Neurobiology of Learning and Memory 129 (2016) 99–106

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme

The representational–hierarchical view of pattern separation: Not just hippocampus, not just space, not just memory?

B.A. Kent [a,*], M. Hvoslef-Eide [b], L.M. Saksida [b,c,d,e,f], T.J. Bussey [b,c,d,e,f]

[a] Department of Medicine, Division of Neurology, University of British Columbia, Vancouver, Canada
[b] Department of Psychology, University of Cambridge, Cambridge, UK
[c] The Brain and Mind Institute, Western University, London, ON, Canada
[d] Molecular Medicine Research Group, Robarts Research Institute, Schulich School of Medicine & Dentistry, Western University, London, ON, Canada
[e] Department of Physiology and Pharmacology, Schulich School of Medicine & Dentistry, Western University, London, ON, Canada
[f] MRC and Wellcome Trust Behavioural and Clinical Neuroscience Institute, University of Cambridge, Downing St, Cambridge CB2 3EB, UK

ARTICLE INFO

Article history:
Received 6 October 2015
Revised 6 January 2016
Accepted 13 January 2016
Available online 1 February 2016

Keywords:
Pattern separation
Memory
Hippocampus
Perirhinal cortex
Representational–hierarchical
Perception

ABSTRACT

Pattern separation (PS) has been defined as a process of reducing overlap between similar input patterns to minimize interference amongst stored representations. The present article describes this putative PS process from the "representational–hierarchical" perspective (R–H), which uses a hierarchical continuum instead of a cognitive modular processing framework to describe the organization of the ventral visual perirhinal–hippocampal processing stream. Instead of trying to map psychological constructs onto anatomical modules in the brain, the R-H model suggests that the function of brain regions depends upon what representations they contain. We begin by discussing a main principle of the R–H framework, the resolution of "ambiguity" of lower level representations via the formation of unique conjunctive representations in higher level areas, and how this process is remarkably similar to definitions of PS. Work from several species and experimental approaches suggest that this principle of resolution of ambiguity via conjunctive representations has considerable explanatory power, leads to wide possibilities for experimentation, and also supports some perhaps surprising conclusions.

1. Introduction

The importance of complex conjunctive representations for the resolution of ambiguity in lower-level representations has been a major focus of our research and that of others. Much work from several species and experimental approaches suggests that this principle of resolution of ambiguity via conjunctive representations has considerable explanatory power, particularly regarding how best to understand the effects of focal brain dysfunction (e.g., Barense et al., 2005, 2012; Bartko, Winters, Cowell, Saksida, & Bussey, 2007; Bussey & Saksida, 2002; Cowell, Bussey, & Saksida, 2006; Cowell, Bussey, & Saksida, 2010a; Graham et al., 2006; Lee, Buckley, et al., 2005; Lee, Bussey, et al., 2005; Lee & Rudebeck, 2010; McTighe, Cowell, Winters, Bussey, & Saksida, 2010). We have referred to this way of thinking about brain organization as the "representational–hierarchical" perspective (R–H) (e.g., Bussey & Saksida, 2002, 2005; Murray, Bussey, & Saksida, 2007; Saksida & Bussey, 2010).

Recently there has been a great deal of interest in a process referred to as "Pattern Separation" (PS). PS has been defined as "… reducing interference among similar inputs by using non-overlapping representations…" (e.g., Reagh et al., 2014) and "…the ability of the network to reduce the overlap between similar input patterns before they are stored in order to reduce the probability of interference…" (Neunuebel & Knierim, 2014). Clearly the main principle of the R–H framework, the resolution of the "ambiguity" of lower level representations via the formation of unique conjunctive representations in higher level areas, is a strikingly similar idea to the above conceptions of PS. The present article is, therefore, aimed at researchers interested in PS, and explores the question: To the extent that researchers are interested in PS because it results in the formation of new, interference-reducing representations, what insights into PS might be offered by considering some of the conclusions resulting from R–H theory?

A quick word about the scope and aims of this review. R–H theory grew out of an interest in the *functional relevance* of processes and representational content in the brain. That is, what such processes and representational content do for us, in aspects of cognition such as memory and perception. The way to determine this is through empirical, causal behavioral experiments that measure

* Corresponding author.
E-mail address: brianne.kent@ubc.ca (B.A. Kent).

http://dx.doi.org/10.1016/j.nlm.2016.01.006
1074-7427/

cognition. Non-empirical approaches such as computational modeling have been hugely important in the area of PS (O'Reilly and McClelland, 1994; Rolls, 1987, 1989, 1990; Norman & O'Reilly, 2003; Rolls & Kesner, 2006; Treves & Rolls, 1994) and indeed, experiments testing the R–H view have been guided by computational modeling (e.g., Barense et al., 2012; Bartko et al., 2007; Cowell et al., 2006). But we must remember that these are models, and at some point experiments on real brains must be done. Correlational experimental approaches, particularly electrophysiology, have shed much light on PS at the cellular and network levels (e.g., Neunuebel & Knierim, 2014; Knierim & Neunuebel, 2016), and focus on PS as a specific mechanism involving the transformation of an input representation to an output representation, in which the output is less correlated than the input (in line with computational models). However this process is of interest because of the representations thus formed and in particular their requirement for cognition (usually memory). Correlational approaches generally do not address this requirement. Additionally, correlational approaches have largely focussed on the hippocampus, whereas a main aim of what follows below is a suggestion that we need to consider other areas of the brain with respect to PS. Fully understanding any function of the brain cannot be accomplished by any single method alone, but requires converging, complementary approaches.

Furthermore, our focus in this article is the relevance to cognition of the *result* of processes such as PS, i.e., the representations that are formed from such processes. We will not, in this review, discuss *how* these representations are formed. For example, we will consider that such representations may be housed in the ventral visual stream, but not the way the formation of such representations might be related to the receptive fields of neurons. Comprehensive treatments of such issues, along with the properties of such representations (e.g., invariance), can be found in other sources such as Rolls (2016).

Insofar as PS can be considered to be a process of forming new representations that help resolve ambiguity, the results of experiments carried out under the auspices of R–H theory lead us to offer a number of possibly surprising hypotheses about PS, including:

1. PS is fundamental to many aspects of cognition including perception; it is not just for memory.
2. PS happens in many cortical regions, not just the hippocampus.
3. PS happens for all stimulus material and not just 'spatial' or 'episodic' material. It happens for different types of representations, in the different regions and different levels throughout the "representational hierarchy".
4. The dentate gyrus (DG) is unlikely to maintain all levels of representation, and thus is not a truly domain-general pattern separator.
5. PS – insofar as this term equates with the formation of interference-reducing conjunctive representations – may have much wider explanatory power for understanding the effects of brain dysfunction than previously suspected.

Below we provide a brief history and summary of R–H theory, with special emphasis on these particular conclusions regarding PS, and then return to these points.

2. The Representational–Hierarchical model: A brief history and summary

The R–H view was first presented in 1998 (Saksida & Bussey, 1998) and discussed in a review article in 1999 (Murray & Bussey, 1999). Eventually the first modeling and experimental studies testing the theory were published (Bussey & Saksida, 2002; Bussey, Saksida, & Murray, 2002, 2003). Although R–H takes a broad view of cognition, initial studies focused on visual cognition, and in particular the issue of whether there are dedicated, anatomically segregated memory (Squire & Zola-Morgan, 1991) and perceptual representation (Schacter & Tulving, 1994) systems in the brain. Initial experimental studies focused on perirhinal cortex (PRh) as a "testing ground" at the anatomical interface between putative memory and perceptual systems. Other authors published similar ideas (e.g., Buckley & Gaffan, 1998; Gaffan, 2002) around this time and since (e.g., Graham, Barense, & Lee, 2010; Nadel & Peterson, 2013).

Our approach was to try to understand impairments in visual cognition following focal brain dysfunction. The prevailing view was that such impairments could be understood in terms of damage to a processing module specialized for a particular aspect of cognition (Schacter & Tulving, 1994; Squire & Zola-Morgan, 1991). In contrast to such a modular view, R–H theory proposed that representations are organized in a hierarchical continuum and are useful for all aspects of cognition that require them, including memory and perceptual discrimination [indeed we initially referred to this idea as the "Perceptual-Mnemonic/Feature-Conjunction" model (Bussey & Saksida, 2002)]. Thus, R–H emphasizes *content* rather than processing. The strongest version of the view is that there are no substantive processing differences within these various regions, and effects of brain dysfunction can be understood entirely on the basis of content (see Forwood, Cowell, Bussey, & Saksida, 2012). Although this very strong version of the view is unlikely to be entirely correct, it is perhaps surprising just how much it has been able to explain, without having to invoke putative differences in processing.

The basic idea of the R–H view is illustrated in Fig. 1.

To summarize a few key principles of R–H theory:

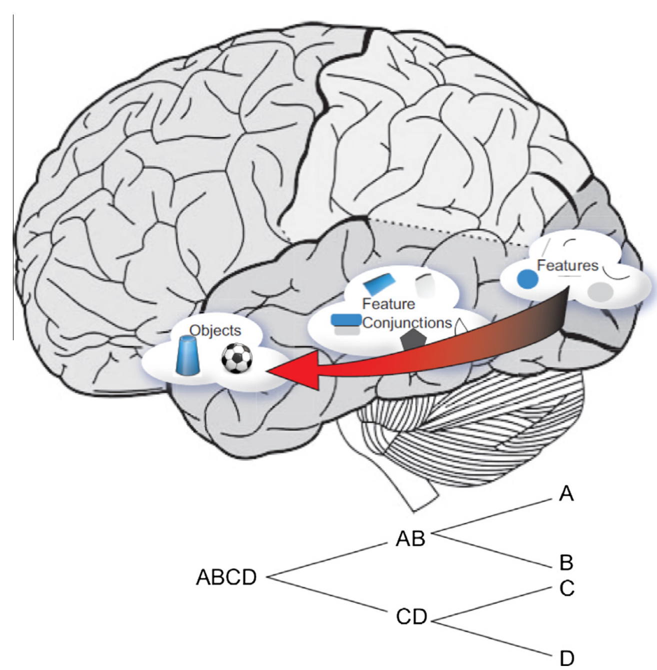

Fig. 1. R–H applied to visual cognition (adapted from McTighe et al., 2010). As representations pass through regions of the visual ventral stream (VVS) and into the medial temporal lobe, they become increasingly more complex in a hierarchical fashion. Representations supported by caudal regions represent "features" (e.g. A, B, C and D), whilst representations supported by more rostral regions represent conjunctions of those features (eventually leading to a representation at the level of a whole object-level representation (ABCD) in PRh and spatial and contextual representations in the hippocampus). The traditional multiple memory systems view suggests that structures within the medial temporal lobe subserve exclusively (declarative) mnemonic function, whereas structures in the ventral visual stream are important for functions such as perceptual discrimination. In contrast, the representational–hierarchical view suggests that stimulus representations throughout the ventral visual–perirhinal–hippocampal stream are useful for any cognitive function that requires them.