Perception and memory in the medial temporal lobe: Deep learning offers a new lens on an old debate

Morgan D. Barense[1,2,*] and Andy C.H. Lee[1,2]
[1]Department of Psychology, University of Toronto, Toronto, ON M5S 3G3, Canada
[2]Rotman Research Institute, Toronto, ON M6A 2E1, Canada
*Correspondence: morgan.barense@utoronto.ca
https://doi.org/10.1016/j.neuron.2021.08.018

In this issue of *Neuron*, Bonnen et al. (2021) use artificial neural networks to resolve a long-standing controversy surrounding the neurocognitive dichotomy between memory and perception. They show that the perirhinal cortex supports performance on tasks that cannot be solved by the ventral visual stream.

One of cognitive neuroscience's most enduring debates is the question of where perception ends and memory begins. The crux of this debate has its origins in a seminal study by Scoville and Milner in 1953, in which a group of patients with intractable epilepsy, including the renowned patient H.M., underwent a radical and experimental treatment—bilateral medial temporal lobe (MTL) removal (Scoville and Milner, 1957). This surgery resulted in dense amnesia characterized by an inability to process long-term memories for facts and events, subsequently termed declarative memory (Squire and Wixted, 2011). Of particular significance, the impairment to declarative memory appeared to be entirely selective, with other cognitive functions, such as perception, working memory, nondeclarative memory, and language, adjudged to be intact. This remarkable pattern of preserved and disrupted cognition became a foundational cornerstone of cognitive neuroscience's prevailing paradigm, in which the brain is understood in terms of a modular organization of cognitive constructs that are mapped to specific brain regions or networks (Gazzaniga, 2013). According to this viewpoint, declarative memory is a distinct brain function mediated by the MTL, and it is separable from other cognitive abilities such as perception (Squire and Wixted, 2011). Visual perception, in contrast, is thought to have its own specialized pathway within the ventral visual stream (VVS) (Kravitz et al., 2013).

The notion of a compartmentalized MTL memory system became the guiding principle of memory research for decades, until the emergence of a collection of contradictory studies in rodents, nonhuman primates, and humans at the turn of the century (Buckley and Gaffan, 2006; Graham et al., 2010; Murray et al., 2007). This work critically implicated the perirhinal cortex (PRC), an MTL structure that is heavily interconnected with the VVS, in the perception of complex objects. Specifically, these studies found that PRC involvement in *both* memory and perceptual tasks was modulated by the visual properties of the stimuli. That is, PRC lesions impaired performance when the stimuli were complex and had many overlapping features but left performance intact when the task could be solved using a simple, single feature, such as an object's shape or color. These findings led to a schism in the field: whereas some argued for the necessity of novel theoretical frameworks to accommodate a role for the MTL in perception, traditionalists disputed these surprising findings and maintained that the MTL was critical for memory alone.

One prominent theoretical framework to emerge during this period is the representational-hierarchical model (Murray et al., 2007), which posits that features of objects are represented with increasing dimensionality throughout the VVS and extending into PRC, such that PRC represents objects at a greater degree of complexity than earlier regions in the VVS (e.g., inferotemporal [IT], V4) (Figure 1). The critical aspect of this theory is that different *operations* (e.g., those that support memory and those that support perception) can act on representations throughout all levels of the hierarchy. Thus, it is not the cognitive process (e.g., memory versus perception) that defines how regions in this pathway contribute to cognition, but rather the complexity of the representational content required for *any* given cognitive process,

be it perception or memory (Cowell et al., 2019).

Detractors of the representational-hierarchical model pointed toward inconsistent evidence and offered alternative explanations for data that appeared to challenge the traditional model of MTL function (Squire and Wixted, 2011). Several studies showed intact performance on seemingly complex perceptual tasks following PRC damage, and valid methodological concerns were leveled from both sides of the debate, with the most empirically intractable focused on issues of stimulus complexity and extraneous memory-related task demands. On one side of the debate, it was argued that PRC-related deficits on supposedly perceptual tasks were, in fact, due to unintended memory demands on those tasks. On the other side, it was suggested that the lack of PRC-related deficits on perceptual tasks was due to those tasks being insufficiently complex to require the PRC and relied on simpler representations housed in intact regions in the VVS (such as IT cortex, see Figure 1). In short, there was little agreement on what constituted a diagnostic task to assess a perceptual role for PRC, and because of a lack of formal, data-driven methods to determine perceptual and mnemonic task demands, the debate was stymied.

In a methodological tour de force reported in this issue of *Neuron*, Bonnen et al. (2021) provide a fresh re-examination of the MTL perceptual-mnemonic debate by utilizing neuropsychological, electrophysiological, and psychophysical data within a deep learning framework. Specifically, they leveraged convolutional neural network (CNN) models that can accurately predict the response of VVS regions to
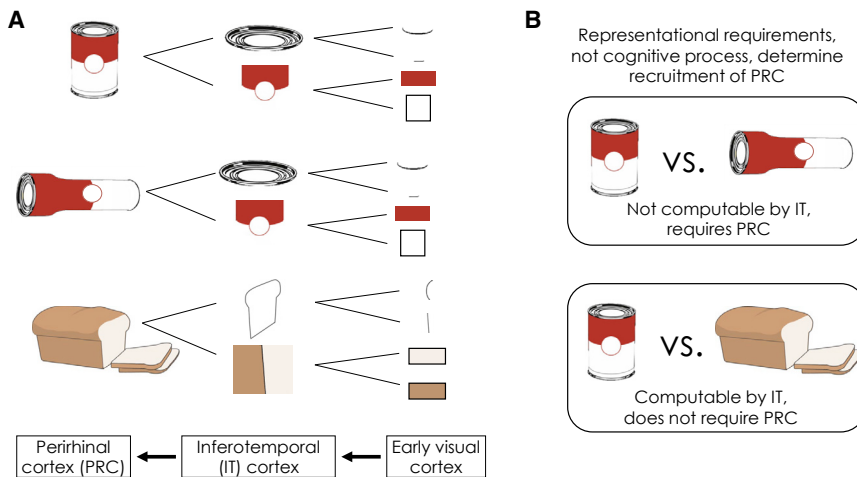
A

B

Representational requirements, not cognitive process, determine recruitment of PRC

VS.

Not computable by IT, requires PRC

VS.

Computable by IT, does not require PRC

Perirhinal cortex (PRC) ← Inferotemporal (IT) cortex ← Early visual cortex

Figure 1. Representational-hierarchical account of cognition
(A) Throughout the VVS and extending into the PRC, the features of an object are represented with increasing complexity. The PRC is at the apex, containing complex representations of objects.
(B) According to this theory, the same cognitive *process* (in this example, perceiving whether two objects are the same or different) is solved by different regions in the hierarchy. Thus, it is *representational content*, rather than cognitive process, that determines the division of labor throughout the pathway. Illustrations by Alexander Jacob.

visual stimuli, as a computational proxy for the visual system following a PRC lesion. The underlying logic is that when these models cannot solve a given task but neurologically intact participants can, it is assumed that the task does not rely on the VVS alone and must recruit additional structures, such as the PRC. To resolve the inconsistent pattern of evidence that has bedeviled the question of whether the PRC contributes to perception, Bonnen and colleagues first tackled the literature retrospectively. Sampling from both sides of the debate, they analyzed a retrospective dataset of 30 published "odd-one-out" perceptual discrimination experiments that had been administered to PRC-lesioned and PRC-intact participants. They generated predictions of VVS-supported performance on these perceptual tasks by passing task stimuli through the CNN and generating a predicted IT response for each individual stimulus. A covariance matrix was then created for the images on each trial, and the item with the lowest covariance was determined to be the model-selected odd-one-out (Figure 1C from Bonnen et al., 2021). Critically, the model employed a lossless decision-making protocol, meaning that the unweighted linear model readout offered a direct assessment of perceptual processing in high-

level visual cortex in the absence of extra-perceptual task confounds such as memory.

What is so important about the new methods developed by Bonnen et al. (2021) is that they are completely objective in assessing the perceptual demands of a task. In the past, experimenters had used their discretion in determining whether a task was sufficiently complex to recruit the PRC, often using a binary distinction of "high" versus "low" complexity. The current method formalizes this process and provides an unbiased continuous assessment of whether a given task can be solved by the VVS. With experimenter guesswork and intuition removed, the method is ideally suited to evaluate task complexity formally and resolve the long-standing debate surrounding PRC function. Perhaps unsurprisingly, this method reclassified 14 experiments from both sides of the debate (including from the present preview first author) as non-diagnostic—VVS model performance was at ceiling for these tasks, indicating that they were perfectly supported by the VVS alone and could not provide insight into the role of the PRC in perception. It was speculated that these tasks were non-diagnostic either because they were insufficiently complex to recruit PRC (in the instances where no PRC-related defi-

cits were previously found) or because the tasks imposed extra-perceptual demands such as memory, which were not captured by the lossless model readout (in the instances where PRC-related deficits were previously found). In either case, these experiments should not be considered relevant to the debate at hand.

Out of the original 30 retrospective experiments, 14 were found to employ perceptual tasks that were not computable by the VVS, suggesting that performance must depend on structures outside the VVS. The key question is whether this structure is the PRC. The results are very clear on this point: there was a striking correspondence between the model of VVS and PRC-lesioned behavior, such that as model estimates of VVS performance decreased, so, too, did performance in those with PRC lesions. That is, when the VVS was incapable of solving a task, those with PRC lesions also failed. Moreover, this relationship was continuous: the "harder" the task was for the VVS, the worse those with PRC lesions fared. In contrast, individuals with an intact PRC outperformed the computational proxy for the VVS. This confirms key predictions of the representational-hierarchical framework: in the absence of an intact PRC, an organism's perceptual abilities are limited to the abilities of the VVS, and an intact PRC provides an advantage over the VVS on some complex perceptual tasks.

In a final analysis employing a novel dataset, Bonnen et al. (2021) zeroed in on which region in the VVS is most closely aligned with PRC-lesioned behavior. Here, they developed a new set of oddity trials using stimuli that were previously shown through electrophysiological recordings in the macaque to separate V4- from IT-supported behavior. Weighted, linear readouts of V4 and IT recordings (learned responses obtained via logistic regression and leave-one-out cross validation) were used to estimate how these two regions would perform on each trial of the oddity task. These readouts were compared directly to the performance of the computational model of VVS and to data from a large sample of PRC-intact human participants who completed the same oddity tasks. PRC-intact humans outperformed both the linear readouts of macaque IT and the computational proxy

for IT, whereas linear readouts from macaque IT outperformed those from macaque V4. Moreover, the computational proxy for IT predicted electrophysiological recordings from IT, providing validation for the use of this computational proxy in the aforementioned retrospective analyses. Finally, an interesting relationship was revealed between electrophysiologically derived IT performance and PRC-intact human reaction times: trials that took more time were associated with lower IT-supported accuracy. Why would PRC-intact participants require more time to solve trials that are not supported by IT? One proposal is that a key role for the PRC is to integrate feature information gathered from multiple saccades across an object into a richer and more cohesive representation than that afforded by IT (Erez et al., 2013).

In sum, by applying a novel unified computational framework of VVS function to a range of experimental data, Bonnen et al. (2021) have provided a formal framework for interpreting perceptual demands and, in doing so, resolved a wealth of conflicting evidence pertaining to the involvement of PRC in perception. An important avenue for future research is the application of the approach developed by Bonnen et al. (2021) to examine the role of other MTL structures, such as the hippocampus, in perception. The representational-hierarchical model situates the hippocampus above the PRC in the hierarchy, representing conjunctions of multiple objects and binding objects to a spatial/temporal context (Cowell et al., 2019). Consistent with this, patients with selective hippocampal damage are impaired on perceptual scene oddity tasks (Erez et al., 2013; Graham et al., 2010), and indeed, Bonnen et al. (2021) demonstrated that these scene tasks could not be solved by the VVS alone (Figure S4 from Bonnen et al., 2021). An intriguing question, therefore, is whether the contribution of the hippocampus to perception can be evaluated computationally, as has been done for the PRC.

In closing, we return to the original question that has motivated a decades-long debate in cognitive neuroscience: do MTL structures contribute to perception in addition to memory? The answer is "it depends," and thus, we argue that the question needs to be reframed. Rather than adhere to cognitive neuroscience's topological tradition of classifying brain regions or networks according to broad cognitive constructs like "memory" and "perception," we propose that a more fruitful approach to understanding brain-behavior relationships in the VVS and MTL will be to precisely quantify the representational content required for any given cognitive operation. Here, Bonnen and colleagues provide the field with novel tools to do just that.

REFERENCES

Bonnen, T., Yamins, D.L.K., and Wagner, A.D. (2021). When the ventral visual stream is not enough: A deep learning account of medial temporal lobe involvement in perception. Neuron 109, 2755–2766. https://doi.org/10.1016/j.neuron.2021.06.018.

Buckley, M.J., and Gaffan, D. (2006). Perirhinal cortical contributions to object perception. Trends Cogn. Sci. 10, 100–107. https://doi.org/10.1016/j.tics.2006.01.008.

Cowell, R.A., Barense, M.D., and Sadil, P.S. (2019). A Roadmap for Understanding Memory: Decomposing Cognitive Processes into Operations and Representations. eNeuro 6. ENEURO.0122-19.2019. https://doi.org/10.1523/ENEURO.0122-19.2019.

Erez, J., Lee, A.C.H., and Barense, M.D. (2013). It does not look odd to me: perceptual impairments and eye movements in amnesic patients with medial temporal lobe damage. Neuropsychologia 51, 168–180. https://doi.org/10.1016/j.neuropsychologia.2012.11.003.

Gazzaniga, M.S. (2013). Shifting gears: seeking new approaches for mind/brain mechanisms. Annu. Rev. Psychol. 64, 1–20. https://doi.org/10.1146/annurev-psych-113011-143817.

Graham, K.S., Barense, M.D., and Lee, A.C.H. (2010). Going beyond LTM in the MTL: a synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. Neuropsychologia 48, 831–853. https://doi.org/10.1016/j.neuropsychologia.2010.01.001.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G., and Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. Trends Cogn. Sci. 17, 26–49. https://doi.org/10.1016/j.tics.2012.10.011.

Murray, E.A., Bussey, T.J., and Saksida, L.M. (2007). Visual perception and memory: a new view of medial temporal lobe function in primates and rodents. Annu. Rev. Neurosci. 30, 99–122. https://doi.org/10.1146/annurev.neuro.29.051605.113046.

Scoville, W.B., and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. J. Neurol. Neurosurg. Psychiatry 20, 11–21.

Squire, L.R., and Wixted, J.T. (2011). The cognitive neuroscience of human memory since H.M. Annu. Rev. Neurosci. 34, 259–288. https://doi.org/10.1146/annurev-neuro-061010-113720.