past and future. It is grounded in episodic memories: memories about experiences you have had. The episodic record of your life provides a sense of continuity with your personal past (recollecting seeing a robin on a certain day and in a particular place); it also allows one to project into possible personal futures (wondering what kinds of birds you might see on your next outing). This ability to visit your personal past and hypothetical future is called mental time travel [11].

Noetic states of consciousness depend on semantic memories of factual and conceptual information about what things are and are not. For example, seeing a bird and recognizing it as a robin not only requires sensory processing of the visual properties of the present robin, but also semantic, including conceptual, memories about what birds and robins are. In sharp contrast to auto-noesis, noetic states do not require explicit self-awareness and do not support mental time travel.

A third kind of state recognized by Tulving is anoetic awareness. Although the most basic, it is the least intuitive, and requires some discussion. Tulving referred to anoetic consciousness as states of 'non-knowing', by which he meant that these do not themselves directly rely on or contribute to our conceptual, explicit knowledge, and hence are not readily available for introspection. They are conceptually similar to what first-order theorists refer to as 'phenomenal consciousness' [2,3], and have also been described as 'non-reflective qualia' that occur in primitive emotional states related to survival behaviors involving defense, mating, feeding and so on [4]. They exist on the 'fringe' or 'penumbra' of consciousness, to borrow expressions from William James.

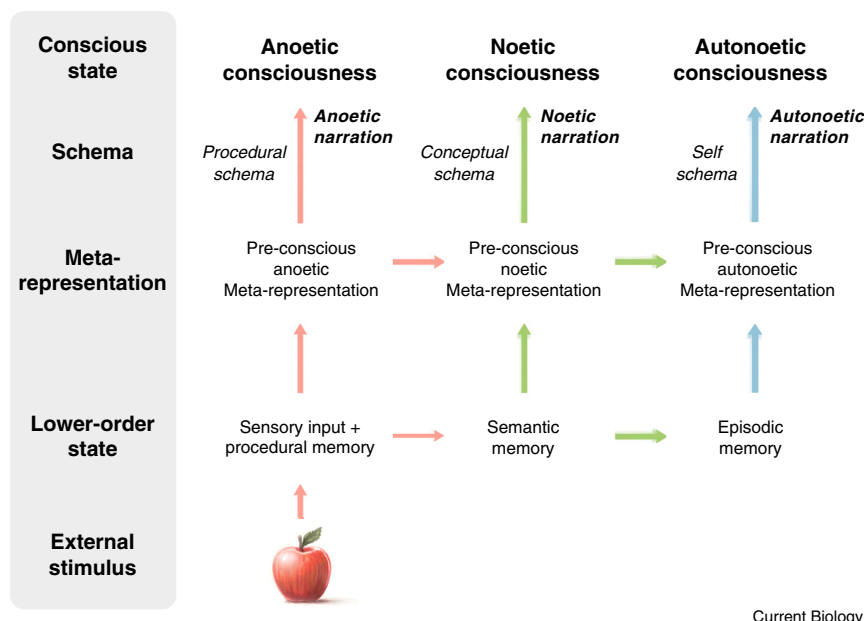
Although Tulving and others describe anoetic states as temporally and spatially bound to the current moment, and mainly about the here and now, we believe this needs qualification. This is true of the sensory component of such states — but the involvement of implicit, procedural memories gives the moment a connection to relevant aspects of one's implicit history with similar stimuli [11–13]. For example, through past exposures to wavelengths within a certain range, the visual cortex, via synaptic plasticity, has implicitly learned

relations between dynamical neural profiles and sensory inputs. These implicit memories are used to classify the wavelength of the present stimulus with other stimuli from the past with similar neural profiles. Indeed, empirical findings show that colors and other low-level visual features are embedded in a local spatial and temporal contexts of other neurons and are processed relative to these [14]. The color of a stimulus, when not conceptualized by semantic memory, can only be experienced relationally with respect to other possible color representations, which is what accumulated procedural memories offer. Accordingly, what we consciously see as red only looks red because of the way it looks similar to and different from other stimuli that have been seen in the past [13,14].

To illustrate the relationship between these three kinds of conscious states, consider the experience of seeing a red apple (Figure 1). An anoetic experience of 'redness', lacking the conceptual support of semantic memory, remains as a sensation of some complex set of wavelengths classified in relation to past experiences with similar wavelengths stored via procedural memories. With semantic memory, the wavelengths become noetic states in which you are informed by what you conceptually understand 'red' to be, and how red relates to other features, such as shape, allowing an awareness that a red apple is present. If your self, via episodic memory, is explicitly brought into the experience, an auto-noetic state of a red apple exists, perhaps involving the recollection of some experience involving apples from your past, or anticipating an experience in your future.

Meta-representations and consciousness

Each of Tulving's three conscious states have been said to be based on a pre-conscious meta-representation



Current Biology

Figure 1. Meta-representations, narrations, schema, and consciousness.

The model developed in this paper involves three key ideas. First, conscious experiences of events in the world involve lower-order elements of memory that are meta-represented and filtered by schemata that narrate the content of the conscious experience. Second, the lower-order state, meta-representation, schema, narrative, and content differ for anoetic, noetic, and auto-noetic experiences. Third, the states are hierarchical, such that the implicit procedural memory element of anoetic consciousness is passed on to noetic and auto-noetic states, providing these explicit conscious experiences with a ‘phenomenal feel’ and sense of ownership.

[12]. As such, Tulving’s model can be thought of as a kind of higher-order theory in which conscious awareness arises from the re-representation of a lower-order state by a higher-order one [5–9,12]. In other words, as in higher-order theory, the meta-representations are pre-conscious antecedents to conscious experiences.

In the case of anoetic consciousness, the lower-order state is not simply a pure sensory state of visual cortex, as claimed by lower-order theories. It is, as noted, a representation in which raw sensory input has been compared with implicit procedural memories from similar past visual experiences. In further contrast to first-order theories, we suggest that even an anoetic, non-reflective, conscious experience of red depends on processing beyond visual cortex. For example, in order to experience red as being about a stimulus in the world the brain must distinguish it from other kinds of visual cortex states based on internal visual imagery or neural noise [15].

For this to happen, some additional mechanisms may have to implicitly retrieve, perhaps via an implicit schema (see below), the dynamical profiles of early sensory activity, in order to draw the correct perceptual inference. In this sense, we propose that meta-representations contribute to even the simplest perception of non-conceptualized redness [6,7,12].

Noetic experiences add semantic knowledge stored in medial and lateral cortical areas of the temporal lobe [8,10] to the anoetic state. This creates conceptualized perceptual content about what the stimulus is, and how it relates to the present situation. If you have the conscious thought *that* you are seeing the redness of an apple it is because you are noetically experiencing the content of the meta-representation.

For auto-noetic states, episodic memories about relevant past personal experiences stored in the temporal lobe are added to the noetic state. Through the mental time-travel feature of episodic memory, we feel consciously connected to our self

over time, including our past, and our possible future, selves.

In summary, the three kinds of meta-representations seem to serve the same purpose — they appear to function as higher-order re-representations of lower-order sensory and memory information. If correct, the implication is that all conscious experiences are preceded by a non-conscious meta-representation, without which there is no consciousness. We suggest that these meta-representations involve areas of prefrontal cortex, possibly including dorsal and ventral lateral prefrontal cortex, the frontal pole, and medial frontal areas [6–9,15–17]. Next, we consider how these pre-conscious meta-representations might be translated into conscious content.

The stories we tell ourselves

Novelists, inspired by William James’ description of ‘the stream of consciousness’, have capitalized on the narrative, story-like quality of the human mind. Scientists, too, have followed this lead [18–20], noting that one’s self has a narrative-like structure that gives us psychological coherence over time. This coherence is probably maintained by perpetual revisions, with the results carried forward through memory. For example, if a new noetic or auto-noetic experience is discordant with stored semantic or episodic memories, either the anomaly has to be accepted or the stored knowledge has to be updated. Although the term ‘narrative’, when used in relation to mental states, most commonly refers to a verbal story line told by our inner voice, our conscious experiences can be visual or even multimodal, and the underlying narratives, in turn, would be as well.

A common assumption is that narrations are outcomes of conscious states — they occur post-consciously. But we propose that narratives are also involved in the initial generation of the content of noetic and auto-noetic consciousness. In particular, we suggest narratives reflect the complex memory processes known as schema [8,21]. These are collections of semantic memories about recurring objects and situations that serve as non-conscious conceptual templates for storing new memories and for

understanding present situations in relation to one's needs and goals [7–9,21]. Of interest is that the ventral medial prefrontal cortex has a role in both schema [8,21] and narrations [22]. This area interconnects with other prefrontal areas that have been implicated in metacognition and consciousness, such as the dorsal lateral frontal cortex and the frontal pole [6–9,15–17,23,24]. We hypothesize that when the relevant prefrontal circuitry is damaged, pre-conscious narrations and the contents of conscious experiences should be affected.

Owning up

The preceding discussion begs the question of how anoetic states fit into our narrative framework of noetic and auto-noetic consciousness. We suggest that anoetic states may play a crucial role in the feeling of ownership we have about our mental states.

Mental state ownership is so natural and automatic that we never have to explicitly affirm it. This fact is put into stark relief by certain neurological or psychiatric conditions in which people lose the sense that their thoughts and feelings belong to them [25]. A common view is that this personal relationship to mental states is due to autobiographical memory, which includes episodic and semantic memories about our self. Certainly, our autobiographical memories, especially episodic ones, are important in giving us a sense of continuity over time. But there may be more to the sense of ownership than that.

As we noted, procedural memories give anoetic states a connection to implicit aspects of one's history that are relevant to the present stimulus situation. The fringe or penumbral quality of these goes unnoticed because anoetic states are often embedded in noetic and/or auto-noetic states with actual conscious content [4]. It is mainly in laboratory studies where explicit consciousness involvement is minimized that a pure anoetic experience can be isolated. But in a real-life situation the procedural memories, and their implicit connection to your past, may be what makes a noetic or auto-noetic state feel it is, by mere acquaintance [25], yours without you ever having to explicitly affirm that.

Now consider this idea in the context of a complex real-life situation in which you are not just seeing a stimulus but are in a life-threatening encounter with a dangerous stimulus such as a rattlesnake. Sensory information about the snake will not simply be transmitted from visual cortex areas to memory and higher cognitive circuits. It will also be transmitted to subcortical circuits, such as those involving the amygdala, that organize defensive behaviors and physiological reactions, and that also activate arousal systems. The behavioral and physiological responses also produce sensory and chemical signals that are represented in the brain as body states [4,8,9,26]. Importantly, these circuits have also acquired procedural memories about the relevant neural responses made in the past. These will determine, for example, how the present degree of activation of amygdala, arousal, or body state circuits match levels from the past. Still other implicit circuits are also likely activated. For example, object recognition or auditory pattern recognition circuits might be primed to detect related stimuli, and motor circuits primed to respond behaviorally.

All of these implicit states will become part of an anoetic background monitored by the meta-representation that also includes semantic conceptual content about snakes and episodic content about your personal experiences with snakes. The content of the meta-representation as a whole, we suggest, will be schematized (implicitly by procedural memory and explicitly by semantic memory) and narrated [27] (Figure 1). Hence you will find yourself in the auto-noetic state of fear [8,9,27], without having to explicitly infer that fear is what you are feeling. You know it is fear because you know what fear feels like to you. Other people do not have your first-hand procedural memories of your neurophysiological responses to threats that accompany your auto-noetic states of fear narrated by your schema. That's why emotions are personal — only you can have your emotions [8,9,27]. But through shared vocabulary, we can extrapolate and infer what others *might* feel in similar situations [8,27]. Complicating our

ability to understand the emotional lives of others is the fact that emotion schema, and hence emotional experiences, differ not only between individuals, but also between social and cultural groups [27,28].

By contrast with other proposals about the role of anoetic states in phenomenal experience, our view differs in two key respects. In our model, first-order neural activity is not all there is to anoetic consciousness. Thus, we maintain that the phenomenal feel of visual consciousness is not simply a state of visual cortex [2,3], and the phenomenal feel of fear is neither an amygdala state [4] nor a state representing body sensations [26]. Only once lower-order states are meta-represented does anoetic phenomenal consciousness result. Second, we suggest in any real-world experience, many other procedural brain systems will also be implicitly activated, besides the obvious ones mentioned above, and all will contribute to the phenomenal feel, the familiarity, of the mental moment, when monitored by the meta-representation.

The philosopher Owen Flanagan [29] noted that theories about the self overrate the role of autobiographical memories, and underestimate the role of the much larger class of events that are experienced but not remembered. We broadly agree but suggest that the reason these experiences may contribute is because they are implicitly remembered via procedural memory.

As noted, prefrontal circuits have been implicated in subjective metacognition, narratives, schemata, consciousness, including self-consciousness and emotional consciousness [6–9,15,16,17,23,24,27], but also in the ownership of mental states [25], and in the 'feeling of rightness', 'feeling of knowing', and 'feeling of familiarity' of such states [21,25]. Perhaps a general purpose higher-cognitive prefrontal network might underlie all kinds of conscious experiences [5–9,27].

Reconciling first-order and higher-order theories

The above provides a possible rapprochement between the first-order and higher-order theories of

consciousness. First-order theorists have claimed that phenomenal consciousness is theirs alone, and that higher-order theory is only about cognitive access and interpretation of the experiences. For them higher-order theory, with its emphasis on cognition, makes consciousness overly complex. On the other hand, higher-order theorists say that first-order theory is too simple, as it ignores conceptualized content, which is essential to real life conscious states.

Tulving's scheme, however, allows us to see the division of labor more clearly. Anoetic states are functionally equivalent to what first order-theorists call phenomenal consciousness. But in our schema anoetic states require sensory representations that are colored by procedural memory, and also requires that they be meta-represented. While noetic and auto-noetic consciousness involve more explicit forms of memories, they too inherit the phenomenal feel provided by anoetic consciousness in the background. In this sense, lower- and higher-order representations can be seen as components of a common theory of consciousness, rather than as elements of competing theories.

Conclusion

The nature of conscious experience in an organism hinges on the kinds of meta-representational states that its brain makes possible. Our brain has the capacity for anoetic, noetic, and auto-noetic meta-representations, and schematic narrations that translate the meta-representational content into corresponding conscious experiences that we feel are ours. Better understanding of the detailed neural circuitries involved would help clarify whether a general-purpose mechanism of consciousness exists. While it is difficult to study consciousness in animals, research on meta-representations face far fewer hurdles and offer a reasonable proxy for comparative studies of the pre-conscious processes related to consciousness.

Given that we are both neuroscientists, why are we so concerned about these 'philosophical' issues? Scientists sometimes

speculate on consciousness for sheer intellectual challenge. But our intentions are actually practical: we believe that progress in psychiatry and clinical psychology has been impeded by the fact that these fields have marginalized the role of subjective experience in understanding and treating mental disorders. This may be because subjective experience has often been seen as too elusive for scientific and clinical investigation. Extension of the science of consciousness into the clinic could both offer novel insights into the nature of consciousness, and suggest new avenues for helping sufferers *feel* better subjectively.

ACKNOWLEDGEMENTS

The authors thank Patrick Haggard and Barton Anderson for comments on earlier versions of the manuscript. J.E.L. and H.L. are supported by a grant from the Templeton World Charity Foundation (TWC0366). J.E.L. is also supported by the Vulnerable Brain Project and the National Institute of Health (R01DA044445).

REFERENCES

- Chalmers, D.J. (1997). Availability: the cognitive basis of experience. *Behav. Brain Sci.* 20, 148–149.
- Block, N. (2019). Empirical science meets higher order views of consciousness: reply to Hakwan Lau and Richard Brown. In *Blockheads! Essays on Ned Block's Philosophy of Mind and Consciousness*, A. Pautz and D. Stoljar, eds. (Cambridge, MA: MIT Press), pp. 199–214.
- Lamme, V.A. (2010). How neuroscience will change our view on consciousness. *Cogn. Neurosci.* 1, 204–220.
- Vandekerckhove, M., and Panksepp, J. (2011). A neurocognitive theory of higher mental emergence: from anoetic affective experiences to noetic knowledge and auto-noetic awareness. *Neurosci. Biobehav. Rev.* 35, 2017–2025.
- Rosenthal, D.M. (2005). *Consciousness and Mind* (Oxford: Oxford University Press).
- Lau, H., and Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends Cogn. Sci.* 15, 365–373.
- Brown, R., Lau, H., and LeDoux, J.E. (2019). Understanding the higher-order approach to consciousness. *Trends Cogn. Sci.* 23, 754–768.
- LeDoux, J. (2019). *The Deep History of Ourselves: The Four-Billion-Year Story of How We Got Conscious Brains* (New York: Viking).
- LeDoux, J.E., and Brown, R. (2017). A higher-order theory of emotional consciousness. *Proc. Natl. Acad. Sci. USA* 114, E2016–E2025.
- Squire, L. (1987). *Memory and Brain* (New York, Oxford: American Neurological Association).
- Tulving, E. (1985). Ebbinghaus's memory: what did he learn and remember? *J. Exp. Psychol. Learn. Mem. Cogn.* 11, 485–490.
- Metcalfe, J., and Son, L.K. (2012). Anoetic, noetic and auto-noetic metacognition. In *Foundations of Metacognition*, M. Beran, J.R.

- Brandl, J. Perner, and J. Proust, eds. (Oxford: Oxford University Press), pp. 289–301.
- Clark, A. (2000). *A Theory of Sentience* (New York: Oxford University Press).
- Schwartz, O., Hsu, A., and Dayan, P. (2007). Space and time in visual context. *Nat. Rev. Neurosci.* 8, 522–535.
- Lau, H. (2019). Consciousness, metacognition, & perceptual reality monitoring. *PsyArXiv* <https://doi.org/10.31234/osf.io/ckbyf>
- Fleming, S.M., and Dolan, R.J. (2012). The neural basis of metacognitive ability. *Philos. Trans. R. Soc. Lond. B* 367, 1338–1349.
- Odegaard, B., Knight, R.T., and Lau, H. (2017). Should a few null findings falsify prefrontal theories of conscious perception? *J. Neurosci.* 37, 9593–9602.
- Gazzaniga, M.S., and LeDoux, J.E. (1978). *The Integrated Mind* (New York: Plenum).
- Bruner, J. (1994). The "remembered" self. In *The Remembering Self: Construction and Accuracy in the Self-Narrative*, R. Fivush and U. Neisser, eds. (Cambridge: Cambridge University Press), pp. 41–54.
- Johnson-Laird, P.N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness* (Cambridge, MA: Harvard University Press).
- Gilboa, A., and Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. *Trends Cogn. Sci.* 21, 618–631.
- D'Argembeau, A., Cassol, H., Phillips, C., Baetue, E., Salmon, E., and Van der Linden, M. (2014). Brains creating stories of selves: the neural basis of autobiographical reasoning. *Soc. Cogn. Affect. Neurosci.* 9, 646–652.
- Fleming, S.M., Dolan, R.J., and Frith, C.D. (2012). Metacognition: computation, biology and function. *Philos. Trans. R. Soc. Lond. B* 367, 1280–1286.
- Dehaene, S., Lau, H., and Kouider, S. (2017). What is consciousness, and could machines have it? *Science* 358, 486–492.
- Klein, S.B. (2015). The feeling of personal ownership of one's mental states: a conceptual argument and empirical evidence for an essential, but underappreciated, mechanism of mind. *Psychol. Conscious. Theory Res. Pract.* 2, 355–376.
- Damasio, A., and Carvalho, G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nat. Rev. Neurosci.* 14, 143–152.
- LeDoux, J.E. (2020). Thoughtful feelings. *Curr. Biol.* 30, R619–R623.
- Barrett, L. (2017). *How Emotions Are Made* (Boston, New York: Houghton Mifflin, Harcourt).
- Flanagan, O. (2011). My non-narrative, non-forensic Dasein: the first and second self. In *Self and Consciousness*, J.L. Liu and J. Perry, eds. (Cambridge: Cambridge University Press), pp. 214–240.

¹Center for Neural Science and Department of Psychology, New York University, New York, NY 1003, USA. ²Department of Psychiatry, and Department of Child and Adolescent Psychiatry, New York University Langone Medical School, New York, NY 1003, USA. ³Emotional Brain Institute, Nathan Kline Institute, Orangeburg, NY 10962, USA. ⁴Department of Psychology and Brain Research Institute, UCLA, Los Angeles, 90095, USA. ⁵Department of Psychology and State Key Laboratory of Brain and Cognitive Sciences, University of Hong Kong, Hong Kong.

*E-mail: jel1@nyu.edu