Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Current Opinion in Behavioral Sciences

Editorial overview: Understanding memory: which level of analysis?

Morgan Barense and Hugo J Spiers

Current Opinion in Behavioral Sciences 2020, **32**:iii–vi

For a complete overview see the Issue

https://doi.org/10.1016/j.cobeha.2020.05.003

Morgan Barense[1,2]

[1] Department of Psychology, University of Toronto, Toronto, Canada
[2] Rotman Research Institute, Baycrest Hospital, Toronto, Canada

Morgan Barense is a Professor and Canada Research Chair in Cognitive Neuroscience at the University of Toronto, an Associate Scientist at the Rotman Research Institute at Baycrest Hospital, and the Director of the Toronto Neuroimaging Facility at the University of Toronto. She received her Ph.D. in 2007 from the University of Cambridge, where she trained in neuropsychology and neuroimaging, and her B. A. in 2002 from Harvard University, where she trained in animal neuroscience. She established her research group at the University of Toronto in 2009 where she and her team bring these approaches together to study the neural underpinnings of memory. Her research has shown that damage memory and perception are inextricably intertwined, relying on shared neural representations and computational mechanisms, and that medial temporal lobe amnesia is not a selective deficit in long-term memory. She and her team have recently developed a smartphone-based memory prosthetic that uses principles from neuroscience and psychology to improve memory for everyday events. She received a Young Investigator Award from the Cognitive Neuroscience Society, an Early Investigator Award and Lifetime Fellowship from the Society of Experimental Psychologists, an Early Career Award from the Canadian Society for Brain, Behaviour and Cognitive Science, an Early Researcher Award from the Province of Ontario, and a Connaught Innovation Award.

Hugo J Spiers

Institute of Behavioural Neuroscience, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, UK

How the brain stores and retrieves memories has been a central question in psychology and neuroscience since the formation of these disciplines. Many questions have dominated the scientific study of memory. These stretch from asking how culture shapes our memories to exploring how molecular interactions create memories. With this diversity comes debate. What are the most productive approaches to studying and conceptualizing memory? For example, should one examine specific brain regions to understand their function or focus on brain-wide networks? Should research concentrate on exploring cognitive processes or on the neural representations? Are highly controlled stimuli essential or should researchers adopt more ecological stimuli to understand memory?

In this special issue, we have brought together experts to reflect on the advantages and pitfalls of studying memory at different levels of analysis. In this overview we explore the common themes that emerge across these reviews, and consider the challenge of integrating across perspectives to provide a more comprehensive understanding of memory. We were grateful to receive reviews covering a wide range of methods which include: behavioural studies with humans [e.g. Shevlin;Ryan and Shen;Simons, Mitregna, Fernyhough], developmental research in infants [Sherman, Graves, and Turk-Browne], human neuroimaging [e.g. Love;Bird;Lee, Bellana, Chen;Olsen and Robin;Koen, Srokova and Rugg], computational modelling [Rogers;Momennejad;Love;Cowell and Huber], patient studies [McAndrews;Irish and Vatansever], rodent lesion and genetic manipulations [e.g. Harel and Ryan;Barker and Warbutron;Milczarek and Vann], human/animal sleep oscillations [Naverrete *et al.*;Swanson *et al.*;Genzel], brain stimulation [Hebscher and Voss], and neurochemical modulation [Decker and Duncan].

Some authors in this special issue directly tackled our question: At which level of analysis should we study memory? Cowell and Huber draw on theories from the philosophy of science combined with convergent neuropsychological, neuroimaging, modelling approaches to argue that an 'intermediate' level, defined in terms of neural representations, cognitive operations, will provide maximal explanatory power in characterizing the mechanisms of memory. This intermediate level (e.g. pattern completion in the hippocampus) lies immediately below the level of phenomena to be explained (e.g. recall ability and its relationship to brain damage). In contrast, higher-levels of description (e.g., recollection and familiarity) are the very phenomena we wish to explain and thus cannot provide a mechanistic explanation for how they themselves are instantiated in the brain. Harel and Ryan focus on methodological precision and a more fine-

Hugo Spiers is Professor of Cognitive Neuroscience at University College London (UCL). He is group leader for the UCL Spatial Cognition Group which is based in the Institute of Behavioural Neuroscience and the Department of Experimental Psychology at UCL. He received his PhD in Neuroscience at UCL in 2002 and trained in neuropsychology, neuroimaging before obtaining a Wellcome Trust Advance Training Fellowship to learn single unit recording in freely moving rodents. He established his research group in 2010 at UCL which integrates research using immersive wireless virtual reality, GPS tracking in wild, world-wide mass online testing, single unit recording in rats and functional neuroimaging in humans. His research has provided insights into how the brain represents the distance and direction to goals, simulates the future at the cellular and brain-network level. He has been elected as a member of the Memory Disorders Research Society and is an Associate Fellow of the Royal Institute of Navigation.

grained approach in their review, advocating that it is the link between precisely mapped networks of neurons and behavior that is the most productive way to understand memory. This view is driven by the impressive advances in optogenetics, chemogenetics and calcium imaging that afford the opportunity to tag, image and manipulate exactly those cells involved in forming and retrieving a memory — ‘engram cells’. This perspective sits in stark contrast to that proposed by Simons, Mitregna, Fernyhough, who argue for the importance of considering broader perspectives from the arts and the humanities if we are to meaningfully understand the subjective experience of memory. A similar perspective that highlights the need for understanding culture is given by Shevlin;Simons, Mitregna, Fernyhough put emphasis on integrating rigorous behavioural, neural studies with the attempt to understand the rich phenomenological aspects of memory.

Other authors highlighted levels of analysis that have traditionally been underappreciated. For example, Decker and Duncan offer a novel and refreshing perspective on the importance of internal brain states in shaping memory. Considering neural mechanisms at the neurochemical rather than the brain structural level, they focus on the different timescales over which neurotransmitters fluctuate and highlight the interdependence of cognitive constructs like memory and attention. Ryan and Shen feature the importance of monitoring eye movements to understand information processing on a moment-to-moment basis. They describe the shared evolutionary history of both the oculomotor and hippocampal systems and demonstrate that eye movements do not reflect the passive expression of the contents of memory, but rather have a functional role in memory formation, retrieval, and construction.

Many authors also highlighted how a convergence of methods has enabled advances in understanding memory. For example, Olsen and Robin explain how new advances in neuroimaging combined with careful neuropsychological studies in patients with focal lesions have enabled a more precise understanding of memory’s neuroanatomical basis. Drawing from varied approaches, they focus on how a specialization gradient along the long-axis of the hippocampus relates to differences in subfield composition, as well as to structural and functional connectivity with a broader network of brain regions outside the hippocampus. Koen, Srokova and Rugg also consider how changes in brain-wide networks are important in understanding memory, in particular how the breakdown in the specificity of neural processing is affected by aging and other individual differences. Novel analyses considering neural dedifferentiation at the network-level will be important as the field shows an increased research focus in how aging impacts brain function.

It was also clear that the field is moving away from a focus on a 1-to-1 mapping between a given brain region and a given cognitive construct, and instead studying memory at the network level. For example, rather than considering the effects of transcranial magnetic stimulation (TMS) in terms of a focal virtual lesion, Hebscher and Voss highlight the importance of considering the larger network within which a targeted brain region lies and the various intermediate-level operations that contribute to episodic memory. Likewise, Irish and Vatansever describe recent discoveries regarding gradients in the brain, showing that macro-scale organizational principles from network-based methods can inform profiles of memory impairment in clinical populations. They place an emphasis on treating functional distinctions between brain regions in terms of gradual transitions, rather than sharp borders based on black-and-white cognitive constructs. Sherman, Graves, and Turk-Browne show that by moving our focus away from 1-to-1 mapping of brain regions to cognitive constructs, we will better appreciate that some kinds of memory - like statistical learning - can reconcile seemingly distinct learning phenomena. They show that statistical learning has a pervasive influence on human cognition and behaviour, making contributions to a wide range of behaviours that are often studied as separate processes, such as episodic memory or spatial navigation.

This approach to considering memory *networks* rather than individual brain areas was evident in the reviews considering the importance of sleep and brain-state for memory encoding and consolidation (Swanson *et al.*; Naverrete;Genzel). Swanson *et al.* provide a new perspective on the way in which information content carried by hippocampal spikes during sharp-wave-ripples (SWR) may be multiplexed for memory consolidation across the brain networks, with different downstream regions able to extract and consolidate different types of information from the same SWR. This perspective is complemented by that of Naverrete *et al.* who provided a model for how three different types of oscillation may be required in combination to support memory consolidation. In developing the model Naverrete *et al.* provide an excellent overview of the varied new discoveries coming from the study of brain oscillations during sleep. Genzel, also further reviews research on SWRs and describes how brain-state (e.g. levels of neuromodulators or recent circuit interactions) will influence how networks, and hubs in networks, establish neural patterns for memory encoding and retrieval. A key point made is that we should avoid seeing memories as static entities, but rather consider them as more dynamic and changeable as to which brain regions are involved and associated with other memories.

A number of reviews highlight the importance of studying how the brain supports natural spontaneous behaviour or processes memories for naturalistic stimuli (Barker and Warbutron;Lee, Bellana, and Chen;Bird). In rodents, Barker and Warbutron argue for the importance of taking a behavioural-level, systems-level, and cellular-level approach to studying natural non-trained behaviours, such as spontaneous object preference tests which exploit an animal's innate preference to explore novel elements of an environment. In humans, the move to more naturalistic stimuli and behaviour has been accompanied by advances in human neuroimaging analyses that exploit the structure in the brain responses to naturalistic stimuli. Bird provides a comprehensive and engaging overview of how the human brain supports memory for complex real-world events, and he notes that much of this progress was made possible by the combination of carefully designed tasks combined with cutting-edge, innovative fMRI analyses. Similarly Lee, Bellana, and Chen review the exciting advances gained by using intersubject correlations between the response time-course or spatial patterns of neural activity across two or more subjects. By focusing on narrative memories, which are frequently shared through interpersonal communication, they highlight how collective experience is reflected in converging neural activity across individuals.

Both Love and Momennejad highlight how analyses with specific models in mind are critical, and review areas in the realms of category learning and reinforcement learning where significant advances were made possible by the application of a model to predict the neural responses. Momennejad concisely covers reinforcement learning, where, rather than simply exploring encoding and retrieval (e.g. watching and recalling a movie), the aim is to understand the learning of optimal policies for efficiently guiding actions to maximize rewards (e.g. how to get to the best cinema). Momennejad focuses on the 'successor representation', which provides a half-way-house between a habit-like caching system for learning optimal actions (model-free) and mental-search system that works through all the possible future actions to make the next move (model-based). By using multi-scale representations and prioritized off-line replay Momennejad shows how a reinforcement learning agent can better generalise over memories and efficiently plan for the future. Love argues that it is essential to use models when analysing fMRI data. He reviews a range of studies to show that constrained cognitive models applied to multiple fMRI datasets are much more successful in uncovering the contributions of brain regions make to a task than standard approaches. Rogers convincingly argues that contemporary neural network models capture a level of description isomorphic to that adopted by by functional brain imaging and connectivity analyses. Building his case with four separate examples in which neural network models have reshaped our understanding of semantic cognition, he demonstrates that neural network models provide us with an essential level of description because they adopt the same simplifying assumptions that have been critical to advances in cognitive neuroscience.

While there was obvious excitement surrounding the many methodological advances that our field has recently enjoyed, in almost every review it was clear that the methods that launched the field of cognitive neuroscience — well-designed behavioural paradigms combined with lesion studies in animals or humans — continue to provide a cornerstone for understanding memory. For example, Milczarek and Vann highlight the contributions of many cutting-edge innovative technologies in behavioural neuroscience that have contributed to understanding memory at both a structural and systems-level, but they remind us that the contribution of these technological advances will be limited if they occur at a faster rate than our understanding of the underlying behaviour that they measure. And finally, McAndrews, Cohn, and Gold remind us of the real-world implications of our work, focusing on how cognitive neuroscience advances have reached the clinic to enhance diagnosis, prognosis, and validation of therapeutic strategies.

Overall, these reviews showcase the exciting diversity of methodological approaches and theoretical perspectives that are currently being used to study memory. Collectively they raise a number of points for consideration. There was a stark contrast between some perspectives: consider the gap between insights gained from tagging and manipulating 'engram cells' [Harel and Ryan] and insights gained from the writings of literary giants such as Woolf, Wordsworth, and Byron [Simons, Mitregna, Fernyhough]. Such diversity makes our field exciting, but it also poses a challenge. How are we to communicate across these vast divides? In our own work we have experienced the challenges of conducting research that bridges across perspectives and species. In some areas, such as the extremely intricate and specialized methods that are becoming the norm in rodent hippocampal memory research, the amount of training and methodological expertise required is considerable and may leave little time to explore other corners of our field. But we argue that memory scientists need to abandon our respective siloes and embrace the research and varied perspectives that emerge from a breadth of approaches. The reviews here highlight that this will be facilitated by shared behavioural tasks that allow comparison across species and the use of explicit models that force the application of a constrained theory to predict behaviour or neural response. This is something we have actively sought to achieve in our past research on spatial navigation [1,2] and visual discrimination [3,4]. While better integration will help bring a more holistic understanding of memory, it is also important to recognise that not all tasks and neural dynamics are universal. There are species differences and cultural differences as well as intrinsic value in exploring topics that do not straddle the breadth of approaches. But even — or especially — in light of these differences, it is essential that memory scientists communicate across the divide. Such was our goal with this special issue, and we are grateful to the authors who graciously accepted our invitation to take part in this discussion.

References

1. Spiers HJ, Barry C: **Neural systems supporting navigation**. *Curr Opin Behav Sci* 2015, **1**:47-55.
2. Spiers HJ, Gilbert SJ: **Solving the detour problem in navigation: a model of prefrontal and hippocampal interactions**. *Front Hum Neurosci* 2015, **9**:125.
3. Barense MD, Groen IIA, Lee ACH, Yeung LK, Brady SM, Gregori M, Kapur N, Bussey TJ, Saksida LM, Henson RNA: **Intact memory for irrelevant information impairs perception in amnesia**. *Neuron* 2012, **75**:157-167.
4. Barense MD, Bussey TJ, Lee ACH, Rogers TT, Davies RR, Saksida LM, Murray EA, Graham KS: **Functional specialization in the human medial temporal lobe**. *J Neurosci* 2005, **25**:10239-10246.