Cognition 178 (2018) 7–25

Contents lists available at ScienceDirect

Cognition

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

Move to learn: Integrating spatial information from multiple viewpoints

Corinne A. Holmes[a,*], Nora S. Newcombe[b], Thomas F. Shipley[b]

[a] Trinity College Dublin, Ireland
[b] Temple University, United States

ARTICLE INFO

Keywords:
Spatial memory
Spatial flexibility
Spatial integration
Mental rotation
Perspective taking

ABSTRACT

Recalling a spatial layout from multiple orientations – spatial flexibility – is challenging, even when the global configuration can be viewed from a single vantage point, but more so when it must be viewed piecemeal. In the current study, we examined whether experiencing the transition between multiple viewpoints enhances spatial memory and flexible recall for a spatial configuration viewed simultaneously (Exp. 1) and sequentially (Exp. 2), whether the type of transition matters, and whether action provides an additional advantage over passive experience. In Experiment 1, participants viewed an array of dollhouse furniture from four viewpoints, but with all furniture simultaneously visible. In Experiment 2, participants viewed the same array piecemeal, from four partitioned viewpoints that allowed for viewing only a segment at a time. The transition between viewpoints involved rotation of the array or participant movement around it. Rotation and participant movement were passively experienced or actively generated. The control condition presented the dollhouse as a series of static views. Across both experiments, participant movement significantly enhanced spatial memory relative to array rotation or static views. However, in Exp. 2, there was a further advantage for actively walking around the array compared to being passively pushed. These findings suggest that movement around a stable environment is key to spatial memory and flexible recall, with action providing an additional boost to the integration of temporally segmented spatial events. Thus, spatial memory may be more flexible than prior data indicate, when studied under more natural acquisition conditions.

1. Introduction

Like all mobile organisms, humans need to learn the spatial layout of their environments. To survive, we must remember the location of food sources and shelter, and avoid areas where we have experienced threats. In achieving these goals, it is vital to be able to flexibly recall a global configuration from various vantage points in and around the space. For familiar spaces, spatial memory is quite flexible (e.g., Easton & Sholl, 1995; Holmes & Sholl, 2005; for discussion, see Meilinger, Riecke, & Bülthoff, 2007). That is, we can easily recall the spatial configuration of our daily environment (e.g., the layout of our kitchen, home, and neighborhood) from multiple orientations (e.g., from the front or back). However, spatial memory for novel spaces seems to be more rigid, with arrays best recalled from a limited number of vantage points – typically the experienced ones (Didwadkar & McNamara, 1997; Mou, McNamara, Valiquette, & Rump, 2004; Shelton & McNamara, 1997, 2001), although there may be other preferred orientations including those aligned with the intrinsic axis of the spatial array (Mou & McNamara, 2002; Mou, Zhao, & McNamara, 2007) or the extrinsic axis of the surrounding area (Adamou, Avraamides, & Kelly, 2014). When participants experience multiple viewpoints, they appear to encode them relative to the frame of reference of the first one encountered (Kelly & McNamara, 2008, Exp. 1; Shelton & McNamara, 2001, Exp. 7; Tlauka & Nairn, 2004), or to encode each representation relative to a unique frame of reference (e.g., Avraamides, Adamou, Galati, & Kelly, 2012; Meilinger, Strickrodt, & Bülthoff, 2016; for discussion, see Meilinger, 2008).

These results pose a puzzle—given the survival value of spatial flexibility, why is spatial memory for novel spaces seemingly so fragmented and rigid? One possibility is that spatial flexibility is effortful, i.e., it requires mentally transforming stored representations to make inferences about spatial relationships from other perspectives (e.g., Fields & Shelton, 2006; King, Burgess, Hartley, Vargha-Khadem, & O'Keefe, 2002; Street & Wang, 2014; Waller, Montello, Richardson, & Hegarty, 2002). For instance, people might imagine rotating a stored spatial configuration (i.e., mental rotation, MR), or they might imagine moving around it (i.e., perspective taking, PT). Even more mental work might be required when there are separate views that present fragmented spatial information encoded with respect to different reference frames, so that mental alignment is required to form a unified

---

* Corresponding author at: Trinity College, Institute of Neuroscience, College Green, Dublin 2, Ireland.
E-mail address: HOLMESCO@tcd.ie (C.A. Holmes).

https://doi.org/10.1016/j.cognition.2018.05.003
Received 25 September 2017; Received in revised form 26 April 2018; Accepted 1 May 2018

representation (Meilinger, Berthoz, & Wiener, 2011; Meilinger et al., 2016). The need to transform spatial information and to make inferences would exact a cognitive cost in terms of accuracy and/or response latency, and these effects have been confirmed (Adamou et al., 2014; Avraamides et al., 2012; Marchette, Ryan, & Epstein, 2017; Meilinger & Watanabe, 2016; Meilinger et al., 2011; Pantelides, Kelly, & Avraamides, 2016). Inferences would also recruit additional brain regions associated with the spatial processing network, and this pattern too has been confirmed (Mellet et al., 2000; Shelton & Gabrieli, 2002).

An alternative approach to spatial flexibility is suggested by considering how flexibility emerges as environments become more familiar. Much prior work has been done under restricted viewing conditions. Experiencing natural transitions among multiple viewpoints during encoding, as occurs in everyday experiences of walking around and through our environments, might lead to integration and spatial flexibility with less need for inference. Indeed, this kind of experience at environmental scale is what theoretical accounts of spatial navigation and wayfinding have emphasized (e.g., Gallistel, 1990). Some research implies the integrated representation has metric properties, as implied by the term "cognitive map", but other findings suggest a linked set of local spatial relations (e.g., Warren, Rothman, Schnapp, & Ericson, 2017).

In an initial study of viewing conditions, Holmes, Marchette, and Newcombe (2017) asked people to learn a tabletop environment viewed with continuous visual flow, with either active or passive movement of two kinds: the tabletop turned on its axis, or they circled the table. Relative to a control group given static snapshots, even passively-experienced visual flow enhanced spatial memory. No difference was observed between active or passive movement, between rotation (table turning) or perspective taking (movement around the table), or their interaction. Thus, these various modes of generating continuous visual flow apparently served similar functions, even though rotation and perspective change have proven to be distinct when they are imagined (e.g., Hegarty & Waller, 2004; Kozhevnikov & Hegarty, 2001; Lambrey, Doeller, Berthoz, & Burgess, 2012; Wraga, Shephard, Church, Inati, & Kosslyn, 2005; Zacks, Vettel, & Michelon, 2003).

Although visual flow may indeed be sufficient to enhance spatial memory in some situations, there were three limitations to this initial look at how movement may be related to spatial memory and flexible recall. First, this study, and many before it, examined spatial memory for a spatial array that could be viewed in its entirety from a single vantage point, thus only considering the effect of experience over time with continuous visibility. Here, we not only studied situations in which all spatial locations could be viewed simultaneously (Experiment 1), but also added the more challenging situation in which the array was viewed piecemeal and thus had to be integrated across views (Experiment 2). The demand to integrate across views begins to address the concern that many of the paradigms used to investigate navigation focus on small – not large-scale spatial cognition (e.g., Wolbers & Wiener, 2014). The second limitation relates to the nature of the spatial array used in our 2017 study. The tabletop environment was an organized layout of wooded areas, buildings and so forth, whereas much prior research has concerned collections of unrelated and discrete objects. It is possible that the more unified organization made the array relatively easy to encode as a whole, hence the equivalent advantage detected for array rotation and observer movement. Here, we use a more discretized spatial array – a dollhouse with separate pieces of furniture located in four rooms. Finally, the third limitation was that *we did not examine spatial flexibility* in our initial investigation. The spatial measures only examined spatial memory from a single vantage point, thus the ability to recall a spatial array from multiple orientations was never empirically assessed. In the current set of experiments, we addressed this point by using multiple vantage points for testing rather than a single one, and once again examined if the type of transition used to generate continuous visual flow between views – i.e., array rotation versus perspective taking – differentially impacts spatial memory and flexible recall, and if active movement is better than passive viewing.

Why might the type of transition matter? When extended spaces are experienced over time, they are generally experienced by walking through and around them and thus perspective change from the vantage point of a moving observer is a more natural way of encoding environmental space than having the space move. In fact, rotation is only possible with tabletop models. Furthermore, we already know that changing one's perspective at retrieval improves spatial performance compared to array rotation, whether it is accomplished by actual movement (e.g., Burgess, Spiers, & Paleologou, 2004; Simons & Wang, 1998; Wang & Simons, 1999; but see Motes, Finlay, & Kozhevnikov, 2006) or imagined movement (e.g., Creem, Wraga, & Proffitt, 2001; Kozhevnikov & Hegarty, 2001; Presson, 1982; Wraga, Creem, & Proffitt, 2000; Wraga, Creem-Regehr, & Proffitt, 2004; Wraga et al., 2005), and that as task difficulty increases, perspective taking is the preferred strategy for imaging a spatial array from alternate viewpoints (Kozhevnikov & Hegarty, 2001). Thus, one might expect that changing perspective during encoding would also be preferred.

Why might self-generated activity matter? The issue seemed worth probing again because the Holmes et al. (2017) studies stand in contrast with other work in this area. Several studies show that action provides an additional advantage over passive viewing, both when performed concurrently (e.g., Frick, Daum, Walser, & Mast, 2009; Gardony, Taylor, & Brunyé, 2014; Wexler & Van Boxtel, 2005) or when performed prior to imagined transformations (e.g., James, Humphrey, & Goodale, 2001; Wiedenbauer & Jansen-Osmann, 2008). These findings align with the idea that motor actions and mental operations are intrinsically intertwined (see Janczyk, Pfister, Crognale, & Kunde, 2012), and develop in tandem (e.g., Frick & Möhring, 2013). Neuroimaging studies provide support for the motor/mental connection, and show that mental transformations elicit activation in supplementary, pre-, and/or primary motor cortices (e.g., MR: Kosslyn, Thompson, Wraga, & Alpert, 2001; Vingerhoets, De Lange, Vandemaele, Deblaere, & Achten, 2002; see Zacks, 2008; PT: Creem et al., 2001; Vogeley et al., 2004; but see Wraga, Flynn, Boyle, & Evans, 2010; MR + PT: Wraga, Boyle, & Flynn, 2010; Wraga et al., 2005). Such findings imply that actively transitioning between viewpoints during learning may improve flexible recall and spatial integration. We hypothesized that active experience may be especially useful as the number of spatial locations increases (i.e., 20 locations versus the 8 used in Holmes et al., 2017; Exp. 1), or when the global configuration is viewed piecemeal and must be integrated across discrete experiences (Exp. 2).

2. Experiment 1

In Experiment 1, we examined the effect of viewpoint transitions on spatial learning and flexible recall when the global configuration of a complex scene could be viewed simultaneously. Participants viewed an array of dollhouse furniture from four viewpoints that presented the global configuration from multiple orientations in one of five between-subjects conditions. The control condition (Static Views, SV) presented the dollhouse as a series of temporally segmented views whereas in the remaining conditions, visual flow was continuous – participants viewed the natural transition from one room to the next. In the passive conditions, the experimenter generated the transition between rooms by rotating the dollhouse (Passive Array Rotation, PAR) or pushing the participant around it (Passive Perspective Taking, PPT). In the active conditions, participants generated each transition by manually rotating the dollhouse (Active Array Rotation, AAR) or walking around it (Active Perspective Taking, APT). Following encoding, participants completed a series of dependent measures to examine non-spatial and spatial memory. The spatial measures were of particular importance, and were designed to assess spatial memory from the preferred orientation and flexible recall from each of the four headings presented at encoding.