Title
Entities also require relational coding and binding

Authors
Brady, Timothy F
Utochkin, Igor S

Publication Date
2019

DOI
10.1017/s0140525x19001924

Peer reviewed

2005; Frenda et al. 2011). Conversely, the Deese–Roediger–McDermott illusion increases linearly from childhood to older adulthood (e.g., Brainerd et al. 2008; Gallo 2010). How might these different data patterns relate to the core versus attributional systems in the integrative memory model? We welcome Bastin et al.'s insights on how developmental patterns in false memory illusions might constrain or validate their model. Indeed, we feel it would be informative to consider the development of recollection and familiarity processes for both true and false memories.

In sum, Bastin et al. should justify the integrative memory model's substantial complexity by addressing how that complexity contributes to our understanding of (1) different types of false memory phenomena (particularly false recollection), and (2) the development of recollection and familiarity for true and false memories across the lifespan. By incorporating these missing elements, we feel the integrative memory model would be more integrative and thus better live up to its name.

Entities also require relational coding and binding

Timothy F. Brady[a] and Igor S. Utochkin[b]

[a]Department of Psychology, University of California, San Diego, CA and
[b]National Research University Higher School of Economics, Moscow, Russia.
timbrady@ucsd.edu isutochkin@inbox.ru http://bradylab.ucsd.edu/
https://www.hse.ru/en/staff/utochkin

doi:10.1017/S0140525X19001924, e285

Abstract

Although Bastin et al. propose a useful model for thinking about the structure of memory and memory deficits, their distinction between entities and relational encoding is incompatible with data showing that even individual objects – prototypical "entities" – are made up of distinct features which require binding. Thus, "entity" and "relational" brain regions may need to solve fundamentally the same problems.

A fundamental tenet of the integrative memory model proposed by Bastin et al. is the distinction between entity representation and relational representations. This distinction is based on the premise that there is a meaningful sense of an "entity" that is holistic and unitized and can be stored and retrieved without recollection or binding; for example, Bastin et al.'s claim that at the "level of the perirhinal cortex and anterolateral entorhinal all visual features are integrated in a single complex representation of the object that can be discriminated from other objects with overlapping features."

It is natural to think that meaningful visual objects might be unitized entities for visual memory, and indeed many memory researchers take this claim for granted. However, we believe this view of unitized, fully bound representation of objects is inconsistent with the cognitive data on object memory. In particular, supposed "entities" such as visual objects are not unitized representations, but themselves are stored as separate features requiring binding in the same fundamental way that objects need to be bound to contexts. This calls into question the core distinction between entities and relations proposed by the integrative memory model and other similar models: If even single objects are stored in a way consistent with the "relational system" – where the "representation keeps components separate and flexibly bound" (target article, sect. 4.1, para. 7) – then it is not clear what an entity would be or whether the fundamental nature of the brain regions subserving object memory are really distinct from those subserving item-context integration.

Work from our labs shows that visual object features are stored and accessed independently in long-term memory: Different features of single individual objects are forgotten at different rates (Brady et al. 2013); and people remember particular features but don't remember which objects these features belonged to (Utochkin & Brady 2019). For example, if people are shown a blue open backpack and then asked to choose among sets of four backpacks that are blue/open, blue/closed, or red/open, red/closed, people can forget the color but still can remember its "open-ness" (Brady et al. 2013). Or, if people are shown one mug (mug A) full of coffee and another mug (mug B) empty, they are well above chance reporting that they saw one full mug and one empty mug, and that they saw mug A and mug B, but they are at chance at ascribing the "fullness" and "emptiness" to the particular mugs A and B (Utochkin & Brady 2019).

Indeed, the idea that some items might be stored in a fully unitized representation (in an "entity" system) seems incompatible with the cognitive problem that object representations are designed to solve. One of the central requirements of a visual memory system is robustness to variation (Schurgin & Flombaum 2018). If object representations were totally integrated entities in memory, their recognition in the real world would be extremely problematic, given the infinite number of poses and states these objects can take, as well as variations in orientation, lighting, and more. One might argue that this invariant recognition is possible if a unitized "concept" of a particular object is formed during multiple episodes – connecting across multiple experiences when an object is presented in different states and viewpoints. But this claim immediately implies an independence of the features forming the core of this concept and those representing the way it changes across contexts.

In contrast to any view based on unitized object memories, we have demonstrated that people's memories are extremely robust to variation even at the level of individual objects, suggesting that representations even at the level of objects are based on separate features that are flexibly bound together rather than unitized. For example, imagine you saw an open-doored cabinet, and then later we asked which cabinet you had seen – but now the "old" item was shown in a new state (the same cabinet now has its doors closed, changing a huge number of visual features). We have shown that people are nearly perfect at generalizing in this way, and can do so even if the "foil" presented at test is a new open-doored cabinet, designed to maximally mislead participants (Utochkin & Brady 2019). Therefore, we believe the flexible nature of binding attributed by the integrative memory model only to item-context distinctions and recollection situations, need to be extended to nearly every level of representation of objects as well as contexts.

Similar evidence for independence and structured representation rather than unitized objects is present in the visual working memory literature, where it is frequently found that both objects and separate features can be stored and objects are not stored as single integrated units (see Brady et al. 2011 for review). Since working memory is critical for consolidation into long-term memory, this may be the beginning of the non-unitized, non-integral storage of items in memory.

Overall, we believe that entities are stored in a way that is not holistic or unitized – and thus, at nearly every level of representation, there is a need for flexible, relational encoding. If this is a common property of memories for individual objects (which is associated with "entities") and complex episodes (requiring the involvement of "relational representations"), then is there a fundamental difference in the representation of these two kinds of information, as proposed by the authors? If there is, then how does one know where an entity ends and a relational representation begins? We believe our work and that of the rest of the visual memory community is more consistent with the idea that there is a *hierarchy* of representations, each requiring the storage of relational information and each allowing for the possibility of misbinding and other retrieval failures. Thus, rather than a strong dichotomy between entities and relational storage, the benefits of flexible, independent storage, and the resulting problem of binding features together, occur at every level of the hierarchy – from the simplest visual feature conjunctions to the binding of objects into contexts and into events.

Improving the integrative memory model by integrating the temporal dynamics of memory

Jonathan Curot[a,b,c] and Emmanuel J. Barbeau[a,b]

[a]Centre de Recherche Cerveau et Cognition, Université de Toulouse, Université Paul Sabatier Toulouse, Toulouse F-31330, France; [b]Centre National de la Recherche Scientifique, CerCo, UMR 5549, Toulouse F-31052, France and [c]Explorations Neurophysiologiques, Hôpital Purpan, Université de Toulouse, Toulouse F-31300, France.
jonathan.curot@cnrs.fr emmanuel.barbeau@cnrs.fr
http://www.cerco.ups-tlse.fr/~barbeau/

doi:10.1017/S0140525X19001973, e286

Abstract

Despite highlighting the role of the attribution system and proposing a coherent large-scale architecture of declarative memory, the integrative memory model would be more "integrative" if the temporal dynamics of the interactions between its components was clarified. This is necessary to make predictions in patients with brain injury and hypothesize dissociations.

"Integrative" is a major asset and is highly relevant to qualify the model presented by Bastin et al. in the target article. *Integration* is inseparable from *multimodality* and *multidimensionality*: the integrative memory model postulates that the systems processing representations, relations, and attributions are linked inside a coherent "architecture" allowing emergent properties. Within this context, one of the major advances proposed by the integrative memory model is the integration and the clarification of the role of the attribution system, which is thought to depend mostly on the prefrontal cortex. In contrast, most previous models of memory were centered on the temporal lobes and Papez circuit.

Some of the aspects of the temporal dynamics of memory that are currently not fully described in the model are: time perception during memory (Eichenbaum 2017a); time sequences that distinguish temporally distinct episodes and stimuli (Ekstrom & Ranganath 2018; Ranganath & Hsieh 2016); projection in the future (Addis & Schacter 2012); and the time scale for building memories at the cellular level (Kukushkin & Carew 2017). However, in this commentary we want to focus on another aspect of temporal dynamics that is essential to clarify the architecture of the integrative memory model. Because the integrative memory model, as its name implies, integrates different components, it is crucial to specify what kind of relation they entertain. This information is missing from the present model. The authors have devoted a large portion of the target article to describing the general architecture of the components, leaving little space to discuss exactly how they relate. (We think that their model could have been dubbed the interaction memory model just as well as the integrative memory model.)

Yet, although not fully specified, the integrative memory model is already based on a few assumptions regarding its temporal dynamics. For example, in line with many previous studies, familiarity is supposed to be rapid. The model also assumes that memory "emerges from *hierarchically* organized representations distributed throughout the brain" (target article, sect. 5.3, para. 1; emphasis added), which suggests a precise order in which the different components are activated. In contrast, most arrows connecting the different components of the model are bidirectional, perhaps due to the lack of knowledge about the connectivity between the components. However, the very presence of these arrows suggests structural and functional connections that have to be characterized.

Using behavioral reaction times for various memory tasks, it is possible to get an idea of the latency of the activation of some of these systems and such latencies can be used as upper time constraints. For example, behavioral paradigms based on time constraints can be used to precisely assess the speed of familiarity (Besson et al. 2012). Recording brain activity using surface EEG (electroencephalography) or MEG (magnetoencephalography), possibly with source reconstruction, or combined EEG-fMRI (functional magnetic resonance imaging) recordings (Hoppstädter et al. 2015) provides a more refined idea of the activation latencies of each component of the model. Intracranial EEG is spatially more precise and reveals, for example, a striking delay between the activity of the perirhinal cortex and the hippocampus that should be taken into account in models of memory (Barbeau et al. 2008; Trautner et al. 2004). Methodological advances even allow comparison of the neuronal activity of different medial temporal lobe regions involved in memory (Mormann et al. 2008). Moreover, it is also possible to calculate the strength of functional interactions between brain regions, as well as causality and synchrony indices, using various approaches such as fMRI (Staresina et al. 2013), intracranial EEG (Krieg et al. 2017; Kubota et al. 2013; Steinvorth et al. 2010), and thorough analyses of neuronal activity (Staresina et al. 2019).

In parallel, validating these dynamics in clinical situations is necessary. Alzheimer's disease – inducing slowly increasing damages to many brain areas involved in both the representation and attribution systems of the integrative memory model – is a pertinent example chosen by the authors. However, it is insufficient to test the model's dynamics. Experiential memory phenomena such as déjà-vu (an erroneous feeling of familiarity) or reminiscences (memories including a mental content and recollection) allow testing of the model on another time scale (Curot et al. 2017). These phenomena are highly transient – hundreds of milliseconds to a few seconds. This is the real-time scale of familiarity feelings, recollection, ecphory, and mental imagery. They become all the