SPATIAL MEMORY AND NAVIGATION IN HUMANS

SPATIAL MEMORY AND NAVIGATION IN HUMANS

By XUE HAN, B.A (HONOURS)

A Thesis Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree Doctor of Philosophy

McMaster University

McMaster University DOCTOR OF PHILOSOPHY (2012) Hamilton, Ontario

(Psychology, Neuroscience, and Behaviour)

TITLE: spatial memory and navigation in humans

AUTHOR: Xue Han, B.A. (University of Windsor)

SUPERVISOR: Dr. Suzanna Becker

NUMBER OF PAGES: xvi, 174

Abstract

We investigated 1) how objects come to serve as landmarks in spatial memory and more specifically how they form part of an allocentric cognitive map and 2) how humans encode multiple connected spatial environments. In both sets of experiments, participants performing a virtual driving task incidentally learned the layout of a town and locations of objects or stores in that town. Their spatial memory and recognition memory for the objects or stores were subsequently tested. To assess whether the objects were encoded allocentrically, we developed a new measurement, pointing consistency. We found that when participants had more limited experience of the environment spatial memory for objects at navigationally relevant locations was more consistent across tested viewpoints than for objects at navigationally less relevant locations. When participants' attention was focused on the appearance of objects, the navigational relevance effect was eliminated, whereas when their attention was focused on the objects' locations, this effect was enhanced, supporting the hypothesis that when objects are processed in the service of navigation, rather than merely being viewed as objects, they engage qualitatively distinct attentional systems and are incorporated into an allocentric spatial representation. The results were consistent with evidence from the neuroimaging literature that when objects are relevant to navigation, they not only engage the ventral "object processing stream", but also the dorsal stream and medial temporal lobe memory system classically associated with allocentric spatial memory. Moreover, in the connected environments, our data were more consistent with the formation of local maps, regardless of whether the neighborhoods were learned together or separately. Only when all visible distinctions between neighborhoods were removed did people behave as if they formed one integrated map. These data are broadly consistent with evidence from rodent hippocampal place cell recordings in connected boxes, and with hierarchical models of spatial coding.

Acknowledgment

I would like to thank my supervisor Dr. Suzanna Becker, to whom I am deeply indebted in terms of both academic and personal support. Also, I would like to thank my supervisory committee Dr. Patrick J. Bennett and Dr. Hong-Jin Sun for their academic support.

I would like to dedicate this thesis to my parents, Mr. Peiran Han and Mrs. Jianhua Zhang, and my husband, Dr. Xiaowei Yu for their unconditional caring and love.

Table of Content **Pages**

List of Figures and Tables **Pages**

Declaration of Academic Achievement

The purpose of the research reported in this dissertation was to explore and understand some of the underlying mechanisms involved in spatial cognition. Specifically, we focused on the contribution of landmarks and objects to egocentric vs. allocentric representations, and the formation of multiple maps. The following two chapters have either been previously published or are under preparation for submission. I am the major contributor to all aspects of the work.

Chapter 2 is a manuscript published in *PloS One*, 2012, volume 7, pages 1-19 by Han, X., Byrne, P., Kahana, M., and Becker, S. titled “When do objects become landmarks? A VR study of the effect of task relevance on spatial memory”. The author of the current thesis is the first author of this published work whose contribution includes the development of the original idea and experimental design, data collection and analysis, and manuscript preparation. The second author is a previous graduate student who programmed some of the test stimuli and provided insightful comments on the manuscript. The third author is a professor at University of Pennsylvania who developed the virtual taxi game used in this research and provided insightful comments on the manuscript. The thesis supervisor is the last author of this paper who provided feedback at all stages.

Chapter 3 is a manuscript in preparation for submission by Han, X., and Becker, S. titled “One spatial map or many? A VR study of how we encode multiple connected environments”. The author of the current thesis is the first author whose contribution includes the development of the original idea and experimental design, data collection and analysis, and manuscript preparation. The thesis supervisor is the last author of this paper who provided feedback at all stages.

1 Chapter One: General Introduction

Spatial memory is necessary for navigation, an essential survival skill for humans and non-human animals. For example, the ability to go out searching for food and then safely return to the nest requires spatial knowledge of the environment and the ability to navigate in the environment. There are different types of navigational strategies one can rely on such as path integration, landmark-based navigation, and cognitive map-based navigation. Moreover, there are many cues one can use, which may be broadly grouped into idiothetic cues and allothetic cues. The idiothetic cues are internal cues such as sensory and proprioceptive input from joints, muscles and the vestibular system of the inner ear during self-movement, while allothetic cues are external cues such as landmarks encountered during navigation.

1.1 Different navigational strategies

Previous research has identified a number of different strategies for navigation, including path integration, route-based navigation (e.g. Dyer, 1991; Etienne, Maurer, Boulens, Levy, & Rowe, 2004; Maaswinkel & Whishaw, 1999; Mather & Baker, 1981; Mittelstaedt & Mittelstaedt,1980; von St. Paul, 1982; Whishaw, Hines, & Wallace, 2001), landmark-based navigation (e.g. Cartwright & Collett, 1982, 1983, 1987; Cheng, 1988; Collett & Land, 1975; Dyer & Gould, 1983), and cognitive map-based navigation (e.g. Burgess, 2006; Etienne & Jeffery, 2004; Foo, Warren, Duchon, & Tarr, 2005; Gould, 1986; Janson, 2007; Kohler & Wehner, 2005; Noser & Byrne, 2007; Tolman, 1948). It is important to note that these different types of navigational strategies are not necessarily mutually exclusive.

1.1.1 Path Integration

Path integration requires one to constantly update one's location with respect to one's starting point by keeping track of the distance and direction travelled. Many species including humans (e.g. Klatzky, Loomis, Golledge, Cicinelli, Doherty, & Pellegrino, 1990), rodents (e.g. Etienne et al., 1998; Séguinot, Maurer, & Etienne, 1993; Siegrist, Etienne, Boulens, Maurer, & Rowe, 2003), birds (e.g. von St. Paul, 1982), and insects (e.g. Müller & Wehner, 1988; Wehner & Wehner, 1986) can employ path integration, which relies on internal idiotheic cues, for navigation. For example, hamsters were able to go to the learned feeding site by only relying on locomotion (Etienne et al., 1998) and return to the nest in complete darkness, even when no other cues were available (Siegrist et al., 2003). Similarly, rats were able to use self-movement cues to return home in the absence of visual or any other reliable external location cues (Maaswinkel et al., 1999). Humans also have this ability to some degree. For example, they were able to return directly to the starting point after travelling a multisegment pathway while blindfolded (Klatzky et al., 1990). However, in the process of continuously updating one's location, errors accumulate (Loomis, Klatxky, Golledge, Philbeck, 1999, for a review). Therefore, external references (allothetic cues) are used to aid in path integration when they are available. For example, ants used skylight patterns as directional cues (Duelli & Wehner, 1973) or wind and sky as compass cues (Müller & Wehner, 2007) in dead-reckoning. However, when the visual references and path integration information were in conflict, rodents relied more on the previously learned visual references (Etienne, Maurer, & Séguinot, 1996; Teroni, Portenier, & Etienne, 1987).

Some brain regions are known to be important for path integration. For instance, rats with lesions to the fimbria-fornix (a major fibre tract connecting the hippocampus with other brain regions) were impaired in path integration (Whishaw & Maaswinkel, 1998). Disruption in the posterior parietal cortex in humans affected path integration (Seemungal, Rizzo, Gresty, Rothwell, & Bronstein, 2008). Moreover, rats with entorhinal cortex or parietal cortex lesions were not able to use path integration to accurately return to the starting point (Parron & Save, 2004) and rats with associative parietal cortex lesions were not able to acquire locomotion-generated information (Save & Moghaddam, 1996). However, the involvement of the hippocampus in path integration is under debate, as some studies have shown that hippocampectomized rats were impaired in path integration (Maaswinkel, Jarrard, & Whishaw, 1999) while others have found that hippocampectomized rats were able to return home via path integration (Alyan & McNaughton, 1999). The reason for these discrepant findings might be due to the degree of damage to the hippocampus, and precisely which sub-regions or pathways were damaged. In Masswinkel and colleagues' study, they found extensive loss of cells in the dentate gyrus but not the fimbria-fornix, whereas in Alyan and McNaughton's study, they found that the dentate gyrus and the most ventral-temporal CA3 were spared.

1.1.2 Landmark-based navigation

Landmark-based navigation requires one to encode the relationship between a goal and the surrounding landmarks. It has been observed in many species including humans (e.g. Smith, Gilchrist, Cater, Ikram, Nott, & Hood, 2008), rodents (e.g. Cheng, 1988; Collett, Cartwright, & Smith,1986; Etienne, Teroni, Hurni, & Portenier, 1990; Suzuki, Augerinos & Black, 1980), birds (e.g. Cheng, 1988; Cheng & Sherry, 1992; Spetch & Edwards, 1988), and insects (e.g. Cartwright et al., 1982, 1983; Collett et al., 1975; Dyer et al., 1983; Wystrach, Schwarz, Schultheiss, Beugnon, & Cheng, 2011). For example, even young children aged between 3 and 7 could use natural landmarks to reorient themselves (Smith et al., 2008). Bees and hoverflies were able to store ‘snapshots’ of the environment which included landmarks, and used them to navigate along a route by matching the incoming images to the stored images (Cartwright et al., 1982, 1983; Collett et al., 1975; Wehner, 1972). However, these ‘snapshots’ were specific retinal images and not governed by the general concepts derived from landmark configurations (Dyer, 1991; Wehner & Räber, 1979). In contrast to bees and hoverflies, gerbils could use both independent landmarks and internal representations of all the landmarks to find a goal (Collett et al., 1986), birds could use distances to edges in addition to the landmarks to find goal locations that involving distance calculation as seen in path integration (Cheng et al., 1992), and desert ants could use both path integration and landmark learning strategies

(Müller & Wehner, 2010; Wehner & Müller, 2010), which suggest that these two types of strategies can operate in concert rather than being mutually exclusive. Moreover, other global orienting cues could be used to help navigation such as skylight and wind (Duelli et al, 1973; Müller et al., 2007). If the above strategies failed, insects would engage in systematic searching behavior for homing (Wehner & Srinivasan, 1981).

With the importance of landmark-based navigation well established in many species, further research has shown that the retrosplenial cortex and parahippocampal cortex in the brain are two of the regions known to play a vital role for this type of navigation. For instance, the human retrosplenial and parahippocampal cortices responded to landmark identities but not to the distances between landmarks (Morgan et al., 2011).

1.1.3 Cognitive map-based navigation

In addition to the path integration and landmark-based navigation strategies, many species are capable of navigating using an internal mental model of an environment. Tolman (1948) introduced the concept of a cognitive map. In Tolman's view, a cognitive map includes routes, paths, environmental features and relationships between these features, as well as relative positions of different places (O'Keefe & Nadel, 1978; Tolman, 1948). If an animal forms an internal cognitive map, it is argued that shortcuts could be taken without prior experience of that specific route. More generally, the current location relative to any other locations in the environment can be calculated.

Whether insects have ‘cognitive maps’ has been debated (e.g. Bennett, 1996; Cartwright et al., 1987; Collett & Collett, 2006; Giurfa & Capaldi, 1999; Menzel, Geiger, Joerges, Müller, & Chittka, 1998; Wehner, Boyer, Loertscher, Sommer, & Menzi, 2006; Wehner & Menzel, 1990). For example, Gould (1986) first suggested the idea that honeybees used cognitive maps to navigate, rather than just following specific serial lists of landmarks in a route, because when they were transported to an unfamiliar region before or after foraging, they were able to fly directly to the foraging site or fly directly back to the hive, respectively. However, Wehner and Menzel (1990) suggested that with careful control, bees did not show any sign of cognitive map-based behaviour after been released from an unfamiliar site, but instead, they were using landmark-based information and path integration in homing. On the other hand, in Menzel and colleagues’ displacement experiment using radar-tracking records, when landmarks were not visible and no path integration could be used at the release site, honey bees made truly novel short cuts suggesting the use of map-like mental representations (Menzel et al., 2005). Therefore, no definite conclusion as to whether insects have cognitive maps could be made (see Collett & Collett, 2006, for a review).

In contrast to insects, the majority of studies suggested that rodents (e.g. Chamizo, Rodrigo, & MacKintosh, 2006; Chapuis, Durup, & Thinus-Blanc, 1987; Nesterova & Hansen, 2009; Roberts, Cruz, & Tremblay, 2007; Schenk, Grobéty, & Gafner, 1997; Tolman, 1948), canines (e.g. Chapuis & Varlet, 1987), and some bird species (e.g. Blaisdell & Cook, 2005; Gibson & Kamil, 2001) have the ability to use cognitive map-based information in navigation. Although Benhamou (1996) claimed that without the opportunities to use other orientation mechanisms, rats were not able to navigate accurately towards a hidden goal, many other studies showed that rats were able to learn the relative positions of different places (e.g. Schenk et al., 1997) and choose correct novel routes and shortcuts in an enclosed maze when no other external cues were available (Roberts et al., 2007), suggesting the rats do have 'cognitive maps'. Similarly, Columbian ground squirrels (Nesterova et al., 2009) and dogs (Chapuis & Varlet, 1987) were able to take detours to find food sources. Furthermore, pigeons (Blaisdell et al., 2005), hamsters (Chapuis et al., 1987), and rats (Chamizo et al., 2006) were able to integrate two different representations of landmark configurations, relative positions of tables, or maze arm arrangements, respectively, supporting their employment of internal representations of the environments.

In contrast to the somewhat mixed findings from insect and rodent studies, it is generally accepted that humans (e.g. Burgess, 2006; Tolman, 1948) and non-human primates (e.g. Janson, 2007; Normand & Boesch, 2009; Noser et al., 2007) have cognitive maps. For example, wild chacma baboons were able to take detours to avoid others (Noser et al., 2007), wild capuchin monkeys chose which foraging site to go based on information including the distance along a detour and reward magnitude (Janson, 2007), and wild chimpanzees went to the same food resource from many different directions (Normand et al., 2009), suggesting that these primates have cognitive map-based information of their living environments.

Besides the findings of behavioural studies supporting the cognitive map-based navigation, brain studies also show evidence of cognitive map-based navigation, and implicate the hippocampus cognitive map formation, in both human (e.g. Abrahams, & Pickering, Polkey, & Morris, 1997; Maguire, Burke, Phillips, & Staunton, 1996; Morgan, MacEvoy, Aguirre, & Epstein, 2011) and non-human animals (e.g. Sutherland, Kolb, & Whishaw, 1982). For example, the left hippocampus responded to distances between landmarks in human (Morgan et al., 2011); temporal lobe damage including the hippocampus impaired encoding of locations of landmarks and the spatial relationships among them (Maguire et al., 1996); right hippocampal formation damaged patients had deficits in spatial memory in the radial arm maze (Abrahams et al., 1997); and bilateral lesions in either the hippocampus or medial frontal cortex in rats impaired the ability to learn spatial representations of the environment (Sutherland et al., 1982).

In addition to the hippocampus, the parahippocampal region (Aguirre & D'Esposito, 1996), postsubiculum (Goodridge et al., 1995), and retrosplenial cortex (Takahashi, Kaeamura, Shiota, Kasahata, & Hirayama, 1997) are also involved in topographical memory tasks and likely form part of a larger circuit that is involved in cognitive map formation. Accordingly, three different streams within the so-called dorsal "where" or "how-to" pathway have been proposed for spatial processing: a parieto-prefrontal branch involved in visuospatial working memory, a parieto-premotor branch involved in visually guided action and a parieto-medial temporal branch involved in spatial navigation (Kravitz, Saleem, &

Mishkin, 2011). It is the latter branch that is the main focus of this thesis.

1.2 Types of mental representations

Given the multitude of navigational strategies discussed up to this point, we now consider what kind of spatial representations are formed in the brain to support navigation in an environment. It is widely held that our spatial memory system supports at least two distinct representations of scenes and objects: egocentric or viewer-centered and allocentric or world-centered (e.g. Becker & Burgess, 2001; Byrne, Becker, Burgess, 2007; Burgess, 2006, Epstein, Higgins, & Thompson-Schill, 2005; Gramann, Müller, Schönebeck, & Debus, 2006; Mou, McNamara, Rump, & Xiao, 2006).

1.2.1 Egocentric representations

Egocentric representations encode locations with respect to the observer, in various body-centered reference frames (e.g., eye-centered, head-centered or body-centered) and may be updated by self-motion (translation and/or rotation). In both the object recognition and spatial memory literature, similar debates have been going on for many years as to which type of representation people use. An individual object could be encoded either as a collection of egocentric or view-based snapshots, or within an allocentric representation. Similarly, object configurations and scenes could be stored either as egocentric snapshots or as allocentric memories. This thesis focuses on the representation of object configurations and scenes rather than on how individual objects are represented. Some have argued for purely egocentric encoding of object configurations, for example, based on evidence that memory for object locations can be disrupted by disorientation (e.g. Wang & Spelke, 2000), and is best when the tested viewpoint (imagined heading) is congruent with one of the studied viewpoints (Roskos-Ewoldsen, McNamara, Shelton, & Carr, 1998; Shelton and McNamara, 1997). On the other hand, as discussed below, others have found evidence for allocentric encoding of object configurations.

In terms of path integration, there may be multiple mechanisms and strategies at play. Although the process of path integration by definition requires dynamically updated information about travel direction and distance using body-referenced idiothetic cues (Mittelstaedt & Mittelstaedt, 1980), the representation of space upon which path integration operates could be either egocentric or allocentric. Several brain regions have been implicated in path integration. Damage to the posterior parietal cortex affected spatial updating of one's position (Farrell & Robertson, 2000) and the precuneus was also implicated in egocentric spatial updating (Wolbers, Hegarty, Buchel, & Loomis, 2008). It has been suggested that parietal circuits may be involved in path integration when egocentric working memory representations of objects are updated after egomotion (e.g. Byrne, Becker, & Burgess, 2007). On the other hand, place-specific firing patterns of entorhinal grid cells and hippocampal place cells (see next section) are known to be updated as the rodent moves around in the dark (Hafting, Fyhn, Molden, Moser, & Moser, 2005; Quirk, Muller, & Kubie, 1990) and are generally believed to perform path integration within an allocentric reference frame. Similarly, landmark-based navigation could be based on egocentric representations as well, if the animal encodes the locations of landmarks as a series of stored egocentric visual snapshots, or it could be based upon allocentric features, which may be combined together at the level of the hippocampus into clusters of landmarks and boundaries that form place representations.

1.2.2 Allocentric representations

Allocentric representations are based on external reference frames, which are independent of one's own body. Like egocentric representations, they could also be updated by self-motion. Compared with egocentric representations, allocentric representations should be less affected by changes in one's viewpoint, and should permit the observer to deal with novel viewpoints with no cost or relatively little cost. There is a large body of behavioural evidence supporting the use of allocentric representations (e.g. Bugress, 2006; Burgess, Spiers, & Paleologou, 2004). For example, a patient who had developmental amnesia with perinatal focal bilateral hippocampal pathology had impaired performance on a shift viewpoint recognition task that required allocentric representation (King, Burgess, Hartley, Vargha-Khadem, O'Keefe, 2002).

Animals that use a cognitive map-based navigational strategy and in some cases a path integration strategy require allocentric representations of environments. The findings of place cells, grid cells, boundary vector cells, head direction cells, view cells, and EC path cells from animal and human studies support the existence of allocentric representations.

Place cells, recorded in the hippocampus and medial entorhinal cortex of freely moving rats, discharge at specific locations, and in the open field, are independent of heading direction (O'Keefe & Dostrovsky, 1971; O'Keefe, 1976). Place cells have also been found in humans (Ekstrom et al., 2003) and non-human primates (e.g. Matsumura, Nishijo, Tamura, Eifuku, Endo, & Ono, 1999; Ono, Nakamura, Nishijo, & Eifuku, 1993). It has been argued that place cells form the basis of a cognitive map (O'Keefe & Nadel, 1978). Correspondingly, damage to the hippocampus impairs performance on allocentric spatial memory tasks (e.g. Maguire et al., 1996; Sutherland et al., 1982), suggesting that the hippocampus is a critical structure for the formation of allocentric representations, across a wide range of species.

Grid cells have multiple firing fields within a single environment, arranged in a grid-like pattern; they have been found in the medial entorhinal cortex (Fyhn et al., 2004; Hafting, Fyhn, Molden, Moser, & Moser, 2005; Sargolini et al., 2006), an area connecting the hippocampus with other association cortices, and in the presubiculum and the parasubiculum (Boccara et al., 2010) in rats. Firing patterns of grid cells are also governed by cues such as geometric boundaries and landmarks in an environment, which, sometimes, has higher priority than path integration-driven information (Moser, Kropff, & Moser, 2008).

Boundary vector cells (BVCs), whose firing is only dependent on the location of an animal relative to the egocentric distance to and allocentric direction of boundaries of an environment, have been found in the subiculum (Lever et al., 2009).

Head direction cells are tuned to an animal's head direction in the horizontal plane, independent of its location within the environment; they have been found in many parts of the rat's limbic system such as the postsubiculum (Ranck, 1984; Taube et al., 1990), the anterior dorsal nucleus of the anterior thalamus (Taube, 1995), the dorsal sector of the caudal lateral dorsal thalamic nucleus (Mizumori & Williams, 1993), the lateral mammillary nuclei (Leonhard, Stackman, & Taube, 1996), the agranular and granular areas of retropsplenial cortex (Chen, Lin, Green, Barnes, & McNaughton, 1994) and other brain regions such as the posterior parietal cortex (the rostral medial prestriate in rat) (Chen et al., 1994). Firing of head direction cells is usually stable within the same recording environment over weeks, but alters with changes of allothetic cues (Taube et al., 1990) and remaps in new environments (Dudchenko & Zinyuk, 2005). Moreover, head direction cells may modulate the orientation of grid cell maps (see Moser et al., 2008 for a review) and their discharge could depend on input from both internal and external sensory cues; however, under certain situations when the two types of cues are in conflict, landmark cues (allothetic cues) could override the internal sensory cues (Goodridge & Taube, 1995).

View cells, found in the hippocampus and parahippocampal region in humans (Ekstrom et al., 2003) and non-human primates (Rolls & O'Mara, 1995**),** fire when one is looking at an object or a location, independent of one's location or head direction. View cells are sensitive to visual input (Rolls et al., 1995) and important for viewpoint change tasks (see Burgess, Maguire, & O'Keefe, 2002, for review). These cells provide information for forming spatial memory of an environment such as where the objects have been seen from which point.

EC path cells, found in the entorhinal cortex, fire in response to the direction of the path traveled (Jacobs, Kahana, Ekstrom, Mollison, & Fried, 2010), providing inputs to the head direction cells, from which one could construct allocentric representations for cognitive map-based navigation.

1.2.3 Egocentric vs. Allocentric representations

People appear to use different spatial representations for different purposes. Presson and Montello (1994) suggested that spatial working memory might depend on a viewer-centered framework. For example, we appear to use egocentric, or viewer-centered, representations when we generate mental imagery in order to guide our actions in space. On the other hand, long-term spatial memory may employ either egocentric or allocentric representations, or a combination of the two.

Another view is that allocentric and egocentric strategies coexist (Burgess, 2006), and recruit separate neural circuits, differentially governed by the traditional rules of associative learning (e.g. blocking and overshadowing) (Doeller & Burgess, 2008). Even within the allocentric circuit, it is an overa-simplification to think of spatial representations as being purely allocentric, because sensory input is inherently egocentric (e.g. retinocentric for visual input);

thus, it follows that there must be a hierarchy of representations from egocentric to allocentric levels, as per the BBB model (Byrne, Becker & Burgess, 2007). Egocentric information about the spatial locations of objects, encoded from spatial information available to the dorsal visual pathway, may be combined with object appearance information from the ventral visual pathway to form allocentric, configural representations of spatial environments in long-term memory at the level of the medial temporal lobe/hippocampal region. Conversely, if long-term memories are stored in an allocentric form, in order to make use of the product of memory retrieval for the purpose of navigation, pointing etc, allocentric memories must be transformed back into an egocentric form. Thus, memories about spatial configurations are hypothesized to be retrieved from (allocentric) long-term memory in the hippocampus and mapped through reciprocal neuronal pathways to generate egocentric mental images (Byrne et al., 2007). According to this view, whenever an allocentric (e.g. place cell based) memory is activated, one has to access it via egocentric cues; conversely, whenever an allocentric memory is retrieved and used to direct a motor response or for internal planning, the product of the retrieval is transformed back into an egocentric mental image. Thus, both the input and output of this top level allocentric code are probably egocentric. This egocentric-allocentric multilevel system can be contrasted with a purely egocentric system which might be used for well learned routes (stored e.g. as sequences of body turns associated with landmarks or choice points), which might be processed in the basal ganglia or through the dorsal visual pathway to the level of the parietal cortex, rather than all the way up to the level of the hippocampus.

1.3 Current research

The current research focuses on both landmark-based and cognitive map-based navigation in humans. It is clear that a variety of species use landmark-based navigation strategies, however, there is some ambiguity as to what constitutes a landmark. It is less clear when objects would be treated as landmarks, not just objects, and how they are represented in the spatial memory. One possibility is that the object's visual features (e.g., size, contrast, etc) inherently differentiate it as a landmark. Another possibility (perhaps not mutually exclusive) is that a given object may or may not be treated as a landmark, according to how it is processed. For example, when they are relevant to navigation or attention was paid to their locations, they might be much more likely to recruit allocentric spatial memory circuits. Therefore, we investigated under what situations an object would be treated as a landmark and incorporated into an allocentric representation of the environment in Chapter 2. We found evidence that suggests objects at navigationally relevant points are more likely to be treated as landmarks and incorporated into an allocentric representation of the environment compared to those at less navigationally relevant points. Moreover, when attention was manipulated to focus on the appearance of objects, all the objects were treated as objects, not landmarks regardless of their positions in the environment.

In Chapter 3, we investigated whether multiple representations would be created in spatial memory when humans experienced a large complex environment consisting of multiple distinct sub-regions. We found evidence that when there were two connected neighborhoods within a town, humans had a very strong tendency to create separate maps of the town.

2 **Chapter Two**

When do objects become landmarks? A VR study of the effect of task relevance on spatial memory.

2.1 Foreword

Previous studies have shown that the hippocampus is crucial for the formation of cognitive maps (e.g. Abrahams, & Pickering, Polkey, & Morris, 1997; Maguire, Burke, Phillips, & Staunton, 1996; Morgan, MacEvoy, Aguirre, & Epstein, 2011; Sutherland, Kolb, & Whishaw, 1982). The parahippocampal region (Aguirre & D'Esposito, 1996), a major input to the entorhinal cortex/hippocampus, is also involved in topographical memory tasks. These regions in the medial temporal lobe likely form part of a larger circuit involved in cognitive map formation. These regions represent the highest level of the parieto-medial temporal branch of the dorsal visuo-spatial pathway, a branch that is involved in spatial navigation (Kravitz, Saleem, & Mishkin, 2011).

Evidence of the involvement of dorsal visuo-spatial pathway in spatial navigation comes from neuroimaging studies. For example, medial temporal and medial parietal structures were activated for recognizing objects at navigationally relevant points (Janzen & van Turennout, 2004). However, it remains to be demonstrated whether navigational relevance causes a switch in favor of allocentric encoding of objects by activating the dorsal visuo-spatial pathway. Thus, the experiments reported here were designed to test the hypothesis that navigational relevance would modulate the degree to which objects would be integrated within their spatial context into allocentric spatial maps. To assess the degree of viewpoint invariance of object memory, we developed a novel VR pointing task and a novel performance metric – pointing consistency across tested viewpoints. The results are consistent with evidence from the neuroimaging literature that when objects are relevant to navigation, they not only engage the ventral “object processing stream”, but also the dorsal stream and medial temporal lobe memory system classically associated with allocentric spatial memory.

2.2 Introduction

In everyday life, an object may be attended to individually, or may be processed within the spatial context of a scene. Traditionally these two styles of processing are associated with two major branches of the visual system, the ventral “what” stream and dorsal “where” or “how to” stream (Goodale & Milner, 1992; Kravitz, Saleem, Baker, & Mishkin, 2011; Ungerleider & Mishkin, 1982). More specifically, when viewed within its spatial context, a single object could be encoded either within a viewpoint-dependent (egocentric) representation, e.g. as a visual snapshot memory, or within a world-centered or allocentric frame of reference. Accordingly, the notion of a single dorsal visual stream for spatial processing has been updated to include three branches: a parieto-prefrontal branch involved in visuospatial working memory, a parieto-premotor branch involved in visually guided action and a parieto-medial temporal branch involved in spatial navigation (Kravitz, et al., 2011). When there are multiple objects, the possibilities are even greater. A collection of objects could be treated as a whole entity and encoded as one configuration, either as an egocentric/view-based snapshot or allocentrically based on inter-object relations. Many studies have tested memory for sets of objects on rotating tabletops to investigate this type of encoding (e.g. Mou, Xiao, & McNamara, 2008). Alternatively, each of the objects could be encoded allocentrically with respect to features of the environment such as buildings or geometric cues. Finally, as we argue here, the brain may employ simultaneously a hierarchy of representations, from egocentric representations of sensory information to allocentric representations in long-term memory. This could allow some egocentric cues, such as a familiar viewpoint or a dominant reference direction within an environment, to have preferential access into allocentric long-term memory.

Some early empirical studies led to rather polarized views on how objects are encoded. For example, memory for object locations can be disrupted by disorientation (e.g. Wang & Spelke, 2000), and is best when the tested viewpoint (imagined heading) is congruent with one of the studied viewpoints (Roskos-Ewoldsen, McNamara, Shelton, & Carr, 1998; Shelton and McNamara, 1997). These data seem inconsistent with an orientation-invariant representation of object configurations. We return to this issue later in the introduction, where we discuss the use of combined egocentric and allocentric representations, as in the BBB model. On the other hand, if one learns the environment by directly experiencing it from multiple perspectives, as opposed to by studying a map, spatial memory of the relation between items is more robust to viewpoint rotations, suggesting allocentric encoding of objects (Evans & Pezdek, 1980). Whereas viewpoint-invariance could simply arise from storing multiple view-based snapshots, strong neurobiological support for allocentric representations comes from evidence of place cells --- neurons that respond selectively when an animal is in a given location. Such cells, which have been identified in the hippocampi of rats (O’Keefe, 1976), non-human primates (Ono, Nakamura, Nishijo, & Eifuku, 1993) and humans (Ekstrom et al, 2003), are often insensitive to the animal’s heading within the environment, suggesting that they encode spatial location within an allocentric representation. Moreover, hippocampal damage impairs allocentric memory function. For example, an individual who suffered perinatal hippocampal pathology showed highly impaired memory for arrays of objects when tested from unfamiliar viewpoints, in spite of highly accurate memory when tested from familiar viewpoints (King, Trinkler, Hartley, Vargha-Khadem, & Burgess, 2004).

The mixed evidence in support of egocentric vs. allocentric representations likely reflects people’s ability to use both types of representation. Methodological differences such as passive versus active navigation and exposure to few versus multiple viewpoints may contribute to the type of processing people engage in. Participants in the Evans et al. (1980) and King et al. (2004) studies learned the environment by active navigation, whereas those in the Roskos-Ewoldsen et al. (1998) and Shelton et al. (1997) studies learned the environment from one or two static views. Thus, active navigation, and/or exposure to a dynamically changing range of views of the environment, may encourage allocentric strategies. Consistent with this notion, rodent place cells tend to be omnidirectional when recorded in the open field but unidirectional when recorded in a linear track or narrow-armed mazes (McNaughton, Barnes, & O’Keefe, 1983; Muller, Bostock, Taube, & Kubie, 1994; O’Keefe & Recce, 1993; Redish, McNaughton, & Barnes, 2000). When humans take a path around the square road in a virtual environment one observes both unidirectional place cells (as in the rat) and also path cells that are sensitive to the direction of motion independent of the (virtual) location within the environment (Jacobs et al., 2010).

An emerging view is that allocentric and egocentric representations coexist and recruit different levels of representation (Burgess, 2006). If we accept that incoming visual input is by definition egocentric (i.e. retinocentric), and that we have the capacity to create allocentric representations (e.g.hippocampal place cells), it follows that allocentric representations of the world can only be constructed from egocentric inputs. Thus, when we encode information, we have the option of employing a purely egocentric strategy or a combined strategy that includes multiple levels, mapping from egocentric to allocentric frameworks. It is likely that we have developed specialized circuits that may be predominantly egocentric, or may also include allocentric representations. This is supported by a wide range of evidence from behavioural, neuroimaging and brain lesion studies in humans and other animal species (e.g. Galati, Lobel, Vallar, Berthoz, Pizzamiglio, & Bihan, 2000; Gramann, Müller, Schönebeck, & Debus, 2006; Hölscher & Schmidt, 1994; Jordan, Schadow, Wuestenberg, Heinze, & Jäncke, 2004; King, Burgess, Hartley, Vargha-Khadem, & O’Keefe, 2002; Mou, McNamara, Rump, & Xiao, 2006; Valiquette & McNamara, 2007; Werner & Schmidt, 1999). These two types representations are differentially governed by the traditional rules of associative learning (e.g. blocking and overshadowing) (Doeller & Burgess, 2008) and vary according to task demands (e.g. Mou, et al., 2006; Mou, Xiao, & McNamara, 2008; Valiquette & McNamara, 2007). Even when performance is behaviourally equivalent, when people employ allocentric representations they activate distinct neural circuits (Jordan et al., 2004). Thus, wayfinding and other allocentric spatial tasks recruit a common neural circuit including the parietal cortex, retrosplenial cortex, fusiform gyrus, precuneus, parahippocampal gyrus, hippocampal complex and several prefrontal cortical regions, while non-spatial navigation tasks such as learning a series of body turns recruit an associative learning circuit involving the striatum (including the caudate nucleus and putamen), insula/ventrolateral prefrontal cortex, and right anterior prefrontal cortex (e.g. Hartley, Maguire, Spiers, & Burgess, 2003; Iaria, Petrides, Dagher, Pike, & Bohbot, 2003; Jordan et al., 2004; Maguire et al., 1998; Morgan, MacEvoy, Aguirre, & Epstein, 2011).

Given the abundant evidence for allocentric representations, an important question is how allocentric representations could arise out of purely egocentric (i.e. retinocentric) sensory input. Byrne, Becker, and Burgess (2007) proposed a computational model, which we shall refer to as the BBB model, suggesting that egocentric information about the spatial locations of objects from the dorsal visual pathway is combined with object appearance information from the ventral visual pathway to form allocentric, configural representations of spatial environments in long-term memory at the level of the hippocampus. Conversely, memories about spatial configurations can be retrieved from (allocentric) long-term memory in the hippocampus and mapped through reciprocal neuronal pathways to generate egocentric mental images. Note that individual objects may also be represented allocentrically within the ventral visual pathway, e.g., there is evidence for view-invariant representations of single objects within inferotemporal cortex [33]; this type of object-based allocentric representation must be distinguished from the *configural* allocentric representations of scenes referred to here, mediated by the medial temporal lobe. Because the BBB model postulates that egocentric level representations provide access cues to allocentric long-term memory, it naturally accommodates preferred viewpoint effects, for example, as defined by intrinsic frames of reference formed from egocentric experiences and environmental cues cues (Mou, & McNamara, 2002; Mou, McNamara, Valiquette, & Rump, 2004; Mou, Xiao, & McNamara, 2008; Shelton & McNamara, 2001). Note, however, that the BBB model does not incorporate the non-spatial associative learning circuit mentioned above. The role of the dorsal visual pathway in the BBB model encompasses both the parieto-medial temporal branch (Kravitz, et al., 2011) for forming allocentric representations, and the parieto-frontal branch (Kravitz, et al., 2011) for maintaining and updating object locations in working memory after real or imagined observer motion.

The BBB model postulates some of the neural mechanisms that may underlie allocentric spatial memory, but it does not tell us what sort of features might contribute to the creation of these memories. As mentioned above, one important factor that may contribute to allocentric coding of features is their utility for spatial memory and navigation. For example, objects placed at choice points should be particularly relevant to navigation. Several experiments have examined the impact of navigational relevance on object recognition memory. Janzen and van Turennout (2004) had participants passively view a movie of a tour through a virtual museum with objects placed at T-shaped intersections (decision points) and simple L-shaped turns (non-decision points), and directed their attention more to some of the objects (toys) than others, half of which were placed at decision points. Although both types of locations lead to a change in one's direction, and as such, could both be considered as decision points (e.g. Schinazi & Epstein, 2010), we adhere to the terminology as used by Janzen & van Turrenout throughout this manuscript. While recognition memory accuracy was not affected by navigational relevance or attention, reaction times were faster for attended objects (toys), and fastest of all to the toys that had been seen at decision points. Moreover, functional imaging studies show greater activation of medial temporal and medial parietal structures associated with spatial cognition (including the hippocampus, parahippocampal cortex, superior parietal lobule/precuneus, parietal-occipital sulcus, retrosplenial/anterior calcarine region) for navigationally relevant objects in recognition memory (Janzen & van Turennout, 2004) and object priming tasks (Janzen and Weststeijn, 2007; Schinazi & Epstein, 2010), and also greater activation in these regions when encoding virtual environments containing landmarks (salient objects) compared to encoding a plain empty virtual environment (Maguire, Frith, Burgess, Donnett, & O'Keefe, 1998). These studies suggest that objects are not always just objects: when they are relevant to navigation, they are much more likely to recruit allocentric spatial memory circuits. One potential confound with the above studies on navigational relevance is that objects at decision points may be inherently more salient. Furthermore, they assessed recognition memory and priming, but not spatial memory. Miller and Carlson (2010) used a setup similar to Janzen and van Turennout's (2004) with an explicit manipulation of object salience, and measured both recognition memory and spatial memory (map drawing). They found that spatial memory for decision-point objects was still superior even when they were less salient than non-decision-point objects, whereas recognition memory was strongly modulated by salience. Thus navigational relevance seems to strongly modulate whether objects are incorporated into spatial memories.

The studies reviewed above suggest that 1) people use both egocentric and allocentric strategies for spatial memory and navigation, depending on task demands, and 2) the hippocampal and parahippocampal regions are crucial for allocentric spatial memory formation and are recruited for encoding objects that are relevant to navigation. It remains to be demonstrated whether navigational relevance causes a switch in favor of allocentric encoding of objects. Thus, the experiments reported here were designed to test the hypothesis that navigational relevance would modulate the degree to which objects would be integrated within their spatial context into allocentric spatial maps. To assess the degree of viewpoint invariance of object memory, we developed a novel VR pointing task and a novel performance metric – pointing consistency across tested viewpoints.

We conducted three experiments to test whether objects would be encoded differently based on navigational relevance, and whether the type of attention paid to objects would modulate this effect. Whereas Janzen and van Turennout (2004) and Miller and Carlson (2010) had participants passively view image sequences of a virtual environment, we wanted a more life-like task where people actively control where they go, how long they spend in each location, and what they pay attention to. They should thereby construct an internal representation of an environment using whatever features are most relevant to navigation and spatial orienting. We constructed a set of virtual towns with grid-like streets lined with stores using Kahana's "Yellow Cab" virtual taxicab simulator (http://memory.psych.upenn.edu/Research). Using this same task, in human intra-cranial recordings, Ekstrom and colleagues (2003) found evidence of place cells and view cells in the human medial temporal lobe, indicating that even this relatively simplistic task and artificial environment engages the standard spatial memory circuits and evokes allocentic spatial representations (see, also, Jacobs et al., 2010). We asked participants to pretend to be a taxi driver in the town and look for and deliver passengers. We placed objects at certain locations in the town, half at decision points (T-shaped intersections) and half at non-decision points (L-shaped intersections). Participants implicitly learned the stores and object locations by playing the taxi game, and were then given tests of recognition memory and spatial memory for the objects after each of the study phases. In spatial memory test trials, memory for the locations of the objects was probed from two different viewpoints, which were views of the town from the two end-points, marked by “Mike’s Restaurant” and “House of Pizza” respectively.

In Experiment 1, participants learned the layout of the virtual town via active navigation, while pretending they were taxi drivers looking for and delivering passengers. In Experiment 2, participants learned the town layout passively by watching videos of trajectories through a town. We also included a between-subjects manipulation in Experiments 1 and 2 to vary the number of starting points that participants would experience. In our study, Experiments 1 and 2 each had two conditions, one in which participants started navigation trials alternatingly from two points, creating two salient viewpoints/reference directions from which spatial memory could be accessed, and one in which participants always started from the same point, creating one salient viewpoint/reference direction (Mou & McNamara, 2002; Mou, Zhao, & McNamara, 2007; Valiquette, McNamara, & Labrecque, 2007). Consistent with previous research by McNamara and colleagues, we expected that participants’ spatial memory would be superior when tested from the most salient viewpoint when they always started from the same end of the town. We also hypothesized that having experienced the town from two different starting points, participants would tend to approach the objects from multiple directions and would thus be more likely to form view-invariant representations of those objects.

One potential confound in Experiments 1 and 2 is that objects at decision points may be attended to more strongly or for more time than objects at other locations. Thus even if memory for objects at decision points is superior, it does not necessarily mean those objects were processed via a different neural circuit or a qualitatively different mechanism. To address this issue, in Experiment 3, we explicitly manipulated the type of attention devoted to objects, by instructing participants to only focus on either the appearance or the location of objects. We hypothesized that the type of attention would modulate the effect of navigational relevance, that is, memory for objects at decision points should benefit from spatial attention and should be hurt by attention to object appearance.

A key issue in the present study is how best to assess the degree of viewpoint invariance of participants' spatial memory for objects. Most previous research in object spatial memory has employed small rooms within which all objects could be viewed from a single location (e.g. Wang & Spelke, 2000; Mou & McNamara, 2002) or a rotating tabletop upon which the entire configuration of objects could be viewed simultaneously (e.g. Mou, Xiao, & McNamara, 2008). In these studies, various measures of memory for object configurations have been employed, such as "configuration error" (Wang & Spelke, 2000) and judgments of relative direction (e.g. ("Imagine you are at the A and facing the B. Point to the D.") (e.g. Mou, & McNamara, 2002; Shelton & McNamara, 2001). These measures of errors in memory for inter-object relations are suitable for testing hypotheses about memory for object configurations, but do not address our main question of whether objects are encoded relative to environmental and geometric cues. In our experiments, we use large virtual towns, with streets lined with buildings and shops, and objects located all around the town. Thus, in our experiments, the objects could not be directly perceived as a configuration within a single location, but would have to be learned individually by actively navigating in the town or watching video tours of the town, integrating the information over larger spatial and temporal extents. We thus expected participants would encode each object relative to the surrounding visible environmental features. We used a novel method to assess viewpoint invariance of spatial memory for objects across different locations within the environment. We calculated the consistency of pointing responses to each object (see Method section) made from two different viewpoints at opposite ends of the town. We reasoned that if participants were encoding object locations relative to an allocentric spatial map of the town, they should make consistent pointing errors when tested from either viewpoint. For example, if an object was in the middle of the town and they mis-localized it by 45 degrees clockwise when pointing from one end of the town, they should mis-localize it by about the same amount and in the opposite direction, 45 degrees counterclockwise, when pointing to it from the opposite end of town. Therefore, we developed a measure of pointing consistency across tested viewpoints. We acknowledge that accuracy for accessing spatial memories from perspectives 180 degrees from the stored perspective is better than from other perspectives (e.g. 45 degrees or 135 degrees), however, it is still worse than accessing it directly from the stored perspective (e.g. Mou & McNamara, 2002; Valiquette & McNamara, 2007; Valiquette, McNamara, & Smith, 2003). Furthermore, in large-scale environments, accessing from the opposite direction of the stored perspective was found to be no easier than from other directions (Werner & Schmidt, 1999).

To summarize our predictions, we hypothesized that 1) objects at more navigationally relevant locations (decision points) should be encoded as landmarks, and become incorporated within an internal cognitive map of space. Although access to this internal representation via egocentric cues could be biased along a preferred orientation (as per Valiquette, McNamara and Smith, 2003), the navigational relevance of objects within the environment should still modulate the degree to which their internal representation is sensitive to changes in viewpoint. 2) When participants experienced the town from fewer viewpoints they should be even less likely to employ allocentric strategies for objects, particularly those at locations not relevant to navigation. Thus, reducing the number of starting points should reduce the number of familiar viewpoints, and thereby enhance the effect of navigational relevance. 3) When participants' attention was manipulated to focus on objects' appearance, the decision-point effect would be eliminated, whereas when participants' attention was directed toward objects' locations, this effect would persist or even be enhanced.

2.3 Experiment 1

Participants implicitly learned the town layout and object locations by playing a virtual taxi game requiring active navigation through a virtual town. We varied the navigational relevance of objects in the environment by placing them either at decision points or non-decision points. In Condition A, participants started passenger pickups alternatingly from the two ends of the town marked by House of Pizza and Mike’s Restaurant respectively, and were subsequently tested from both of those viewpoints. This would establish two salient viewpoints/reference directions, which were also the tested viewpoints, from which either type of object could be encoded. In Condition B, participants entered the town from only one direction, facing the House of Pizza, thereby establishing only a single salient viewpoint/reference direction during study. Nevertheless, in both conditions, we tested participants’ memory from the same two viewpoints, one facing House of Pizza and the other facing Mike’s Restaurant. In Condition B, by always having the participants start navigating from one end of town rather than two, we introduced an encoding bias. If indeed decision-point objects were encoded as part of an allocentric map of the town whereas non-decision-point objects were not, spatial memories for decision-point objects should be less affected by this manipulation relative to other objects. Therefore, we predicted that the pointing responses would be less accurate and less consistent for non-decision-point objects in Condition B relative to those in Condition A, but memory for decision-point objects should be similar across the two conditions, if objects at decision points were encoded as part of an allocentric map, relative to other objects. Moreover, when participants experienced the two tested viewpoints equally, the difference between decision points and non-decision points would be reduced.

2.3.1 Method

2.3.1.1 Participants

Sixty McMaster University students ranging in age from 18 to 25 years (mean age 19.63) participated in the experiment. There were 30 participants in each condition (12 males and 18 females in Condition A; 13 males and 17 females in Condition B). Participants had normal or corrected-to normal vision, and received partial course credit or $10 for taking part in the experiment. This study was reviewed and approved by the McMaster University Research Ethics Board. Written informed consent was obtained from all participants involved in this study.

2.3.1.2 Materials

We employed Kahana's "Yellow Cab" virtual driving simulator (see http://memory.psych.upenn.edu/Research) for constructing the environment and for simulating the virtual taxi game for the study phase of the experiment. There were two rectangular shaped towns (14 by 9 VR units in size) with different layouts (see Figure 2.1). Each participant experienced only one of the two towns. In each town, there were two distinctive buildings, Mike's Restaurant (see Figure 2.2A) and House of Pizza (see Figure 2.2B), respectively located at the two ends of the town, which marked the two alternative starting points for each passenger pickup and also the two tested viewpoints for the spatial memory tests. There were four stores designated as passenger drop-off or goal locations. The two starting points were not used as drop off locations. There were 19 objects (see

Figure 2.3) used in the experiment, 8 of which appeared in the town and the remaining eleven of which were only shown at the object pre-exposure phase and served as distractors for the subsequent recognition memory task. Four of the 8 objects in the town were placed at decision points (T-shaped intersections), where the participants could decide to turn right, turn left, or continue straight. The other four objects were placed at non-decision points (L-shaped intersections), where the participants could only turn in one direction. The locations of decision-point and non-decision-point objects were matched pairwise with respect to their distance from the town midline and distance from the town end-point, that is, with the viewer at location (0,0), for every decision-point object at location (x, y) there was a corresponding non-decision-point object at location (-x, y) (see Figure 2.1). Thus, during the pointing task with the participant placed at either end of the town, the average distance of objects to the town midline and to the observer was equal for the two groups of the objects. The locations of individual objects in each town remained constant across blocks. During the study phase, participants used either the joystick or the arrow keys on the keyboard to control their navigation, allowing them to turn in any direction, control their speed, or do a U-turn.

The memory test was implemented in Matlab with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). On each memory test trial, there was a half compass shaped figure (a navigator) with pictures of multiple views of the actual town seen from either one end or the other end of the town on the top of the navigator (see Figure 2.4). Images of different views of the two were shown at the ends of the compass lines, so that the view from straight ahead was located at the top of the compass (forward direction), the views when looking to the left to varying degrees were located at corresponding points to the left of forward, etc. Additionally, at the tip of the compass pointer an image of the target object for the current trial appeared; this target object moved with the pointer. The target was always one of the 19 objects (8 of which had been in the town and 11 of which were distractors), and the participant could use the mouse to rotate the compass pointer (along with the target) to point in the remembered direction of the target object from the displayed viewpoint and make a mouse-click to indicate a pointing response or press the space bar to indicate that an object was not recognized as having been in the town.

2.3.1.3 Procedure

There was a total of four blocks, each consisting of an object pre-exposure, a study and a test phase. In the object pre-exposure phase, each of the 19 objects appeared for two seconds followed by a blank screen for one second. This established some degree of familiarity of the distractor objects so that the subsequent recognition memory task would be more challenging. In each study phase, the participant was asked to act as a taxi driver whose task was to roam around and find passengers and deliver them to specific locations (i.e., stores). A trial began with the participant located at one of the two ends of the town, facing toward the middle of the town, and he/she was asked to freely navigate until a passenger was found and "collected" by bumping into the passenger. A textual cue then appeared, e.g. “Please take me to the Flower Patch store, I will give you 100 points”, and the participant was instructed to navigate as quickly and efficiently as possible to drop off the passenger to the goal location by bumping into it. In Condition A, each time the participant dropped off a passenger, he or she was re-located to the opposite end of the town from the previous trial, facing either Mike’s Restaurant or House of Pizza, before being cued to collect the next passenger. In Condition B, participants were always relocated to House of Pizza after each pick up. There were five passenger deliveries in each of the four blocks, hence a total of 20 passenger deliveries. The participant’s location and viewing direction were recorded every 30-40 ms throughout the entire study phase. The memory test combined simultaneous tests of recognition memory and spatial memory. On each trial, if the participant thought the object had not appeared in the town, he or she pressed the ‘space bar’, and the next object would be displayed. Otherwise, he or she then pointed in the direction of the object’s remembered location from the displayed viewpoint (see Figure 2.4) by using the mouse to move the compass pointer to the desired direction, and then pressing the left mouse button. We did not measure recognition memory reaction time separately, but we did measure pointing latency, as our memory test combined recognition and spatial memory. For both conditions A and B, in each memory test phase, the participant had to respond to each object twice, once from each end of the town (see Figure 2.4). The recognition responses, pointing directions and total reaction time for the combined spatial/recognition memory response were recorded during the memory test phase.

2.3.1.4 Data Analysis

In McNamara et al's experiments, a gender effect was not consistent: in most of their studies, there was no gender effect (e.g. Kelly & McNamara, 2008; Mou & McNamara, 2002; Mou, Xiao, & McNamara, 2008; Shelton & McNamara, 2004; Valiquette & McNamara, 2007; Valiquette, McNamara, & Labrecque, 2007) and in other studies, males were more accurate (Valiquette, McNamara, & Smith, 2003) or faster (Mou, McNamara, Valiquette, & Rump, 2004) than females. Moreover, Lavenex and Lavenex (2010) did not find a gender effect on spatial relational learning. Gender is not a focus in our study here and we did not find a gender effect in our measurements. Therefore, we did not include gender as a factor in our analyses here.

Bonferroni correction was used for all the multiple comparisons throughout this paper. Loftus and Masson's (1994) method was used to calculate 95% confidence intervals shown in the figures.

2.3.1.4.1 Recognition Accuracy

The accuracy of the participants' recognition memory was calculated as follows: If the participants indicated that they had seen the object, but the object was not used in the town, the response was counted as a false positive. If the participant indicated that they had not seen the object, but the object was used in the town, the response was counted as a false negative. Otherwise, the response was counted as a correct recognition. We calculated percent correct recognition separately for objects at decision points and those at non-decision points, averaged across blocks. We then compared the difference in recognition accuracy between decision and non-decision points in Conditions A and B by using a two way repeated measures Place x Condition ANOVA.

2.3.1.4.2 Pointing Latency

There were two reaction time scores (in seconds) for each object in each block: one for each of the two tested viewpoints. We averaged the reaction time over correct responses across blocks for objects at decision points for each viewpoint, and did the same for those at non-decision points to get pointing latency for decision and non-decision points for each of the two tested viewpoints. We compared the pointing latencies using a three-way repeated measures ANOVA, with Place (decision vs. non-decision points) and tested Viewpoint (Mike's Restaurant vs. House of Pizza) as within-subject factors and Condition (one vs. two starting points) as a between-subject factor.

2.3.1.4.3 Pointing Error (Average Absolute Pointing Errors)

In each block, participants had to point to each object once from each of the two starting points. Therefore, each participant had two pointing responses for each object, one from the viewpoint of Mike's Restaurant and one from the viewpoint of House of Pizza. A pointing error was defined as the signed value, in degrees, of the difference between the pointing response and the object's actual direction (participant's response in degrees minus the object's correct direction in degrees). Therefore, for each starting location we have four raw signed pointing errors for decision-point objects and four for the non-decision-point objects for each participant in each block if the participants correctly recognized all the objects used in the town. The signs of the pointing errors from one end of town were reversed so that consistent spatial memory errors for the same object from the two viewpoints would have the same sign; for example, a pointing error of ten degrees clockwise from Mike's Restaurant and ten degrees counterclockwise from House of Pizza (the opposite end of town), after this sign change correction, would be coded equivalently as signed errors of +10.

We calculated the absolute value of all the raw signed pointing errors for each participant across blocks for decision points and similarly for the non-decision points. Thus, each participant had two average absolute pointing errors for each of the two tested viewpoints, one for decision points and one for non-decision points. We used a three -way repeated measures ANOVA to compare differences in pointing error between decision and non-decision points and between the two tested viewpoints in the two conditions (one vs. two starting points).

2.3.1.4.4 Pointing Consistency: Standard Deviations

To test the hypothesis that participants were more likely to incorporate decision-point objects into an allocentric map of space, we developed a measure of viewpoint consistency in pointing errors. We reasoned that if a participant is using an allocentric representation of an environment to recall an object's location, their pointing errors for that object should be consistent across the tested viewpoints, regardless of overall magnitude. Thus, if an objects' location is remembered accurately from one viewpoint, it should be equally accurately remembered when tested from the other viewpoint. On the other hand, if an object is remembered incorrectly, resulting in a high pointing error from one viewpoint, the participant should make an error of the same magnitude but opposite sign when tested from the opposite viewpoint. In contrast, if s/he has an egocentric representation of an object's location from a given direction within an environment, the pointing errors made between familiar and unfamiliar viewpoints would be more variable, because the participant may have to mentally rotate the representation in order to align it with the familiar stored viewpoint. We used the standard deviation of pointing responses (signed pointing errors) across the two tested viewpoints as a measure of the consistency of the pointing responses. If, for example, a participant consistently mis-located an object as being 10 degrees clockwise when tested from the Mike's Restaurant viewpoint and 10 degrees counterclockwise when tested from the House of Pizza viewpoint, the signed errors for this object would both be +10 and the standard deviation across the two viewpoints would be zero. Note that we counted pointing errors in the clockwise direction from the Mike's Restaurant viewpoint and counterclockwise from the House of Pizza viewpoint as both positive errors, while counterclockwise errors from Mike's and clockwise errors from House of

Pizza were counted as negative errors, see Pointing Error (Average Absolute Pointing Errors) for details.

The pointing consistency across the two tested viewpoints was calculated by taking the standard deviation of the two signed pointing errors that the participant made for each decision-point object from the two ends of the town, and then averaging these standard deviation scores across objects and across blocks at decision points, and similarly averaging those at non-decision points. Any object with less than two signed pointing errors was dropped from the consistency analysis. We thereby obtained two average pointing consistency scores for each participant, one for decision-point objects and one for non-decision-point objects. We analyzed these scores with a two-way repeated measures ANOVA (Place x Condition) to test the hypothesis that pointing responses across two tested viewpoints would be more consistent for objects at decision points than for other objects, particularly when there was only one starting point.

2.3.1.4.5 View Time

To assess whether participants spent more time viewing objects at decision points than viewing other objects during study, we calculated the “view time” of each object, for each block in each study phase, as the percentage of the total time that participants spent at locations where their facing direction placed the object within their field of view, and then calculated average percentage view times for the two types of objects. We averaged these scores across blocks and then used a two-tailed paired sample *t*-test to assess differences in view time.

2.3.2 Results

2.3.2.1 Recognition accuracy

Recognition accuracy was better when there was a single starting point and it was better for decision point objects.

A two-way repeated measures Place (decision point vs. non-decision point) x Condition (one starting point vs. two) ANOVA revealed significant main effects of Place [*F* (1,58)=8.706, *p*=0.005] and Condition [*F* (1,58)=5.342, *p*=0.024], but no interaction between Place and Condition [*F* (1,58)=0.455, *p*=0.503]. Recognition accuracy was significantly better for objects at decision points (*mean*=90.7%, *SE*=0.011) than for those at non-decision points (*mean*=87.1%, *SE*=0.013) across conditions. Unexpectedly, recognition memory was also significantly better when participants used one starting point (Condition B *mean*=91.4%, *SE*=0.015) than two (Condition A: *mean*=86.4%, *SE*=0.015).

2.3.2.2 Pointing latency

Response latency was faster when there was a single starting point, but it was not affected by navigational relevance or viewpoint.

A three-way repeated measures Place (decision points vs. non-decision points) x Condition (one starting point vs. two) x tested Viewpoint (Mike's Restaurant vs. House of Pizza) ANOVA of the pointing/recognition latencies revealed a significant main effect of Condition [*F* (1,58) = 5.705, *p* =0.02], but no significant effect of Place [*F* (1,58) = 0.024, *p* =0.877] or tested Viewpoint [*F* (1,58) = 1.388, *p* =0.244] and no significant interactions. Responses were significantly faster in Condition B (one starting point) (*mean*=4.578, *SE*=0.409) than in Condition A (two starting points) (*mean*=5.959, *SE*=0.409) across object types and tested viewpoints.

2.3.2.3 Pointing errors

Navigational relevance affected pointing accuracy when there was a single starting point.

A three-way repeated measures Place x Condition x tested Viewpoint ANOVA of the pointing errors revealed significant main effects of Place [*F* (1,58) = 6.751, *p* =0.012] and tested Viewpoint [*F* (1, 58) = 7.369, *p* =0.009], and significant interactions between Place and Condition [*F* (1, 58) = 5.964, *p* = 0.018] and between tested Viewpoint and Condition [*F* (1,58) = 14.275, *p* < 0.001], but no main effect of Condition [*F* (1,58) = 3.047, *p* = 0.086] alone and no other significant interactions, see Table 1. Pointing errors were significantly smaller for objects at decision points (*mean*=26.479, *SE*=1.254) than those at non-decision points (*mean*=28.239, *SE*=1.441) across conditions and pointing errors were significantly smaller when they were made from House of Pizza viewpoint (the starting point in Condition B, *mean*=26.043, *SE*=1.295) than when they were made from the Mike's Restaurant viewpoint (*mean*=28.675, *SE*=1.487) across both types of object and both conditions (one start point or two).

To further investigate these significant interactions, we conducted separate two-way repeated measures Place x tested Viewpoint ANOVAs for the two conditions. In Condition A, when participants started alternatingly from both ends of the town during study, there was no main effect of Place [*F* (1, 29)=0.013, *p*=0.911] or tested Viewpoint [*F* (1, 29)=1.012, *p*=0.323], and no interaction between Place and tested Viewpoint [*F* (1, 29)=0.065, *p*=0.801]. In Condition B, when participants always started from the same end of the town, there were significant main effects of Place [*F* (1, 29)=12.082, *p*=0.002] and tested Viewpoint [*F* (1, 29)=14.625, *p*=0.001], but no interaction between Place and tested Viewpoint [*F* (1, 29)=0.521, *p*=0.476]. Pointing errors were significantly smaller for objects at decision points (*mean*=27.934, *SE*=1.822) than for objects at non-decision points (*mean*=31.349, *SE*=2.171) across viewpoints and pointing errors were significantly smaller when they were made from House of Pizza viewpoint (the starting point, *mean*=26.493, *SE*=1.935) than when they were made from the less familiar Mike's Restaurant viewpoint (*mean*=32.79, *SE*=2.272) across object types in Condition B.

2.3.2.4 Pointing consistency

Navigational relevance affected pointing consistency when there was a single starting point.

A two-way repeated measures Place x Condition ANOVA of pointing consistency scores revealed significant main effects of Place [*F* (1, 58) = 7.794, *p* =0.007] and Condition [*F* (1, 58) = 4.964, *p* = 0.03] and a significant interaction between Place and Condition [*F* (1, 58) = 8.264, *p* = 0.006]. Pointing responses were significantly more consistent for objects at decision points (*mean*=20.659, *SE*=1.013) than for those at non-decision points (*mean*=23.272, *SE*=1.37) across conditions, and were significantly more consistent in Condition A (*mean*=19.491, *SE*=1.571) than in Condition B (*mean*=24.44, *SE*=1.571) across object types. To further investigate the interaction between Place and Condition in terms of pointing consistency, two-tailed paired sample *t*-tests were used. In Condition A, there was no difference in pointing consistency by object location (*t* = 0.095, *df* = 29, *p* = 0.925). In Condition B, pointing responses were significantly more consistent at decision points than at non-decision points (*t* = -3.147, *df* = 29, *p* = 0.004). Moreover, because we hypothesized that reducing the number of starting points (Condition B) would reduce pointing consistencies for objects at non-decision points. Two-tailed independent *t*-tests showed that pointing consistencies for objects at decision points were no different between the two conditions (*t* = -1.115, *df* = 56.21, *p* = 0.27), but significantly worse for objects at non-decision points (*t* = -2.787, *df* = 49.006, *p* = 0.008) in Condition B than Condition A (Note: controlling for multiple comparisons, significant *p* value is 0.0125; Equal variances were not assumed). The analysis revealed that the navigational relevance effect was only significant in the single-starting-point condition (Condition B), in which pointing responses were more consistent for objects at decision points than for those at non-decision points, but not in the two-starting-point condition (Condition A). Moreover, reducing the number of starting points during the study phase detrimentally affected the pointing consistencies for objects at non-decision points, but had little effect on objects at decision points, see Figure 2.5.

2.3.2.4.1 Pointing Consistency: Correlation between two tested viewpoints

One reason pointing responses were more consistent (i.e. less variable) across viewpoints in the case of objects at decision points could simply be that the pointing errors themselves were smaller for decision-point objects. Even if the pointing responses from the two viewpoints were uncorrelated, smaller magnitude pointing errors would lead to smaller standard deviations in pointing errors. To rule out this possibility, we also employed a secondary measure of pointing consistency that is insensitive to overall error magnitude: The Pearson correlation coefficient (*r*) between the two signed pointing errors at the two tested viewpoints across blocks was calculated for each type of object. Therefore, each participant had one Pearson's *r* for decision-point objects and one for non-decision-point objects. Because we hypothesized that the decision-point objects would be less affected by viewpoint changes, these correlation scores were compared using one-tailed nonparametric Wilcoxon Signed Rank tests, which revealed that signed pointing errors were significantly more correlated for objects at decision points (*mean r*=0.377, *SE*=0.052) than those at non-decision-point objects (*mean r*=0.213, *SE*=0.068) (*p* = 0.014) in Condition B (one starting point), but not in Condition A (decision points *mean r*=0.326, *SE*=0.057; non-decision-point objects *mean r*=0.402, *SE*=0.055) (*p*=0.1495) (Note: controlling for multiple comparisons, significant *p* value is 0.025) Thus, the correlation analysis was in complete agreement with our standard deviation measure of pointing consistency, indicating that memory for decision-point objects in Condition B was more view-invariant, and not just more accurate.

2.3.2.4.2 Consistency of signs of pointing errors

Another limitation of our pointing consistency measure is that it is sensitive to the locations of the objects in the town, such that if an object was closer to one end of the town than the other, even if the participant consistently mis-localized it to the same location from both ends of the town, the angular error magnitudes would differ. This is not a confound, because it is equally true for both decision and non-decision-point objects. However, an alternative measure that is insensitive to the angular error magnitude is the consistency of the signs of the errors. If a participant consistently mis-localizes an object, for example, clockwise from one end and counterclockwise from the other end, the signs of the errors would be consistent. Note: The signs of the pointing errors from one end of town were reversed; see Method-Exeperiment 1-Data Analysis-Pointing Error (Average Absolute Pointing Errors) for details. There were 4 pairs of pointing responses made for 4 decision-point objects and another 4 for the four non-decision-point objects in each block, if the participant correctly identified all of the objects, and there were 4 blocks. We calculated the percentage of pairs of pointing errors that had the same sign over blocks for decision-point objects and then for non-decision-point objects. Because we hypothesized that the decision-point objects would be less affected by viewpoint changes, one-tailed Wilcoxon Signed Rank tests were used revealing that the percentage of same signed pointing errors for decision points was significantly higher than those for the non-decision-point objects in Condition B (*p*=0.0135, DPs mean=66.93%, SE=0.033; NDPs mean=58.14%, SE=0.032), but not in Condition A (*p*=0.457, DPs mean=62.88%, SE=0.035; NDPs mean=62.26%, SE=0.035) (Note: controlling for multiple comparisons, significant *p* value is 0.025).Thus, the analysis of consistency of signs of pointing errors was in complete agreement with our standard deviation measure of pointing consistency, indicating that memory for decision-point objects in Condition B was more view-invariant, and not just more accurate.

2.3.2.5 View Time.

A two-way repeated measures Condition x Place ANOVA of view time revealed a significant main effect of Place [*F* (1, 58) = 146.56, *p*<0.0001] and a significant interaction between Place and Condition [*F* (1, 58) = 7.571, *p* = 0.008], but no main effect of Condition [*F* (1, 58) = 0.021, *p* = 0.885]. Viewing time for objects at decision points (*mean*=29.9%, *SE*=0.004) was longer than for objects at non-decision points (*mean*=21.4%, *SE*=0.005) across conditions (Condition A: DP *mean*=28.9%, *SE*=0.006, NDP *mean*=22.3%, *SE*=0.007; Condition B: DP *mean*=30.9%, *SE*=0.006, NDP *mean*=20.5%, *SE*=0.007). To investigate the interaction between Place and Condition, two-tailed paired sample *t*-tests were conducted, which revealed that viewing time for decision-point objects was significantly longer than for non-decision-point objects in both conditions (*p*s<0.001).

2.3.2.5.1 View time correlations.

View time was not correlated with spatial memory accuracy or consistency.

Given the significant difference in viewing time between objects at decision points and non-decision points, any potential differences we might observe in spatial memory for these objects in the current experiment could be due to more attention and encoding time being devoted to decision-point objects (a potential confound). Alternatively, the viewing time differences may be entirely due to participants engaging other processes at decision points, such as imagining the route along alternate paths and making navigation decisions. While the lack of spatial memory differences between the two types of objects in Condition A (two starting points) suggests the latter interpretation, viewing time differences could still be a potential confound in Condition B (one starting point). If the reason participants spent more time viewing decision-point objects was partly due to greater time devoted to attending to and encoding those objects' locations, we would expect viewing time to correlate with memory for those objects. We therefore assessed whether any of the pointing error and consistency measures were correlated with view time for both decision and non-decision points. These correlational analyses revealed that none of our memory measures were significantly correlated with viewing time.

2.3.3 Discussion

We hypothesized that navigational relevance would strongly modulate whether objects were treated as landmarks and encoded within an allocentric cognitive map, particularly when objects were seen from a limited range of viewpoints. Although participants were free to navigate around the town and potentially approach each object from multiple directions, the single starting point would bias participants to approach each object from fewer directions, on average. This led to our prediction that spatial object memory would be more accurate and more viewpoint-invariant for objects at decision points than for other objects, particularly when we reduced the number of starting points. Our results confirmed this prediction. While the two object types showed differences in viewing time and recognition memory accuracy in both conditions, there was no effect of navigational relevance on any of the spatial memory measures in Condition A, where participants used two different starting points. On the other hand, in Condition B, when there was only one starting point, spatial memory for objects that were not at decision points suffered, such that pointing responses were less accurate and less consistent across the two tested viewpoints. Thus, as predicted, spatial memory for non-decision point objects was sensitive to the number of starting points, whereas spatial memory for decision point objects was less affected.

Interestingly, when participants began navigation from both ends of the town, the navigational relevance effect was not merely diminished but disappeared altogether, see Figure 1. One reason for this lack of effect of navigational relevance in Condition A could be that when experienced from more viewpoints, even objects at “non-decision points”, i.e. L-intersections, come to be treated as landmarks. Although L-intersections are less navigationally relevant than T-intersections, they do involve a turn in the route and are thus more relevant when compared to straight portions of a route. Future studies could investigate this possibility, by including objects along straight roads. Another possibility is that when experienced from both ends of the town, the objects at non-decision points were encoded as multiple egocentric snapshots. In either case, pointing consistency differences between decision and non-decision-point objects would disappear. One way to tease apart these two alternative explanations would be to repeat the fMRI study by Janzen and van Turrenout (2004) in which participants viewed a trajectory through a virtual museum containing objects at both decision and non-decision points. However, rather than viewing the tour in one direction only, they could view the tour in both directions as in our Experiment 2. If this caused a switch from egocentric to allocentric / dorsal visual stream encoding for the objects at non-decision points then those objects should now activate the parahippocampal region.

Unexpectedly, the number of starting points also affected pointing response latencies and recognition accuracy, but in the opposite direction to the consistency effects. As mentioned above, pointing errors were more consistent in Condition A, the condition with two starting points. In contrast, pointing responses were faster and recognition memory was more accurate in Condition B, the single starting point condition, across both object types and both tested viewpoints. One possible explanation for these results is that some participants were using a mental navigation strategy to recall object locations. Such a strategy would be fastest when there was a single starting point, and more likely to break down as the number of to-be-remembered routes increased. Individual differences in strategy are often seen in spatial cognition studies, and certainly warrant further investigation in the tasks studied here.

Not surprisingly, we saw an effect of the specific viewpoint in Condition B: In the case of a single starting point at House of Pizza, pointing errors were smaller from the more familiar House of Pizza viewpoint than from Mike's Restaurant viewpoint for both types of objects. This is consistent with the findings of Mou and Colleagues (Mou, & McNamara, 2002; Mou, et al., 2004; Mou, et al., 2008; Shelton & McNamara, 2001) reviewed in the introduction, and fits within the BBB model which postulates that egocentric retrieval cues are used to index long-term allocentric memory (Byrne, et al., 2007).

Importantly, in spite of the preferred viewpoint effect on both types of objects, our pointing consistency analysis revealed that spatial memory for the two types of objects was differentially affected by the reduced number of starting points in Condition B (relative to Condition A). Pointing errors were significantly more consistent across the tested viewpoints for objects at decision points than for those at non-decision points, using the standard deviation (pointing consistency), the correlation analysis and the consistency of signs of pointing errors analysis. This finding is consistent with the hypothesis that objects at decision points are more likely to be incorporated within an allocentric map, less affected by the number of salient reference directions, and more robust to changes in viewpoint at test time. It also supports our claim that pointing consistency across tested viewpoints is a useful measure of allocentric coding when objects are seen in large-scale spaces, as opposed to being viewed from a single location.

Another possible explanation for the superior spatial memory for decision-point objects in Condition B is that they were not encoded in a qualitatively different manner, but were simply better encoded than were non-decision-point objects. For example, participants may have devoted more attention to decision-point objects. Consistent with this alternative interpretation, recognition memory was superior and viewing times were longer for these objects. However, it is important to note that our “view time” measure was not a pure measure of the time a participant was actually attending to each object, as it would also include the time spent making navigational decisions. Accordingly, participants often stopped at intersections and looked around before deciding where to go next. More importantly, this alternative explanation cannot account for the lack of significant differences in spatial memory for decision-point and non-decision-point objects in Condition A (two starting points) in spite of equivalent differences in viewing times. Moreover, there was no correlation between viewing time and any of our spatial memory measures.

Although attentional differences between the two types of objects do not seem to be the most likely explanation for the superior memory for decision-point objects in Condition B, we cannot entirely rule out this possibility when participants are freely navigating in the environment and are free to re-visit any location as often as they like. Thus viewing times and experienced viewpoints of each object are not strictly controlled. Moreover, objects at decision points could be seen from three directions, whereas objects at non-decision points only could be seen from two directions, when participants were actively driving in the town. Thus, even when we eliminated one starting point in Condition B, the inherent difference in the number of experienced views for objects at L-shaped versus T-shaped intersections may have contributed to the superior memory for decision-point objects. Janzen et al. (2004) and Miller et al. (2010) controlled for potential factors such as viewing time and number of experienced viewpoints by having their participants passively transported through the virtual environment rather than actively navigating; in spite of equal viewing time for both types of objects, and only experiencing a single view of each object, they still saw evidence of encoding differences in both the fMRI and behavioural results.

To rule out the difference in number of experienced views or in viewing times as possible explanations of our decision-point effects in Condition B, we conducted a second experiment in which we showed participants videos of trajectories through a town instead of asking them to actively drive. As in Experiment 1, half the participants had one starting point and the other half had two.

2.4 Experiment 2

Participants watched videos showing a fixed route through the town. In Condition A participants saw the same route in both the forward and the reverse direction, while in Condition B they only saw the route in one direction, starting from a view facing Mike’s Restaurant and ending at a view of House of Pizza. As in Experiment 1, we interleaved blocks of study trials with blocks of memory test trials from two different tested viewpoints. Because participants’ trajectories through the town were highly constrained, relative to the free navigation conditions in Experiment 1, we were able to use much larger towns with more stores and objects while keeping the total study time to within a reasonable limit. Although active navigation might be more effective, we predicted that passively viewing a continuous trajectory through the town would still lead to the generation of a continuous cognitive map of the environment. Using a similar passive navigation paradigm and a recognition memory test, Janzen and van Turennout (2004) found greater parahippocampal activity for decision-point objects even when the participants did not correctly recognize them. Thus, as in the previous two experiments, we predicted that objects at decision points would more likely be encoded as part of an allocentric cognitive map, and should therefore be remembered more consistently across the two tested viewpoints, particularly in Condition B (only one starting point).

2.4.1 Method

2.4.1.1 Participants

Fifty McMaster University students ranging in age from 18 to 38 years (*mean* 20.44) participated in the experiment. There were 25 participants in each condition (8 males and 17 females in Condition A and 7 males and 18 females in Condition B). Participants had normal or corrected-to normal visions and received either partial course credit or $10 for taking part in the experiment. This study was reviewed and approved by the McMaster University Research Ethics Board. Written informed consent was obtained from all participants involved in this study.

2.4.1.2 Materials

As in the previous experiment, we used Yellow Cab to create a rectangular town (20 by 13 VR units in size), see Figure 9. Although the layout of the town was different than those used in Experiment 1, we imposed the same constraint on the locations of the decision-point and non-decision-point objects, namely, they were equally distributed about the town midline (see Figure 2.6), so that the average distance from each object to the town midline was equal for the two groups of objects. There were 20 objects, all of which were used in the pre-exposure phase and subsequent recognition memory test, and 10 of which were located in the virtual town during the study/navigation phase, five at decision points and five at non-decision points. Two video clips were created by recording the experimenter driving in the town following the route shown in Figure 9, in which each time an object was approached, there was a turn in the route; one going from Mike's Restaurant to House of Pizza, and the other traversing the reverse route from House of Pizza to Mike's Restaurant.

2.4.1.3 Procedure

There were six blocks of trials, each including an object pre-exposure phase, a study phase and a test phase, as in Experiment 1. In the study phase, participants in Condition A watched both videos alternatingly three times each, and participants in Condition B watched video 1 six times. Prior to each study phase, participants were shown the rectangular outline of the town with Mike's Restaurant and House of Pizza marked at each end, and told that they would be tested for their spatial memory of the objects after each block, and that they would be asked to draw a map of the layout of the town with all of the objects in it at the end of the experiment, to encourage participants to pay attention to the layout of the town during the study phase. The spatial memory test phase was the same as in Experiment 1.

2.4.1.4 Data Analysis

Recognition accuracy, pointing latency, pointing errors, and pointing consistency were calculated and analyzed as in Experiment 1, except that we had five decision-point and five non-decision-point objects in each town rather than 4 of each object type.

2.4.2 Results

2.4.2.1 Recognition accuracy.

Recognition accuracy was better for decision-point objects.

A two-way repeated measures Place x Condition ANOVA revealed a significant main effect of Place [*F* (1,48) = 4.917, *p* =0.031], but no main effect of Condition [*F* (1,48) = 0.902, *p* =0.347] and no interaction between Place and Condition [*F* (1,48) = 1.317, *p* =0.257]. As in Experiment 1, recognition memory for objects at decision points (*mean*=93.1%, *SE*=0.012) was significantly more accurate than for objects at non-decision points (*mean*=90.6%, *SE*=0.011) across conditions.

2.4.2.2 Pointing latency.

Pointing latencies were faster for decision-point objects.

A three-way repeated measures Place x Condition x tested Viewpoint ANOVA revealed a significant main effect of Place [*F* (1,48) = 5.673, *p* =0.021], but no main effects of tested Viewpoint [*F* (1,48) = 1.939, *p* =0.17] or Condition [*F* (1,48) = 0.244, *p* =0.623], and no interactions. Pointing latencies for objects at decision points (*mean*=3.465, *SE*=0.165) were significantly faster than those for objects at non-decision points (*mean*=3.643, *SE*=0.189) across conditions, although the effect was very small (*mean* difference of less than 0.2 seconds).

2.4.2.3 Pointing errors.

Pointing errors were affected by viewpoint, but not by navigational relevance.

A three-way repeated measures Place x Condition x tested Viewpoint ANOVA of the pointing errors revealed a significant interaction between tested Viewpoint and Condition [F (1,48) = 12.283, p =0.001], but no other main effects or interactions. Thus in contrast to the results obtained in Experiment 1 under active navigation conditions, navigational relevance did not significantly affect pointing errors when participants engaged in passive navigation. To identify the source of the viewpoint by condition interaction in terms of pointing errors, two separate two-tailed paired sample t-tests were conducted for each condition (for controlling for multiple comparison, significant p value was 0.025). There was no difference in the pointing errors between the two tested viewpoints in Condition A (t=2.116, df=24, p=0.045), but pointing errors made from the familiar Mike's Restaurant viewpoint were significantly smaller than those made from the House of Pizza viewpoint (t=-2.895, df=24, p=0.008). The results showed the viewpoint effect only in Condition B, but not in Condition A, see Table 2.1

2.4.2.4 Pointing consistency.

Navigational relevance affected pointing consistency when there was only one starting point.

A two-way repeated measures Place x Condition ANOVA of the pointing consistency standard deviation scores revealed a significant interaction between

Place and Condition [*F* (1,48) = 5.681, *p* =0.021], but no main effects of Place [*F* (1,48) = 0.081, *p* =0.777] or Condition [*F* (1,48) = 0.756, *p* =0.389], see Figure 2.7. To investigate the interaction between Place and Condition in terms of pointing consistency, two-tailed paired sample *t*-tests were used. In Condition B, pointing responses were significantly more consistent for decision-point objects than for non-decision-point objects (*t* =-2.484, *df* = 24, *p* = 0.020), but no such difference in Condition A (*t* =1.244, *df* = 24, *p* = 0.226). Note: for controlling for multiple comparisons, significant *p* value is 0.025. As in Experiment 1, reducing the number of starting points resulted in greater consistency of pointing errors across viewpoints for decision-point objects relative to non-decision-point objects, but there was no such difference when there were two starting points. There are two possible sources of the reduced variability in pointing errors to decision-point objects: the errors themselves could be smaller, and/or the errors could be more systematic across viewpoints. Our analysis of the pointing errors rules out the former interpretation, as there was no effect of navigational relevance on pointing error magnitude. Thus, the effect of navigational relevance on consistency, but not on accuracy, indicates that if an object was mis-localized when tested from one end of town, it tended to be mis-localized to the same (allocentric) direction when tested from the other end of town.

An alternative explanation for the consistency difference in Condition B could be alignment effects: some of our objects (3 decision-point objects and 2 non-decision-point objects) were viewed from directions aligned with the main longitudinal axis of the town and thus aligned with the tested viewpoints, whereas others (2 decision point and 3 non-decision-point objects) were viewed along the perpendicular axis. To rule out this alternative explanation, we performed the same analysis in Condition B on consistency scores for a subset of the objects, including two decision-point objects and two non-decision-point objects, which were pairwise matched for their average distances to the midline of the town, with one object of each type located on a part of the route aligned with the main longitudinal axis of the town and one object of each type located on a part of the route that was aligned with the perpendicular axis (two objects at far left of the town and two at the far right of the town, see EXPERIMENT 2 Materials section for details). Only trials where there were pointing errors for both objects (one decision-point object and one non-decision-point object) in each pair were used in each block, and then averaged by object types and over blocks. One participant's data were eliminated in this analysis due to unsuccessful recognition of all four objects over two blocks. We hypothesized a priori that even with this reduced set of responses to the matched pairs of objects, navigational relevance would still be a modulating factor, leading to greater pointing consistency for objects at decision points. A one-tailed paired sample *t*-test of the consistency scores revealed that, as with the full set of data, pointing responses for just these alignment-matched objects were significantly more consistent across viewpoints for decision-point objects (*mean*=23.847, *SE*=1.638) than for non-decision-point objects (*mean*=27.877, *SE*=2.492) (*t* =-1.958, *df* = 23, *p* =0.0315). Although the effect was weakened by only analyzing less than half (8 out of 20) of the responses, the navigational effect was still significant.

2.4.3 Discussion

The results of both Experiments 1 and 2 suggest that when participants have more limited experience with an environment (one starting point rather than two), objects at decision points are remembered more consistently, and are thus more likely to be encoded in a view-invariant manner. Janzen and van Turennout's [37] findings of greater fMRI parahippocampal activation during recognition memory judgments for objects placed at T-junctions relative to L-junctions suggest that different encoding mechanisms may be employed for these two types of objects. However, they did not explicitly test spatial memory. Building on their results, we saw a difference in the consistency of spatial memory errors, as hypothesized, with the responses for objects at non-decision points showing less consistency across tested viewpoints in spite of similar pointing error magnitudes for the two types of objects. Unlike in Experiment 1, the total viewing time and number of experienced viewpoints for the two types of objects were held constant in Experiments 2. The greater consistency of pointing errors for decision-point objects, in spite of a lack of difference in average absolute pointing errors for these objects, means that even when participants could not accurately recall the correct locations of the decision-point objects, they mis-localized these objects in a manner that was consistent across the two tested viewpoints, whereas pointing to non-decision-point objects was no less error-prone but less consistent across viewpoints. This finding provides strong support for the hypothesis that decision-point objects were more likely to be encoded within an allocentric frame of reference.

Across both experiments, whether participants navigated freely or passively, when they were biased to have fewer spatial reference directions (one starting point rather than two), pointing errors were less consistent for non-decision-point objects compared to decision-point objects. This was true even when participants only saw objects from a single view (Experiment 2, Condition B), suggesting that for objects that are highly relevant to navigation, even exposure to a single view may be sufficient for their incorporation into an allocentric representation, whereas for objects less relevant to navigation, exposure from multiple viewpoints may be required.

Our original hypothesis was that objects could either be 1) treated as landmarks and incorporated within allocentric maps of space, or 2) encoded egocentrically. The object’s relevance to navigation and spatial cognition, rather than the amount of attention paid to the object, was hypothesized to be a critical factor in determining whether the allocentric spatial memory system is engaged in object encoding. To further investigate this possibility, we designed another experiment in which we manipulated explicitly the type of attention participants paid to objects.

2.5 Experiment 3

We manipulated participants' attention explicitly by asking half of them to pay particular attention to the appearance and the other half to attend to the locations of objects. We hypothesized that when attending to appearance, participants would encode objects simply as objects, not as landmarks. In this case, navigational relevance would not contribute to memory encoding, and they would be primarily engaging their object recognition system (associated more with the ventral visual pathway) to process the objects. On the other hand, asking participants to pay attention to the locations of the objects was hypothesized to engage visuo-spatial attention and navigation circuits associated with the dorsal visual stream (and more specifically, with the parieto-frontal and parieto-temporal branches of the dorsal stream [1]) to a greater degree, leading to the incorporation of the object into a configural, allocentric representation of space in the medial temporal lobe. Thus, we predicted that when attending to objects' appearance participants' spatial memory would be equally accurate and consistent for decision and non-decision-point objects, whereas when attending to objects' locations, the greater navigational relevance of decision-point objects would favor their encoding as landmarks within an allocentric framework, relative to non-decision-point objects. Moreover, we tested whether video game experience would have an effect on spatial memory or navigational strategies.

We hypothesized that when attention was directed toward objects' appearance, the pointing consistency results we observed in the above experiments would disappear, and spatial memory would be less accurate for all objects, whereas when attention was directed toward objects' locations, the decision-point effect would be enhanced compared to results in Experiment 1-Condition B.

2.5.1 Method

2.5.1.1 Participants

Sixty McMaster University students participated in the experiment. Three participants whose recognition memory accuracy was less than 25% were excluded from the final data analysis. Therefore, there were fifty-seven participants; age ranged from 19 to 29 years, and the mean was 20.64. There were 29 participants (22 females and 7 males) in the Appearance condition and 28 (20 females and 8 males) in the Location condition. Participants received partial course credit or $10 for taking part in the experiment. This study was reviewed and approved by the McMaster Research Ethics Board. Written informed consent was obtained from all participants involved in this study.

2.5.1.2 Materials

The same materials were used as in Experiment 1.

2.5.1.3 Procedure

We used the same procedure as that used in Experiment 1 (active navigation) Condition B (single starting point) with the following changes.

At the beginning of the experiment, the participants were pseudo-randomly assigned to one of two attention conditions, either Appearance or Location, and were respectively asked in advance to pay particular attention to either the appearance or the locations of objects. They were told that their memory for the objects would be tested at the end of the experiment, and they would either have to recall as many visual details as possible of the objects in the appearance condition, or they would be asked to map out the locations of objects on a piece of paper in the location condition.

Whereas our previous experiments incorporated multiple blocks of interleaved study and test phases, in this experiment there was only one block of trials, including a single study phase and single test phase, in order to discourage participants from switching their attentional focus more towards the locations of objects after undergoing the first spatial memory test. The study phase was terminated once participants had found and delivered ten successive passengers or had reached the cutoff time of 35 minutes.

After the study phase and the pointing task, participants were asked: Do you play video games?

2.5.1.4 Data Analysis

Participants' recognition memory accuracy, pointing latency, pointing errors and pointing consistency across viewpoints were calculated as in the previous experiments. Additionally, we calculated the average navigation efficiency for each participant.

2.5.1.4.1 Navigation Efficiency

We subtracted the optimal time for each delivery based on the shortest route between the pick-up location and the destination from the actual time the participants took to deliver each passenger after the first 5 minutes navigating in the town. Hence, if the participant chose the shortest route to deliver the passenger, their efficiency score for this delivery would be zero. The first 5 minutes navigation was excluded from the analysis assuming participants used this time to learn the layout of the town.

2.5.2 Results

2.5.2.1 Recognition accuracy

Recognition accuracy was better for decision-point objects.

A two-way repeated measures ANOVA with Place (Decision vs. Non-Decision Points) as a within subject factor and Attention (Appearance vs. Location condition) as a between subject factor revealed a significant main effect of Place [*F* (1,55)=4.348, *p*=0.042], but no main effect of Attention [*F* (1,55)=0.094, *p*=0.76] and no interaction between Place and Attention [*F* (1,55)=0.001, *p*=0.981] on recognition memory accuracy. Recognition memory accuracy was better for objects at decision points than for those at non-decision points (Appearance: DP *mean*=86.69%, *SE*=0.03, NDP *mean*=81.17%, *SE*=0.029; Location: DP *mean*=87.64%; *SE*=0.031, NDP *mean*=82.25%, *SE*=0.029).

2.5.2.2 Pointing latency

Pointing latency was faster at the familiar viewpoint.

A three way repeated measures Place x Attention x tested Viewpoint ANOVA revealed a significant main effect of tested Viewpoint [*F* (1,55)=4.224, *p*=0.045] on pointing latency, but no other significant main effects or interactions. Pointing responses were faster when tested from the House of Pizza viewpoint (the starting point) than the Mike's Restaurant viewpoint for both types of object locations and both attention conditions (House of Pizza *mean*=4.804, *SE*=0.291; Mike's Restaurant *mean*=5.487, *SE*=0.476).

2.5.2.3 Pointing errors

Pointing errors were smaller at the familiar viewpoint across conditions.

A three-way repeated measures Place x Attention x tested Viewpoint ANOVA showed that there was a significant main effect of tested Viewpoint [*F* (1,55)=5.204, *p*=0.026] on pointing errors, but no other significant main effects or interactions. Pointing errors were smaller at the more familiar House of Pizza viewpoint than at the Mike's Restaurant viewpoint for both types of object locations and both attention conditions, see Table 2.1 for means and SEs.

2.5.2.4 Pointing consistency

Pointing responses were more consistent for objects at decision points in the Location condition, but not in the Appearance condition.

A two-way repeated measures Place x Attention ANOVA revealed a significant interaction between Place and Attention [*F* (1,55)=5.156, *p*=0.027], but no main effects of Place [*F* (1,55)=0.605, *p*=0.44] or Attention [*F* (1,55)=0.135, *p*=0.715] on pointing consistency. Based on the results of our previous experiments, we predicted a priori that pointing consistency would be worse for objects at non-decision points than for those at decision points in the

Location condition. To test this prediction, we therefore used a one-tailed paired sample *t*-test, which revealed that pointing scores were significantly more consistent for objects at decision points versus non-decision points in the Location condition (*t* = -2.186, *df* = 27, *p* =0.019, DP *mean*=23.457, *SE*=2.424; NDP *mean*=29.2486, *SE*=3.461), but not in the Appearance condition (*t* = 1.043, *df* = 28, *p* =0.153, DP *mean*=29.1428, *SE*=2.382; NDP *mean*=26.3072, *SE*=3.401), see Figure 2.8. (Note: for controlling for multiple comparisons, significant *p* value is 0.025).

2.5.2.5 Navigational efficiency

Pointing consistency was significantly correlated with navigational efficiency.

An analysis of the correlation between participants' navigation efficiency and pointing consistency revealed a positive correlation for both decision-point objects [*r* (28)=0.59, *p*=0.001] and non-decision-point objects [*r* (28)=0.51, *p*=0.006] in the Location condition, but no such correlation in the Appearance condition (DPs: [*r* (29)=0.284, *p*=0.136]; NDPs: [*r* (28)=-0.06, *p*=0.757]), see Figure 2.9. This result suggests that in the Location condition, objects were encoded as landmarks and facilitated efficient navigation, while in the Appearance condition they were not.

2.5.2.6 Questionnaire results

2.5.2.6.1 Video Game

Video Game players were more efficient at navigation, and more accurate but no more consistent in pointing.

There were 32 participants who self-identified as video game players (17 in the Appearance condition) and 25 who did not (12 in the Appearance condition). Separate two-way repeated measures Place x Video Game Experience ANOVAs were conducted on pointing latency, pointing errors and consistency. There was a significant main effect of video game experience on pointing errors [F (1,55)=6.581 p=0.013], but no other significant main effect or interaction with any other measure. Video game players had significantly smaller pointing errors (*mean*=27.602, *SE*=2.757) than non-players (*mean*=38.279; *SE*=3.119). However, video game experience was not a significant factor in pointing latency or pointing consistency.

We conjectured that video gamers might be more accurate at pointing to objects, even though they were no more consistent in the errors they made across viewpoints, because of their superior ability to navigate and encode routes, and subsequently to recall and/or imagine specific routes in the town. A two-tailed independent t-test to compare navigational efficiency of gamers to that of non-gamers revealed that participants who played video games (*mean*=16.2519) were also more efficient in delivering passengers to their destinations than non-players (*mean*=25.9189) (t = -2.125, df = 55, p =0.038).

2.5.3 Discussion

Much research has been devoted to the roles of the dorsal and ventral visual pathways, commonly referred to as the “what and where”, “what and how to”, or "perception and action" pathways (see e.g. Goodale & Milner, 1992; Kravitz, Saleem, Baker, Mishkin, 2011; Ungerleider & Mishkin, 1982). However, there have been relatively few attempts to manipulate the degree to which objects are processed by one pathway or the other within a single study. While we did not measure directly what neural circuits were involved, in Experiment 3 we manipulated the type of attention participants paid to objects, in an attempt to bias them in favor of either the object processing stream or the visuo-spatial stream. The results of Experiment 3 support our hypothesis that directing participants' attention to appearance vs. location affected how the objects were encoded. When asked to attend to the appearance of the objects, participants did not show any differences in pointing consistency between decision-point and non-decision-point objects. We suggest that this is because the objects were not treated as landmarks; therefore, navigational relevance would not contribute to memory encoding. On the other hand, asking participants to pay attention to the locations of the objects encouraged them to treat the objects as landmarks, not just simply as objects. This type of processing is postulated to engage the dorsal visual pathway, both the parieto-prefrontal branch for top-down executive control of visuospatial processing and the parieto-medial temporal branch for encoding within a world-centred reference frame (Kravitz, et al., 2011), leading to the incorporation of the object into a configural, allocentric representation of space in the medial temporal lobe. Consistent with this prediction, pointing responses were more view-invariant for decision-point objects than for non-decision-point objects when attention was directed to objects' locations, but not when attention was directed toward objects' appearances.

It is somewhat surprising that our attention manipulation did not produce any main effects on recognition memory, pointing latency or pointing errors. It could be that both attention conditions resulted in equally strong, but qualitatively different attentional resources being devoted to the objects in the two attention conditions. The differential effect of the attentional manipulation on pointing consistency supports this notion, but further experiments are required to demonstrate that distinctly different neural circuits were recruited in the two conditions.

Chun and Jiang (1998) suggested that memory for context could be implicitly learned and used to guide spatial attention for detecting the target among distractors. We suggest that without an explicit attentional manipulation, people might automatically pay attention to the locations of objects or building that are relevant for navigation in everyday life; this could explain the decision-point effect shown in our first two experiments. When attention was manipulated explicitly toward the objects' locations, this decision-point effect was enhanced, whereas when attention was focused on the objects' appearance, the effect was eliminated.

Both decision and non-decision points benefit from aligning the tested viewpoint with a salient reference direction, as the pointing responses were faster and more accurate when tested from House of Pizza viewpoint, the starting point, than from Mike's Restaurant. However, the analysis of pointing consistency revealed that spatial memory was less affected by viewpoint changes for decision-point objects than for non-decision-point objects.

An interesting double dissociation is apparent in the results of Experiment 3: video game experience was associated with faster and more accurate pointing responses and greater navigation efficiency but no greater pointing consistency. In contrast, the attentional manipulation affected pointing consistency but not pointing accuracy or latency. Further experiments would be required to tease apart what systems or strategies are at play that could explain these differences. One possibility is that gamers are more adept at employing egocentric route recall strategies, whereas attention to location versus appearance causes a (within-subjects) processing switch between spatial and non-spatial object encoding systems.

2.6 General Discussion

The results of our three experiments are consistent with our prediction that navigational relevance contributes to whether objects are encoded as landmarks within an allocentric framework. It is important to note that allocentric encoding might not be unique to the dorsal visual pathway; objects might individually be encoded in a view-invariant manner within the ventral visual pathway (the so-called "what pathway") (for more recent interpretations of the role of the ventral visual stream, see e.g. Goodale and Haffenden, 1998; Schenk, 2006). However, this type of allocentric or view-invariant coding of individual objects is distinctly different from the notion of allocentric spatial coding of a conjunction of the objects and features within a large-scale environment into a “cognitive map”, as typically attributed to the hippocampus. It is the latter type of allocentric encoding that we focus on in the present experiments.

Our novel pointing consistency measure proved to be sensitive to the navigational relevance manipulation, across all three experiments, in both active and passive navigation conditions. This greater viewpoint-invariance in memory for objects at decision points was modulated by whether participants began their navigation from both ends of town or just one (Condition A vs. B in Experiments 1 and 2), and whether participants attended to the objects’ locations or appearance (Experiment 3). When participants only began navigating from one end of the town, making one tested viewpoint more accessible than the other, spatial memory from the less familiar viewpoint was more disrupted for objects that were not at decision points. Similarly, when attention was explicitly directed toward the objects’ locations, memory for objects at decision points was even more consistent across tested viewpoints. Even when participants were no more accurate at pointing to decision-point objects (Experiments 2 and 3) they were still more consistent across tested viewpoints for objects at decision points relative to other objects. Taken together, our results suggest that when people process objects in the service of navigation, and when they are exposed to multiple views of objects, both factors contribute to the encoding of objects within their broader spatial context as allocentric spatial maps. These results are broadly consistent with the framework of the BBB model (Byrne, Becker, & Burgess; 2007), which proposes that a landmark's visual attributes are processed within the ventral visual stream, its spatial attributes are processed within the dorsal visual stream, and both are integrated into a large-scale spatial representation within the medial temporal lobe. Here we suggest a further refinement of the BBB theory, that is, objects in the environment may or may not be treated as landmarks, according to where they are located and how they are attended to.

If navigational relevance leads to more view-invariant location memory, does this necessarily imply allocentric coding? An alternative interpretation is that navigational relevance improves spatial encoding within an egocentric memory system. However, our results argued against this interpretation. If memory is allocentric, then reducing the number of viewpoints experienced at study should have little impact on the variability of pointing errors across tested viewpoints. As predicted, the manipulation of reducing the number of starting points in Experiments 1 and 2 only affected the consistency in pointing errors for non-decision-point objects, but not for decision-point objects. Thus, the results of our experiments are consistent with our prediction that objects at decision points were more likely than other objects to be encoded allocentrically. Moreover, it may only require exposure to a single viewpoint (Condition B in Experiment 2) to generate an allocentric representation of an object that is navigationally relevant.

While the effect of navigational relevance on viewpoint invariance was consistent across all three experiments, its effect on pointing accuracy versus latency differed. When there was a single starting point, pointing responses to objects at decision points were more accurate but no faster in Experiment 1 (active navigation), faster but no more accurate in Experiment 2 (passive navigation), and neither faster nor more accurate in Experiment 3. While participants in Experiment 3 only had a single block of study and test trials, those in Experiments 1 and 2 had multiple interleaved study-test blocks, affording the opportunity to develop different strategies over blocks on the pointing task in the active versus passive navigation conditions. There are several different strategies that could be used, including 1) employing an allocentric representation, 2) recalling multiple view-based snapshot memories and judging the alignment of the test view with the stored snapshots, 3) mentally rotating a single stored view of the scene to match the test view, or 4) imagining navigating from the tested view to the experienced view. Strategies 3 and 4 should result in longer reaction times relative to the strategies 1 and 2, while strategy 2 would be less accurate than an allocentric strategy, particularly when fewer stored viewpoints are available, as in the single starting point condition. In the passive navigation experiment in Condition B where only a single view of each object was seen, strategy 2 would be infeasible, but either strategy 3 (mental rotation of a stored view) or 4 (route recall / mental navigation) could have been employed. Adopting either of these egocentric strategies for non-decision-point objects and an allocentric strategy for decision-point objects would explain the observed reaction time differences. On the other hand, the active navigation conditions of

Experiment 1 permitted participants to approach each object from multiple directions and via multiple routes. This might bias participants to favor strategy 2, attempting to match the test viewpoint to multiple viewpoint-specific snapshot memories; such a strategy would be less accurate than an allocentric one, and could explain the lower accuracy for non-decision-point objects. Future research is required to determine which if any of the strategies discussed here might be employed, and under what conditions, to cause the observed differences between active and passive navigation.

Unexpectedly, video game experience was associated with greater pointing accuracy and greater navigational efficiency, but no greater pointing consistency. One reason for this pattern of results could be that the advantage conferred by video game experience in our task is due to better egocentric encoding and recall of routes rather than superior allocentric encoding of objects in their spatial context. While individual differences in encoding and retrieval strategies were not the main focus of the present experiments, there is a growing literature on spatial navigation supporting the notion that individuals do tend to favor either an allocentric strategy or an associative response strategy, each of which is associated with its own distinct neural circuits (see e.g. Bohbot et al, 2007; Doeller & Burgess, 2008). Moreover, preferential use of either the former or the latter strategy is associated with corresponding grey matter differences in the hippocampus versus basal ganglia (Bohbot et al., 2007). Future studies could probe in greater detail what strategy participants were employing in the tasks studied here, and how individual differences may contribute to when objects are incorporated into cognitive maps.

While the experiments reported here focused on how objects are encoded within large-scale spaces, other studies have identified additional factors at play in smaller spaces where the collection of objects can be viewed simultaneously. Mou, McNamara and colleagues proposed that the interobject relations form an intrinsic reference system that contributes to long-term spatial (Mou, Liu, & McNamara, 2009; Mou & McNamara, 2002; Mou, McNamara, Valiquette, & Rump, 2004; Shelton & McNamara, 2001; Valiquette & McNamara, 2007). When intrinsic structure (object defined) and extrinsic structure (environmental defined) are congruent, they jointly define the reference direction of spatial memory, while when they conflict, the first learning perspective (egocentric experience) defines the reference direction of spatial memory (Kelly & McNamara, 2008). Having a preferred reference direction for accessing spatial memory is not inconsistent with the use of an allocentric representation. As predicted by the BBB model, an access cue such as a view of a specific landmark arrives as an egocentric sensory input pattern, and must first be transformed into an allocentric representation, and then subjected to an associative recall process to retrieve a complete allocentric spatial memory. Consistent with the preferred reference direction findings of Mou and colleagues (Mou, Liu, & McNamara, 2009; Mou & McNamara, 2002; Mou, McNamara, Valiquette, & Rump, 2004; Shelton & McNamara, 2001; Valiquette & McNamara, 2007), we observed a viewpoint familiarity effect across all three experiments when participants started navigation from one end of town.

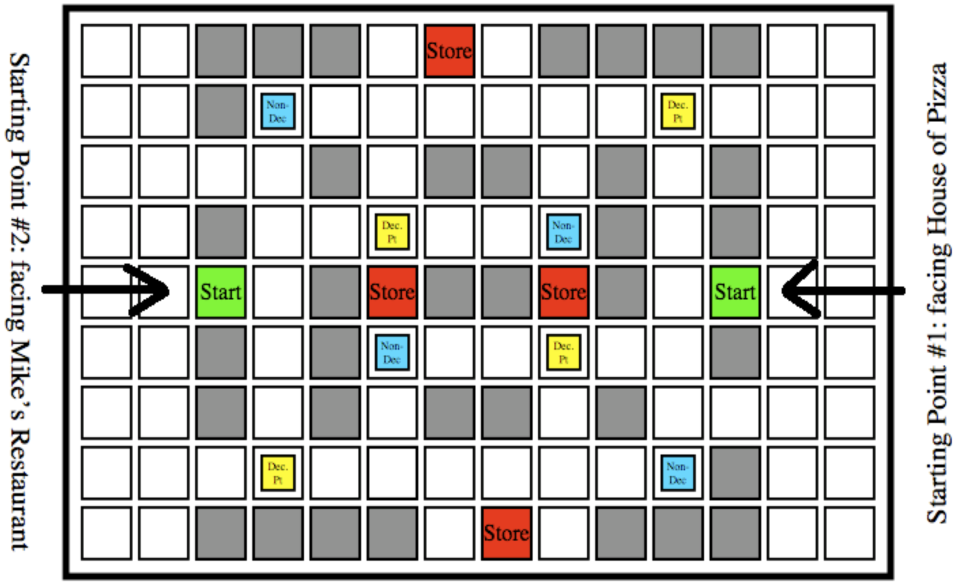
Our findings suggest that people may switch flexibly between egocentric and allocentric representations, according to the type of attention paid to objects. However, there are a number of limitations to the present set of experiments that warrant further study. First, we focused on within-subject encoding differences for objects at different locations, but we did not assess in detail potential between-subject strategy differences. The latter may be a product of both short-term context and long-term experience. For example, our results hinted at strategic differences between video gamers and non-gamers, and between participants engaged in active versus passive navigation. Future studies could investigate in greater detail the basis of such individual differences, with additional measures of spatial strategy use including questionnaires, secondary allocentric tasks such as navigation with detours or short-cuts, and fMRI to determine whether distinct neural circuits are recruited. Moreover, we focused on objects that are relatively small and contained within the confines of the larger space, but more profound encoding differences might be seen with larger distal cues. Finally, there is a growing literature on the encoding of large-scale spaces at multiple spatial scales. Individual objects might be encoded differently at multiple spatial scales when they are clustered in different regions of space, creating both inter-object relations at a local scale, and object-environment relations at a global scale.

Table 2-1. Average pointing errors in degrees for Experiments 1, 2 and 3.

DPs represents decision points; NDPs represents non-decision points; numbers in the bracket represent standard errors; HoP represents that House of Pizza is the starting point and Mike represents that Mike's Restaurant is the starting point.

Tested Viewpoint	Place	Expt. 1-A Two starting points	Expt. 1-B One starting point (HoP)	Expt. 2-A Two ways	Expt. 2-B One way (Mike)	Expt. 3-Appearance (HoP)	Expt. 3-Location (HoP)
House of Pizza	DPs	25.41 (1.80)	25.05 (1.81)	30.11 (1.84)	31.93 (2.5)	31.85 (4.00)	32.79 (4.07)
	NDPs	25.78 (1.92)	27.94 (2.21)	30.42 (1.63)	32.06 (2.38)	28.03 (3.83)	29.28 (3.90)
Mike's Restaurant	DPs	24.64 (1.84)	30.82 (2.15)	31.81 (2.07)	28.3 (2.32)	33.48 (3.69)	33.79 (3.75)
	NDPs	24.48 (2.22)	34.76 (2.56)	32.59 (1.92)	29.78 (2.26)	31.30 (3.70)	39.20 (3.77)

Figure 2.1 Virtual town used in Experiments 1 and 3.



Town's layout (14 by 9 VR units in size) used in Experiments 1 and 3. The grey squares are non-distinctive uniformly textured buildings at locations where the participants are not able to drive into. The "Store" squares are the stores that serve as passenger drop-off locations. The two "Start" squares are the two starting points, locating at either end of the town (Mike's Restaurant and House of Pizza). The "Non-Dec" squares are the non-decision points where the objects were placed; at these locations participants can only turn in one direction. The "Dec. Pt." squares are the decision points where the objects were placed; at these locations participants can either turn left or right. All the white squares indicate locations along the routes that participants can navigate in the town. All the objects are located in the middle of the street; participants can go around them.

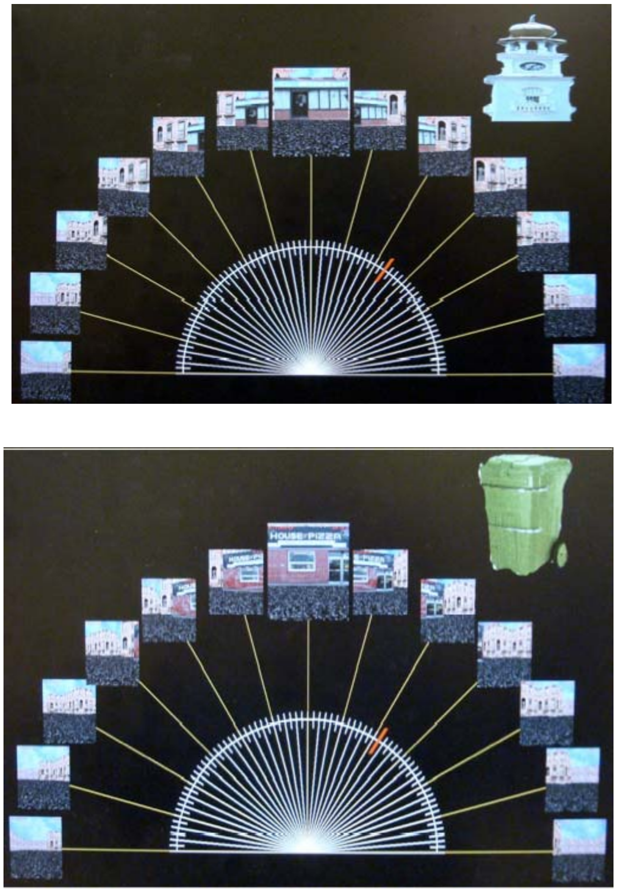
Figure 2.2 Two starting points (also tested viewpoints).

(a) Mike's Restaurant

(b) House of Pizza

Figure 2.3 Sample objects used in the towns

Figure 2.4 Pointing task: two testing viewpoints

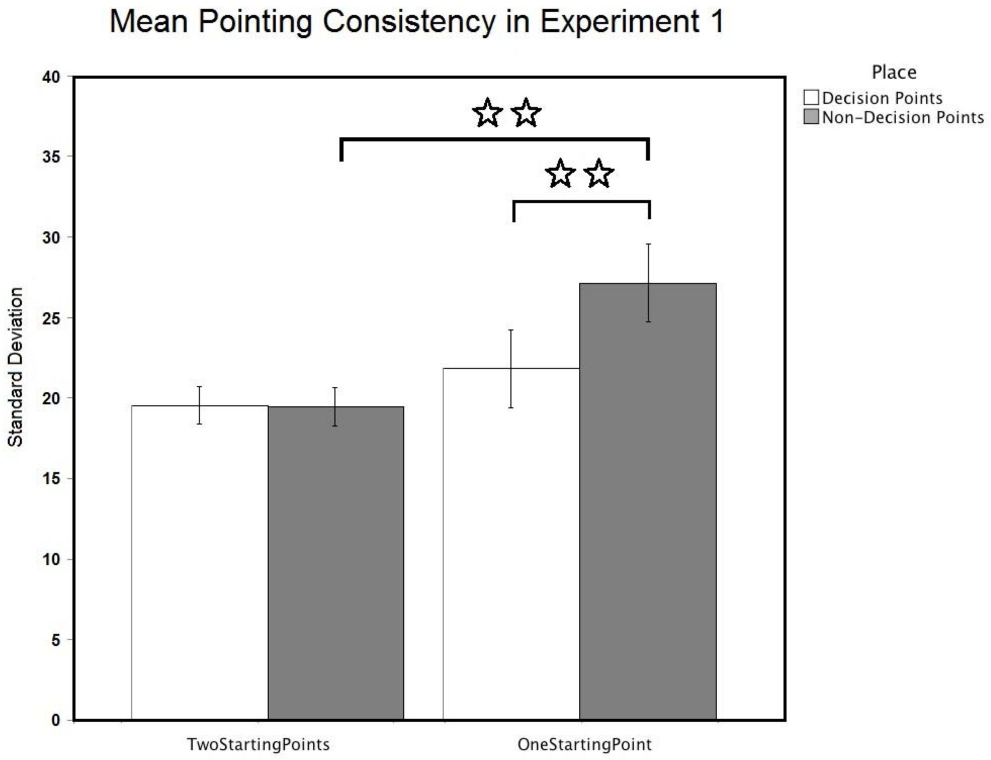


(a)

(b)

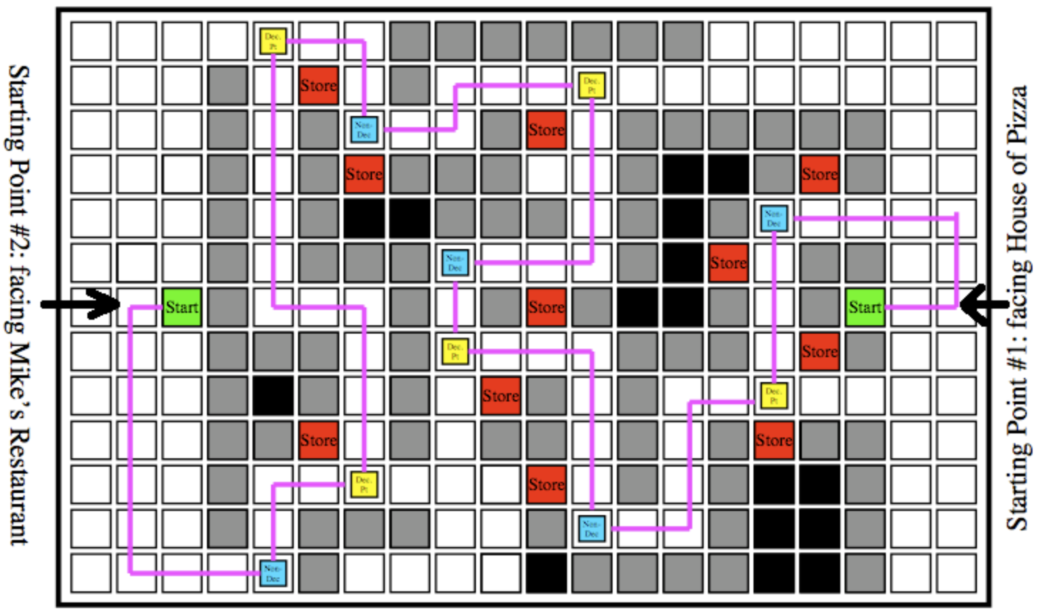
The navigators used in the pointing task. In a semi-circular arc along the top of the navigator, pictures of different views of the actual town are shown, as seen from different angles at the starting location at either one end or the other end of the town. At the tip of the compass pointer (red) an image was shown of the target object for the current trial. It could be moved by moving the pointer. (a). Navigator from Mike's Restaurant point of view; (b). Navigator from House of Pizza point of view.

Figure 2.5 Pointing consistency in Experiment 1.



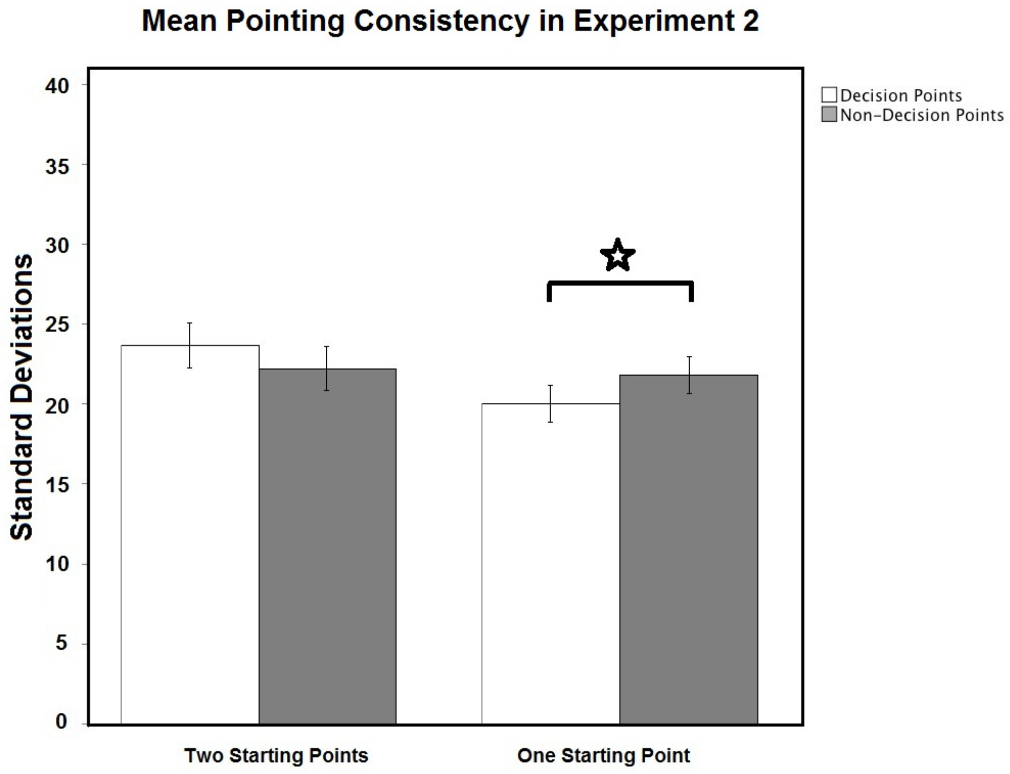
Active Navigation: Mean pointing consistency scores (95% confidence intervals) for decision and non-decision-point objects in Experiments 1-Active navigation (two starting points vs. one starting point). White bar is for decision points and grey bar is for non-decision points. The pointing responses were significantly more consistent for objects at decision points than for those at non-decision points in Experiment 1b (one starting point), but not in Experiment 1a (two starting points). ☆☆ means $p<0.01$.

Figure 2.6 Virtual town used in Experiment 2.



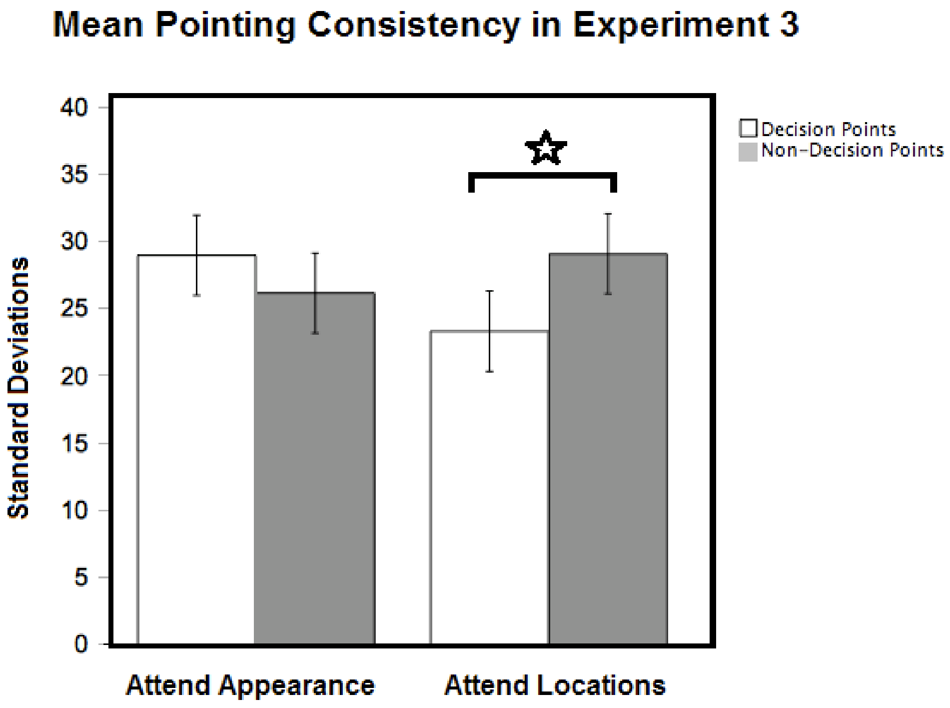
Town's layout (20 by 13 VR units in size) used in Experiments 2, similar to the one used in Experiment 1, but larger. The pink line is the travel trajectory that participants watched in the video clips.

Figure 2.7 Pointing consistency in Experiment 2.



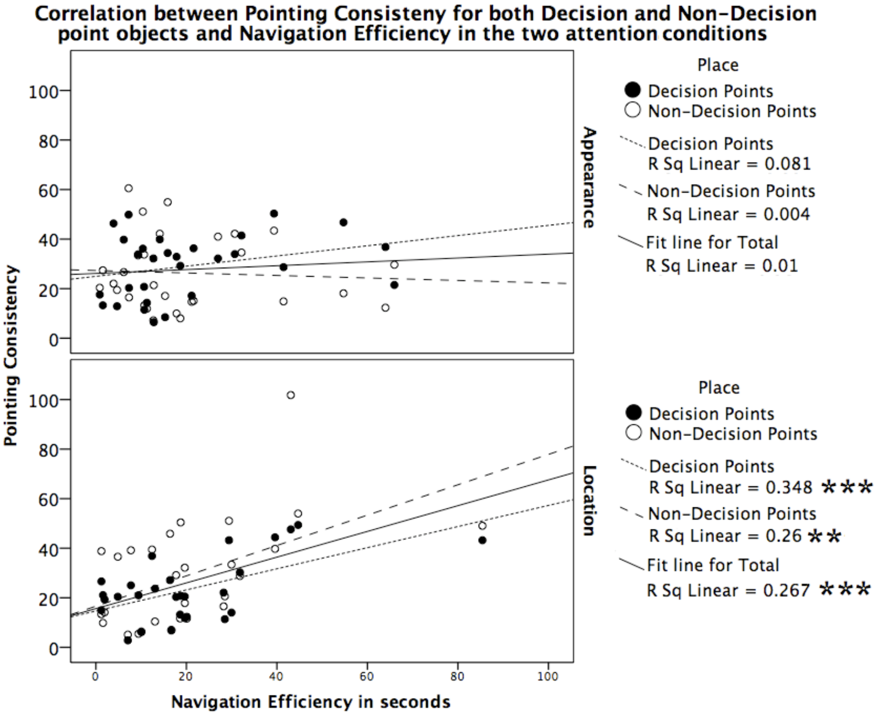
Passive Navigation: Mean pointing consistency scores (95% confidence intervals) for decision and non-decision-point objects in Experiments 2-Passive navigation (two starting points vs. one starting point). White bar is for decision points and grey bar is for non-decision points. Pointing responses were more consistent for objects at decision points than for those at non-decision points in Experiment 2b (one starting point), but not in Experiment 2a (two starting points). ☆ means $p<0.05$.

Figure 2.8 Pointing consistency in Experiment 3.



Attention Manipulation: Mean pointing consistency scores (95% confidence intervals) for decision and non-decision-point objects in Experiments 3- Active navigation with attention manipulation (attend Appearance vs. attend Locations). White bar is for decision points and grey bar is for non-decision points. Pointing responses were more consistent for objects at decision points than for those at non-decision points in the Location condition, but not in the Appearance condition. ☆ means $p<0.05$.

Figure 2.9 Correlation between navigation efficiency and pointing consistency in Experiment 3.



Correlation between navigation efficiency and pointing consistency for decision and non-decision-point objects for both attention conditions in Experiments 3. Black dot is for decision points and white dot is for non-decision points. There are significant positive correlations between navigation efficiency and pointing consistency for decision-points and non-decision-point objects in the Location condition, but not in the Appearance condition. The more consistent the pointing responses made from two tested viewpoints, the more efficient the participants were in delivering passengers. ** means p<0.01, *** means p<0.001.

3 **Chapter Three**

One spatial map or many? A VR study of how do we encode multiple connected environments.

3.1 Foreword

The results from Chapter 2 revealed the importance of the locations of objects within an environment, and the type of attention devoted to those objects, for constructing cognitive maps. This sheds light on what sort of features may provide input to the process of cognitive map formation. However, many unanswered questions remain regarding the process of cognitive map formation. For example, in large-scale environments, do we construct a single cognitive map or multiple cognitive maps of different sub-regions, is the representation flat or hierarchical, how many of these maps would be formed for multiple spaces and how are larger regions of space are integrated. A previous study suggested that subjects were able to point to landmarks across multiple levels of the building, but there was a cost in pointing latency and accuracy when they were required to integrate across multiple routes (Montello & Pick, 1993), suggestive of multiple local representations. Subjects might either form a hierarchy of spatial maps at multiple scales or integrate spatial information across multiple locally-constructed maps in this study. However, the study by Montello and Pick required subjects to integrate spatial information across three dimensions (i.e., across multiple floors of a building), which may have posed an unnatural barrier to map integration. It could be that within a single two-dimensional plane, people could more readily integrate their spatial knowledge from local regions into a global perspective.

In our study, we investigated how humans encode multiple connected spatial environments using a virtual taxi game. We hypothesized that if two connected neighborhoods are explored jointly, people will form a single integrated spatial representation of the town. If the neighborhoods are first learned separately and later observed to be connected, then people will form separate spatial representations and there will be an accuracy cost when inferring directions from one neighborhood to the other. Interestingly, our data were inconsistent with these hypotheses, and instead suggest that people have a very strong bias to form separate local maps, regardless of whether the neighborhoods were learned together or separately. Only when all visible distinctions between neighborhoods were removed did people behave as if they formed one integrated spatial representation. These data are broadly consistent with evidence from rodent hippocampal place cell recordings in connected boxes, and with hierarchical models of spatial coding.

3.2 Introduction

From the pioneering studies of Tolman (1948), it has been widely accepted that mammals including rodents form cognitive maps. The construction of cognitive maps is crucial to many of our spatial abilities, including our ability to navigate from one place to another. And yet, after many decades of research on this topic, the precise nature of our internal spatial representations remains a topic of considerable debate. One of many unresolved issues is whether we construct a single cognitive map of a large-scale space, or many local maps of smaller regions.

Some theories of spatial cognition have emphasized local representations. For example, Worden (1992) proposed that mammals store memories of their geographical environment as a collection of independent fragments. Each fragment consists of a set of landmarks with their geometric relationship and non-geometric properties. For example, subjects had longer response times when successive trials probed different environments, suggesting the representations of these environments were independent, sequentially accessed fragments (Brockmole & Wang, 2002).. Others have suggested that peoples’ spatial memories are organized hierarchically based on global and local properties of an environment (e.g. Hirtle & Jonides, 1985; McNamara, 1986; Meilinger, 2008; Stevens & Coupe, 1978). In support of this notion, people are more accurate in estimating the distance for within-cluster landmarks, defined as landmarks that were recalled together during a recall task, than for between-cluster landmarks

(Hirtle & Jonides, 1985), more readily primed by locations in the same region than locations in different regions (McNamara, 1986), and make more errors in judging relations between landmarks located at widely separated geographical locations (Stevens & Coupe, 1978). Conversely, the between-array geometry may influence the performance on within-array judgments (Greenauer & Waller, 2010).

Another factor that may be important in determining whether an environment is represented as one global map or separate local maps is whether the local regions were learned together or separately. Few studies have investigated this question. Montello and Pick (1993) found that people who learned two routes on different levels of a building were able to point to landmarks across multiple levels of the building, but there was a cost in pointing latency and accuracy when they were required to integrate across multiple routes. This result is consistent with the idea that subjects based their responses on a hierarchy of spatial maps at multiple scales. Alternatively, subjects may have employed post-retrieval strategies to integrate spatial information across multiple locally-constructed maps. The study by Montello and Pick required subjects to integrate spatial information across three dimensions (i.e., across multiple floors of a building), which may have posed an unnatural barrier to map integration. It could be that within a single two-dimensional plane, people could more readily integrate their spatial knowledge from local regions into a global perspective. However, Ishikawa and Montello (2006) showed that people were able to integrate two routes (i.e. estimate directions between S-shaped route and U-shaped route) on a horizontal plane, but the subjects were less accurate in estimating landmarks on one of the two routes from the other route, which suggests that they might have constructed multiple spatial maps or a hierarchy of spatial maps at multiple scales even within a single two-dimensional plane.

Strong evidence for local spatial representations comes from electrophysiological recordings of place cells. Place cells, first discovered by O'Keefe and Dostrovsky (1971) and subsequently reported in many species including humans (Ekstrom et al., 2003), fire when the animal is within a specific local area and are often insensitive to heading direction (O'Keefe, 1976). A place cell's spatial tuning is based on inputs from a multitude of cues, including the location of local boundaries from subicular "boundary vector cells"(Burgess, Jackson, Hartley, & O'Keefe, 2000; Hartley, Burgess, Lever, Cacucci, & O'Keefe, 2000; Lever, Burton, Jeewajee, O'Keefe, & Burgess, 2009), path integration cues from entorhinal grid cells (Hafting, Fyhn, Molden, Moser, & Moser, 2005), and contextual cues such as odour (Anderson & Jeffery, 2003). Moreover, recordings made in very large spaces suggest that place cells form the basis of a hierarchical or multi-scale representation, with some place fields spanning the entire length of an 18-meter track (Kjelstrup et al., 2008). Further support for multiscale spatial representations comes from the properties of grid cells in the medial entorhinal cortex, each of which fires at multiple locations arranged in a hexagonal grid (e.g. Fyhn, Molden, Witter, Moser, & Moser, 2004;

Hafting et al., 2005; Sargolini et al., 2006). Like place cell firing fields, grid cell firing fields vary in spatial scale (Barry, Hayman, Burgess, & Jeffery, 2007).

Given the above evidence for local multiscale representations of space, a key question is how larger regions of space are integrated, within very large-scale complex environments such as cities. One possibility is that local spatial representations are flexibly combined, during the learning process, into larger scale, more complex representations. If this is the case, then one would expect to see local representations of connecting paths between regions, and place cells that fire in one region or the other, but not both. Alternatively, separate maps might be formed for different local regions, and flexibly combined via post-retrieval processes. In the latter case, one would expect completely distinct sets of place cells firing in different sub-regions, and place cells with multiple unrelated firing fields in different regions. A wealth of electrophysiological evidence sheds light on this question. For example, in complex spatial environments with multiple turning points both place cells and grid cells recorded in rats in the hairpin maze showed similar remapping patterns at the turning points (Derdikman et al., 2009), which suggests that regions separated by a turning point may be encoded as separate distinct maps; moreover, turning points may be encoded separately, providing a representational bridge between the different regions. Other evidence seems to support piecemeal, fragmented maps of space (Derdikman et al., 2009). For example, when two regions of an environment are learned separately, the place cells in the two environments bear no relation to each other and many cells have (unrelated) place fields in both regions (Tanila, 1999). Furthermore, opening a connection between the two environments causes many of the place cells to re-map and/or develop a single place field in just one part of the environment (Paz_Villagran, Save, & Poucet, 2004), suggesting that the animal is treating the unified space as a new environment and generating distinctive internal representation for the latter case.

Based on the above evidence from human behavioural and animal electrophysiological studies, we predict that when there is local spatial structure in a large-scale environment, humans will tend to construct multiple local maps, particularly when the multiple environments are explored separately. Therefore, in our study, we had participants explore two connected virtual neighborhoods and investigated under what conditions they would be treated as one large unitary spatial map versus multiple piece-meal maps. In the latter case, combining two maps together should incur a cost in accuracy and/or speed when inferring directions from a location in one map to a location in the other map.

In Experiment 1, participants learned two neighborhoods separately by playing a virtual taxi game. They were then shown a video clip of a connecting pathway the two neighborhoods, and after that, they explored the neighborhoods jointly. Spatial memory was tested from two different viewpoints, by having participants point to different passenger drop-off (PDO) locations, which, given the goals of the virtual taxi game, were expected to be well-learned, salient landmarks. Because the results were most consistent with participants forming separate maps, in Experiment 2 we varied the means by which participants learned how the two neighborhoods were connected by either allowing them to i) view, but not navigate, the connection between the neighborhoods; ii) view a video clip of being teleported along the connection; or iii) freely navigate along the connection. As in Experiment 1, within-neighborhoods pointing to PDO locations was always more accurate than between-neighborhoods pointing, which suggests that participants based their responses on multiple local maps. Experiment 3 examined the hypothesis that participants' between-neighborhood errors were due to an inability to accurately judge the length of the connection between neighborhoods by removing the fences that separated the two neighborhoods. Again participants were most accurate for within-neighborhood PDO locations. Finally, in Experiment 4 we removed all distinct features that distinguished the town's two neighborhoods. In this final experiment, the direction estimation was not different between the two types of PDO locations. Thus, participants were able to encode the large town as one single environment when there were no differentiating cues to spatially group it into local regions; in all other cases, their behaviour was more consistent with the construction of multiple local maps.

3.3 Experiment 1

Participants implicitly learned the town layout and PDO locations by playing a virtual taxi game requiring active navigation through a virtual town. Learning and spatial memory test phases were repeatedly interleaved in 5 blocks.

3.3.1 Method

3.3.1.1 Participants

Twenty-six McMaster University students (7 males and 19 females) of ages ranging from 18 to 32 years (mean age 19.31) participated in the experiment. Participants had normal or corrected-to-normal vision and received partial course credit for taking part in this 1-hour experiment. Written informed consent was obtained from all participants. This study was reviewed and approved by the McMaster Research Ethics Board.

3.3.1.2 Materials

We employed Kahana's "Yellow Cab" virtual driving simulator (see http://memory.psych.upenn.edu/Research) for constructing the environment and for simulating the virtual taxi game for the study phase of the experiment. There was one rectangular shaped town (21 by 10 VR units in size) consisting of two neighborhoods (see Figure 3.1) connected by a navigable pathway that was initially occluded by an opaque, non-navigable barrier. Each of the two neighborhood included eight distinctly textured and signed passenger drops off (PDO) locations as well as multiple uniformly textured grey background buildings. One of the neighborhoods, which we shall refer to as neighborhood A (left side of the town, colored in purple in Figure 3.1), was designed as the restaurants district , and the other, which we shall refer to as neighborhood B (right side of the town, colored in blue in Figure 3.1), was designed as a shopping district. Each neighborhood was surrounded by distinctly coloured and textured fences. In addition to the 8 PDOs, each neighborhood had another distinctive building -- “Mike’s Restaurant” (see Figure 3.2a) in neighborhood A and “Aaron Chang Gallery” (see Figure 3.2b) in neighborhood B -- which marked the two respective starting points for passenger pickups in the two neighborhoods, and also the two tested viewpoints for the spatial memory tests; these starting points did not serve as PDOs. There were 30 restaurants/shops used in the experiment, eighteen of which appeared in the town and the remaining twelve of which were only shown at the restaurant/shop pre-exposure phase and served as distractors for the subsequent memory task. The locations of six of the eight restaurants (numbered in Figure 3.1) and the location of the starting point in neighborhood A were matched pairwise with respect to their distance from the midline of the town to shops (numbered in Figure 3.1) and starting point in neighborhood B: with the viewer at location (0,0), the center of the town, for every restaurant at location (x, y) in neighborhood A there was a corresponding shop in neighborhood B at location (-x, y) (see Figure 3.1). Thus, during the pointing task with the participant placed at starting point A (or B), the distance and angles of restaurants 1a, 2a, and 3a (or shops 4b, 5b and 6b) in neighborhood A (or B) to the observer were equal to those of the pairwise matched shops 1b, 2b, and 3b (or restaurants 4a, 5a, and 6a) in neighborhood B (or A). The identities and locations of the restaurants and shops in the town remained constant across blocks. During the study phase, participants used the arrow keys on the keyboard to control their navigation, allowing them to turn in any direction, control their speed, or do a U-turn.

The memory test was implemented in Matlab with the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). On each memory test trial, there was a half compass shaped figure (a navigator) with a picture of one of the two starting point views on the bottom of the navigator (see Figure 3.3a and 3.3b). Additionally, at the tip of the compass pointer an image of the target restaurant or shop for the current trial appeared; this target moved with the pointer. The target was always one of the 30 restaurants/shops (18 of which had been in the town and 12 of which were distractors), and the participant could use the mouse to rotate the compass pointer (along with the target) to point in the remembered direction of the target object from the displayed viewpoint and make a mouse-click to indicate a pointing response in the remembered direction of the target, or press the space bar to indicate that a restaurant or a shop was not recognized as having been in the town.

3.3.1.3 Procedure

There was a total of five blocks. Blocks 1, 2, 4, and 5 each included a restaurant/shop pre-exposure phase, a study phase and a test phase. Block 3 consisted of a restaurant/shop pre-exposure phase, a video clip viewing and a test phase. In the restaurant/shop pre-exposure phase, each of the 30 restaurants or shops appeared for two seconds followed by a blank screen for one second. The purpose of this pre-exposure phase was to establish a degree of familiarity with the distractors so that the subsequent recognition memory task would be more challenging. In the study phase, the participant was asked to act as a taxi driver whose task was to roam through the neighborhood to find passengers, which were randomly located in the town, and deliver them to specific locations (restaurants or shops, a.k.a PDO locations). A trial began with the participant located at one starting point, facing toward the town, and he/she was asked to freely navigate until a passenger was found and "collected" by bumping into the passenger. A textual cue then appeared, e.g. "Please take me to the Computer Store, I will give you 100 points", and the participant was required to navigate as quickly and efficiently as possible to drop off the passenger to the goal location by bumping into it.

In Block 1, participants explored only half of the town (e.g. neighborhood A), while in Block 2 participants explored the other half of the town (e.g. neighborhood B). There was no visible or accessible connection between the two neighborhoods. Which neighborhood was explored first was counterbalanced between participants. In Block 3, participants were taken through the restaurant/shop pre-exposure phase and were then shown a video clip that illustrated how the two neighborhoods were connected by presenting a trajectory of driving from one starting point to the other starting point. The direction of the trajectory was from starting point A to B for half the participants and in the reverse direction for the other half. In Blocks 4 and 5, the visible barrier between the two neighborhoods was removed and the connection was open and navigable, and therefore participants were able to travel back and forth between the two neighborhoods and explore the entire space of the town in locating and dropping off passengers.

In each of the active navigation blocks (1, 2, 4 and 5), the study phase was terminated when either participants had successfully found and delivered five passengers or 10 minutes had elapsed. In Blocks 1 and 2, after each passenger delivery, the participant was relocated to the same starting point within the neighborhood to start another passenger pickup while in Blocks 4 and 5 (after removal of the barrier) he or she was relocated to the other starting point in the adjacent neighborhood alternatingly between pickups. The participant’s location and viewing direction were recorded every 30-40 ms throughout the entire study phase. The memory test combined simultaneous tests of recognition memory and spatial memory.

In Blocks 1 and 2, immediately after the study phase, the participants' spatial memory for the restaurants or shops in the neighborhood that they explored was tested by having them point to the remembered locations of the buildings from the starting point in the same neighborhood. In Blocks 3, 4 and 5, spatial memory for the restaurants and shops in each of the two neighborhoods was tested by having participants point to the remembered locations of each of the restaurants and shops in the town twice, once from each of the two tested viewpoints/starting points (see Figure 3.3a and 3.3b). On each test trial, if the participant thought the restaurant or shop had not appeared in the town, he or she pressed the ‘space bar’, and the next restaurant or shop would be displayed. Otherwise, he or she then pointed to the direction of the remembered location of the restaurant or shop from the displayed viewpoint by using the mouse to move the compass pointer to the desired direction, and then pressing the left mouse button[1]. The recognition responses, pointing directions and total reaction time for the combined spatial/recognition memory response were recorded during the memory test phase. After 5 blocks of study and test phases, participants were asked to draw a map of the virtual town on a piece of paper (i.e. the mapping task).

3.3.1.4 Data Analysis

The mapping task performed at the end of the experiment required participants to draw a map of the town consisting of two neighborhoods with restaurants and shops. This map allowed us to determine if a participant understood the correct relationship between the two starting points, which means they knew the correct spatial relations of the two neighborhoods. Only data from participants who correctly mapped the two starting points in the main statistical analyses.

---

[1] Note: We did not measure recognition memory reaction time separately, but we did measure pointing latency, as our memory test combined recognition and spatial memory.

Within-neighborhood pointing responses were defined as responses made from starting point A to restaurants 1a, 2a, and 3a in neighborhood A, and from starting point B to shops 4b, 5b, and 6b in neighborhood B. Between-neighborhood pointing responses were defined as responses from starting point A to shops 1b, 2b, and 3b in neighborhood B, and from starting point B to restaurants 4a, 5a, and 6a in neighborhood A.

Recognition accuracy, pointing latency, and pointing errors from Blocks 1 and 2 for the 6 matched restaurants and shops in each neighborhood also were analyzed to determine if task difficulty differed between the two neighborhoods.

3.3.1.4.1 Recognition Accuracy

A recognition response was correct if the participant indicated that they had seen the restaurant or shop that was used in the town; otherwise the response was incorrect. We calculated percent correct recognition separately for PDO locations within the neighborhood and those between neighborhoods. A two-tailed paired sample t-test was used to compare the difference in recognition memory accuracies between two types of PDO locations (within versus between neighborhoods).

3.3.1.4.2 Pointing Latency

There were two reaction times (in seconds) for each PDO location in each block: one for each of the two tested viewpoints. We averaged reaction times for correct responses across blocks for pointing to within- and between-neighborhood PDO locations. Reaction times were analyzed with a two-tailed paired sample t-test.

3.3.1.4.3 Pointing Error (Average Absolute Pointing Errors)

In each block, participants had to point to each PDO location once from each of the two starting points. Therefore, each participant had two pointing responses for each PDO location, one from the viewpoint of “Mike’s Restaurant” and one from the viewpoint of “Aaron Chang Gallery”. A pointing error was defined as the signed value, in degrees, of the difference between the pointing response and the PDO location’s actual direction (participant’s response in degrees minus the PDO location’s correct direction in degrees). Therefore, for each starting location we had six signed pointing errors for within- and between-neighborhood PDO locations for each participant in each block (assuming that participants correctly recognized all PDO locations) and then averaged across blocks. We calculated the absolute value of all the raw signed pointing errors for each participant. A two-tailed paired sample t-test was used to compare the difference in pointing errors between two types of PDO locations (within versus between neighborhoods).

3.3.2 Results

Based on the mapping task at the end of the experiment, six participants (5 females and 1 male) failed to understand the spatial relationship between the two starting points; their data were excluded from further analyses. Therefore, we only analyzed data of twenty participants (6 males and 14 females).

There was no difference between the PDO locations in the two neighborhoods in terms of recognition accuracy (t=1.633, df=19, p=0.119), pointing latency (t=0.554, df=19, p=0.586) or pointing errors (t=0.065, df=19, p=0.949), which suggests that the PDO buildings in the two neighborhoods did not differ in terms of encoding difficulty.

3.3.2.1 Recognition accuracy

A two-tailed paired sample t-test of the recognition accuracy revealed no significant difference between the two types of PDO locations (t=1.65, df=19, p=0.116; Within neighborhoods PDO locations: mean=95.56%, SE=0.012; Between neighborhoods PDO locations: mean=92.78%, SE=0.017).

3.3.2.2 Pointing latency

A two-tailed paired sample t-test of the pointing latencies revealed a significant difference between the two types of PDO locations (t=-6.61, df=19, p<0.001). Responses were faster for pointing to PDO locations within neighborhoods (mean=2.87, SE=0.24) than between neighborhoods (mean=3.76, SE=0.29).

3.3.2.3 Pointing errors

A two-tailed paired sample t-test of the pointing errors revealed a significant difference between the two types of PDO locations (t=-4.18, df=19, p=0.001). Errors were significantly smaller when pointing to PDO locations within neighborhoods (mean=24.49, SE=1.72) than between neighborhoods (mean=37.74, SE=3.57) (see Figure 3.4).

3.3.3 Discussion

The results are consistent with the construction of separate local spatial maps for the two neighborhoods. We hypothesized that combining two local maps would result in reduced accuracy and/or an increase latency of between-neighborhood pointing responses, both of which were seen in our results: pointing responses for PDO locations in the same neighborhoods as the testing viewpoints were more accurate and faster than for those in the adjacent neighborhoods, even when we only analyzed PDO locations that were on average, equi-distant from the observer (i.e. each restaurant or shop within the same neighborhood was paired with a shop or restaurant at equal distance away from the observer in the other neighborhood). The longer time they took did not improve their accuracies for PDO locations in the adjacent neighborhoods; on the contrary, pointing errors were larger for those PDO locations. Moreover, there was a difference between the two types of PDO locations in terms of pointing errors even in the last block (see Figure 3.4).

One reason the two neighborhoods were treated as separate environments could be because they were explored separately at the beginning of the experiment. To investigate this possibility, in Experiment 2 we added a condition in which the two neighborhoods were explored together right from the start of the experiment. We also allowed participants to explore the environments for a greater number of blocks, and varied the conditions under which participants learned how the two neighborhoods were connected.

3.4 Experiment 2

3.4.1 Method

3.4.1.1 Participants

Sixty McMaster University students participated in the experiment; age ranged from 18 to 37 years, and the mean was 20.02. There were 20 participants in each condition (10 males and 10 females). Participants had normal or corrected-to-normal vision. Participants received two course credits or $20 for taking part in this 2-hour experiment. This study was reviewed and approved by the McMaster Research Ethics Board. Written informed consent was obtained from all participants.

3.4.1.2 Materials

The materials were the same as those used in Experiment 1. In addition, we had participants complete a questionnaire adapted from the Santa Barbara Sense of Direction Scale (Hegarty, Richardson, Montello, Lovelace, & Subbiah, 2002) to investigate what factors may be associated with performance on the pointing task (see Appendix A).

3.4.1.3 Procedure

The experiment had six blocks, each consisting of a restaurant/shop pre-exposure, a study and a test phase. The restaurant/shop pre-exposure phase was the same as in Experiment 1. Each study phase used procedures that were similar to those used in Experiment 1, except we had three conditions, which varied between subjects. In the *View* condition, in the first half of the study phase in each block, participants explored one neighborhood (e.g. neighborhood A) but could see the other neighborhood (e.g. neighborhood B) through a pathway connection, but they could not drive through it. In the second half of the study phase in each block, participants explored the other neighborhood (e.g. neighborhood B), and again could see the first neighborhood along the connecting path but could not cross through it. The starting neighborhoods were counterbalanced between participants. After each passenger delivery, participants were relocated to the starting point in the same neighborhood to start another passenger pick-up. Within each half of the study phase, participants had to successfully find 5 passengers and deliver them to their destinations (PDO locations) in order to finish the study phase.

The *Teleport* condition was similar to the View condition except after learning both neighborhoods separately in each block of the six blocks, the participants watched a video clip showing how the two neighborhoods were connected: moving from starting point A to starting point B (or from B to A). The video clips and starting neighborhoods were counterbalanced between participants.

In the *Whole* condition, the pathway was visible and accessible; therefore, the participants could navigate back and forth between the neighborhoods throughout the study phase in each block. Each time the participant dropped off a passenger, he or she was re-located alternatingly to one or the other starting point, facing either “Mike’s Restaurant” or “Aaron Chang Gallery”, before being cued to collect the next passenger. In all three conditions, which neighborhood was explored first was counterbalanced between participants and there were ten passenger deliveries in each of the six blocks, hence a total of 60 passenger deliveries.

The same memory test and mapping task were used as in the last three blocks in Experiment 1. In all three conditions, at the end of all six blocks, participants were asked to answer a sense of direction questionnaire (see Appendix).

3.4.1.4 Data Analysis

Analyses were the same as those used as in Experiment 1, except we had three conditions. Therefore, we conducted two-way repeated measures ANOVAs, with Location (within vs. between neighborhoods) as a within-subject factor and Condition (View vs. Teleport vs. Whole) as a between-subject factor. We also analyzed the final performance in the last block to investigate whether, with more learning time, the two neighborhoods would eventually be treated as one environment.

In addition, based on the questionnaire participants answered, we were able to investigate individual differences by using correlation analyses. Answers for Questions 1 to 15 (the questions in the original version of the SBSOD) were coded such that larger values indicated a good sense of direction. These ratings were then summed together, yielding final SBSOD scores; these scores were then correlated to recognition accuracy, pointing latency, and pointing errors. For the purpose of our study, we also analyzed Questions 16 and 17 separately. Question 16 had participants rate on a 7-point scale how likely they were to rely on a GPS when they travel to new places, while Question 17 asked participants whether they were video game players (a yes/no question).

We also compared SBSOD scores of participants who correctly mapped the two starting points with those who did not, as a way to test the validity of the questionnaire.

3.4.2 Results

Based on the mapping task, four participants (2 females and 2 males) in the View condition, three participants (2 females and 1 male) in the Teleport condition, and eight participants (6 females and 2 males) in the Whole condition got the relationship between the two neighborhoods wrong. Therefore, their data were excluded from further analyses.

In preliminary analyses, sex and block were included as factors. However, there were no sex differences in any of the spatial measurements; therefore, sex was dropped from further analyses. There were main effects of blocks (i.e. participants showed learning curves) in all the measures, but there were no block by memory score interactions, therefore, we averaged data over blocks for all subsequent analyses although we showed learning curve in figures.

3.4.2.1 Recognition accuracy

A two-way repeated measures Location x Condition ANOVA revealed no significant main effects or interactions (Location [F(1, 42)=3.85, *p*=0.057];

Condition [F(2, 42)=0.48, *p*=0.624], Location and Condition [F(2, 42)=3.09, *p*=0.056].

3.4.2.2 Pointing latency

A two-way repeated measures Location x Condition ANOVA of the pointing latencies revealed a main effect of Condition [F(2, 42)=3.87, *p*=0.029], but no main effect of Location [F(1, 42)=0.062, *p*=0.805] or an interaction between Condition and Location [F(2, 42)=0.85, *p*=0.434]. Post hoc comparisons showed that pointing latency was significantly faster in the Whole condition than in View condition (*p*=0.033), but no other pairwise differences were significant (View condition: mean=4.23, SE=0.34; Teleport condition: mean=4.45, SE=0.33; Whole condition: mean=3.07, SE=0.40).

3.4.2.3 Pointing errors

A two-way repeated measures Location x Condition ANOVA revealed significant main effects of Location [F(1, 42)=29.95, *p*<0.001] and Condition [F(2, 42)=4.34, *p*=0.019], but no interaction between Condition and Location [F(2, 42)=0.58, *p*=0.565]. Post hoc comparisons showed that errors in pointing between neighborhoods (mean=41.48, SE=1.99) were significantly larger than those made when pointing within neighborhoods (mean=29.84, SE=1.97). Pointing errors in the Whole condition (mean=42.64, SE=3.20) were significantly larger than those in the View condition (*p*=0.017, mean=30.34, SE=2.77), while there was no difference between View and Teleport conditions (*p*=1.00) or between Teleport and Whole conditions (*p*=0.135) (see Figure 3.5).

3.4.2.4 Final performance

In all measures participants' performance improved over blocks (e.g. see Figure 3.6 for pointing errors). We therefore investigated whether Location effects dissipated after several blocks of learning. In particular, we predicted that the differences in pointing between- versus within-neighborhoods might disappear once sufficient learning had taken place. Participants' final performance in the last block was analyzed in terms of recognition accuracy, pointing latency, and pointing errors by using two-way repeated measures Location x Condition ANOVAs. There were no significant main effects or interactions in our analyses of recognition accuracy or pointing latency. However, in terms of pointing errors, even in the final block of learning; there was a significant main effect of Location [F(1, 42)=26.07, p<0.001], but no main effect of Condition [F(2, 42)=2.34, p=0.109] and no interaction between Location and Condition [F(2, 42)=0.128, p=0.88] (see Figure 3.6). As in the overall analysis across all blocks, in the final block, pointing within neighborhoods resulted in smaller errors (mean=22.65, SE=1.87) than pointing between neighborhoods (mean=36.71, SE=2.36) across conditions.

3.4.2.5 Questionnaire results

T-tests revealed that participants who correctly mapped the two starting points had higher SBSOD scores (t=2.17, df=58, p=0.039) and recalled more stores in the mapping task (t=3.372, df=58, p=0.001) than those who did not, suggesting that participants who had a better sense of direction performed better overall in our experiment, thus validating the SBSOD questionnaire as a measure of spatial abilities.

3.4.2.5.1 Video Game Players

Across conditions, participants with video game experience (gamers) (n=27, mean=16.04, SE=0.47) recalled significantly more restaurants and shops during the mapping task than those without video game experience (non-gamers) (n=18, mean=13.83, SE=0.8) (*t*=2.53, *df*=43, *p*=0.015), but were no better in terms of recognition accuracy, pointing latency, or pointing errors.

3.4.2.5.2 View condition

In the View condition, SBSOD scores were not significantly correlated with any of the measurements (see Table 3-1). GPS usage (Q16, on a 7-point scale) was significantly positively correlated with recognition accuracy (*r*(16)=0.533, *p*=0.033) and negatively correlated with pointing errors (*r*(16)=-0.555, *p*=0.026) when pointing between neighborhoods. Note that high scores for Q16 indicated that participants were less likely to rely on GPS when they travel to new places.

3.4.2.5.3 Teleport condition

In the Teleport condition, SBSOD scores were significantly negatively related to pointing errors for within-neighborhood PDO locations, *r*(16)= -0.718, *p*=0.001 (see Table 3-1). However, there were no significant correlations between any of the other measures and the SBSOD scores, nor was GPS usage correlated with any of the measurements.

3.4.2.5.4 Whole condition

In the Whole condition, the SBSOD scores were significantly negatively correlated with pointing latency for between-neighborhood PDO locations, *r*(12)=-0.591, *p*=0.043, and marginally negatively correlated pointing latency for within-neighborhood PDO locations, *r*(12)=-0.572, *p*=0.052 (see Table 3-1). No other correlations with between the SBSOD scores were significant in this condition (see Table 3-1), nor were there any significant correlations between GPS usage and any of the measurements in the Whole condition.

3.4.3 Discussion

In all three conditions, regardless of whether the two neighborhoods were explored separately or jointly, participants were more accurate at pointing to PDO locations within neighborhoods than at pointing between neighborhoods. This difference persisted even in the final block, in all three conditions (see Figure 3.6), suggesting that even after extensive learning, the two neighborhoods were still treated as two separate environments.

There clearly were individual differences in how participants carried out the pointing task. For example, in the View condition, participants who relied less on a GPS in daily life were better at pointing to PDO locations between neighborhoods. The View condition poses the greatest challenge to participants in integrating their knowledge of the two neighborhoods. In contrast to the teleport and whole conditions, participants never visually experience moving along the corridor to see how the neighborhoods are connected. They have to piece the two neighborhoods together without actually either actively or passively navigating between them. Those who do not rely on a GPS may be more adept at visuo-spatial imagery, and therefore better able to imagine moving along the connecting pathway without having directly experienced it.

Interestingly, whereas GPS use negatively predicted performance in the View condition, sense of direction scores positively predicted performance in the teleport and whole conditions. Those who had higher SBSOD scores had the advantage in both conditions, as evidenced by greater pointing accuracy in the Teleport condition and shorter pointing latencies in the Whole condition. It is possible that these two conditions favor two distinctly different response strategies, a topic for future research.

Although the results of this experiment are consistent with the notion that participants treated the two neighborhoods as two separate environments in all three conditions, an alternative explanation is that the participants had a global map of the two neighborhoods (or could precisely combine the two separate maps), but made errors in judging the *length* of the pathway connecting the two neighborhoods, resulting in larger errors in pointing to PDO locations between neighborhoods. This alternate hypothesis is weakened by the observation that participants in all three conditions exhibited the same PDO location effect even though they had different experiences with the connection between the neighborhoods. Nevertheless, we wanted to provide a stronger test of the hypothesis. Therefore, in Experiment 3, the fences between the two neighborhoods were removed while all other features of the neighborhoods remained the same. As a result, the possibility of participants misjudging the length of the pathway should be reduced, and therefore, pointing errors for within- and between-neighborhood PDO locations ought to be the same.

3.5 Experiment 3

In Experiment 3, in order to encourage participants to form a single global map of the environment, they explored both neighborhoods jointly from the very beginning, as in the “Whole” condition in Experiment 2, and fences separating the two neighborhoods were removed (see Figure 3.7). However, in this experiment, the town still consisted of two distinct Restaurant vs. Shopping districts surrounded by differently colored fences, which would still evoke the perception of two separate neighborhoods as in Experiment 2. Therefore, we hypothesized that pointing errors for within-neighborhood PDO locations would be smaller than pointing errors for between-neighborhoods. This result would suggest that the pointing error difference observed in Experiment 2 was not due to the misjudgment of the length of the pathway, but rather to participants basing their responses on two separate local maps.

3.5.1 Method

3.5.1.1 Participants

Twenty McMaster University students (9 males and 11 females) participated in this experiment; age ranged from 18 to 30 years, and the mean was 20.9. Participants had normal or corrected-to-normal vision. Participants received either two course credits or $20 for taking part in this 2-hour experiment. This study was reviewed and approved by the McMaster Research Ethics Board. Written informed consent was obtained from all participants involved in this study.

3.5.1.2 Materials

The materials were the same as those used in Experiment 2, except that the fences between the two neighborhoods were removed (see Figure 3.7). Therefore, there was no pathway connecting the two neighborhoods. However, the two neighborhoods were still visually distinct because the fences surrounded each neighborhood differed in color and texture, and one neighborhood contained only restaurants while the other contained only shops.

3.5.1.3 Procedure & Data Analysis

The procedures and analyses were the same as those used in Whole condition in Experiment 2.

3.5.2 Results

3.5.2.1 *Recognition accuracy*

A two-tailed paired sample t-test of the recognition accuracy revealed a significant difference between the two types of PDO locations (*t*=-2.57, *df*=19, *p*=0.019). Recognition accuracy was better for between-neighborhood PDO locations (mean=98.61%, SE=0.005) than for within-neighborhood PDO locations (mean=96.67%, SE=0.008).

3.5.2.2 Pointing latency

A two-tailed paired sample t-test of the pointing latency revealed no significant difference between pointing to PDO locations within versus between neighborhoods (t=0.122, df=19, p=0.904; within-neighborhood PDO locations: mean=3.50, SE=0.21; between-neighborhood PDO locations: mean=3.48, SE=0.24).

3.5.2.3 Pointing errors

A two-tailed paired sample t-test of the pointing errors revealed a significant difference between the two types of PDO locations (t=-14.00, df=19, p<0.001), in which errors were smaller for pointing PDO locations within neighborhood (mean=22.20, SE=1.78) than between neighborhood (mean=41.95, SE=1.63) (see Figure 3.8).

3.5.2.4 Questionnaire results

Participants with video game experience (n=10) and those without video game experience (n=10) did not differ in terms of recognition accuracy, pointing latency, pointing errors, or number of restaurants and shops recalled.

None of the questionnaire measurements were significantly correlated with the SBSOD scores (see Table 3-2). GPS usage was significantly negatively correlated with pointing latency for between-neighborhood PDO locations, r(20)= -0.502, p=0.024. In summary, less GPS usage was associated with faster pointing responses for between-neighborhood PDO locations.

3.5.3 Discussion

The results were similar to those obtained in the Whole condition in Experiment 2: pointing to PDO locations was more accurate within neighborhoods than pointing between neighborhoods, even in the last block (see Figure 3.9), even in the face of better recognition memory for between-neighborhood PDO locations. These findings argue against the alternative hypothesis that larger between-neighborhood pointing errors were due to the misjudgment of the length of the pathway connecting the two neighborhoods. Even without fences to separate the two neighborhoods, participants still behaved as if they treated them as two separately encoded environments. Moreover, participants who relied on a GPS more often in their daily lives were slower at pointing to PDO locations in the adjacent neighborhoods.

One possible alternative explanation for the current findings is that the environment simply was too big to be encoded as one map. Or perhaps the two types of PDO locations differed in important ways: For example, it is possible that the difference in pointing errors reflects the fact that all within-neighborhood PDO locations were located around the edge of the town whereas between-neighborhood PDO locations were located in the centre of the town. To rule out these possibilities, we conducted another experiment in which the types of PDO locations were mixed between the two neighborhoods and the environment no longer contained distinct boundaries delineating the two neighborhoods.

3.6 Experiment 4

The difference between this experiment and Experiment 3 was that the fences were replaced by uniformly textured walls and the restaurants and shops were intermixed within the town. Therefore, unlike previous experiments, there were no spatial or visual features to differentiate the two neighborhoods. Nevertheless, we still analyzed the same 6 pairs of PDO locations (see Figure 3.10) as we did in the previous experiments, and to aid comparisons across experiments we therefore still refer to them as within- and between-neighborhood PDO locations. We hypothesized that there would be no difference in pointing errors for within and between neighborhoods PDO locations. This result would suggest that the large pointing error differences observed in Experiments 1-3 were caused by participants constructing and using separate maps of the distinct neighborhoods, whenever there were features available to differentiate and cluster local landmarks into sub-regions.

3.6.1 Method

3.6.1.1 Participants

Twenty (11 males and 9 females) McMaster University students participated in the experiment; age ranged from 18 to 35 years, and the mean was 20.6. Participants had normal or corrected-to-normal vision. Participants received either two course credits or $20 for taking part in this 2-hour experiment. This study was reviewed and approved by the McMaster Research Ethics Board. Written informed consent was obtained from all participants involved in this study.

3.6.1.2 Materials

The materials were the same as those used in Experiment 3, except that the layout of the town was changed: The fences separating the two neighborhoods were removed, and restaurants and shops were intermixed across the town rather than being localized, respectively, within two distinct neighborhoods (see Figure 3.10). Therefore, there was no pathway connecting the two neighborhoods, and there were no other distinctions between the two neighborhoods. In this case, therefore, the town should be perceived as one environment.

3.6.1.3 Procedure & Data Analysis

The procedures and analyses were the same as those used in Experiment 3.

3.6.2 Results

3.6.2.1 *Recognition accuracy*

A two-tailed paired sample t-test of the recognition accuracy revealed no significant difference between the two types of PDO locations (t=-1.19, df=19, p=0.249; within-neighborhood PDO locations: mean=96.53%, SE=0.008; between-neighborhood PDO locations: mean=97.36%, SE=0.007).

3.6.2.2 Pointing latency

A two-tailed paired sample t-test of the pointing latency revealed no significant difference between the two types of PDO locations (t=1.43, df=19, p=0.169; within-neighborhood PDO locations: mean=3.91, SE=0.33; between-neighborhood PDO locations: mean=3.69, SE=0.30).

3.6.2.3 Pointing errors

A two-tailed paired sample t-test of the pointing latency revealed no significant difference between the two types of PDO locations (*t*=-1.74, *df*=19, *p*=0.098; within-neighborhood PDO locations: mean=35.25, SE=3.16; between-neighborhood PDO locations: mean=39.29, SE=2.21) (see Figure 3.11).

3.6.2.4 Questionnaire results

Participants with video game experience (n=12) and those without video game experience (n=8) did not differ on any of the performance measures.

The questionnaire data were analyzed using correlation. Because there was no difference between the two types of PDO locations, we correlated these averaged scores to the SBSOD scores and GPS usage: none of the correlations were significant (see Table 3-2).

3.6.3 Discussion

The results demonstrated that performance improved across blocks, which indicates that participants were learning the task during the experiment. However, unlike previous experiments, there was no indication that errors depended on PDO location. We interpret this result as showing that participants encoded the large scale environment within a single cognitive map. Furthermore, the results of this experiment rule out the possibility that the larger pointing errors for between-neighborhood PDO locations found in the previous experiments were due to the size of the environment or the positions of within- and between-neighborhood PDO locations relative to the edges of the environment.

3.7 General Discussion

Our results support the view that people have a strong tendency to use separate maps to encode spatial information in adjacent neighborhoods, whether they are explored separately or jointly, provided that the neighborhoods are distinct.

Initially we analyzed sex differences, but failed to find any significant differences between males and females in any of the spatial memeory measurements. Therefore, we dropped this factor from our current analyses. However, there were individual differences in terms of sense of direction and GPS usage. For example, GPS usage and video game experiences have different effects on spatial memory and recall memory. When the two neighborhoods were distinct, video game players showed better recall memory during the mapping task than non-players. This difference disappeared when the virtual town was presented as one environment. Moreover, the more GPS usage, the less accurate the pointing responses for between neighborhoods PDO locations in Experiment 2, but only in the View condition. It is unclear what is producing these correlations. It could be that people who rely on a GPS when travelling to new places have poor mental imagery abilities, which would translate into poorer navigation abilities in real life, as well as greater difficulty imaging how the two neighborhoods are connected without direct visual experience in the View condition.

What types of spatial representations did participants use in our tasks? Two possible types of cognitive map have been proposed to underlie spatial cognition: strip-like maps or relatively broad and comprehensive maps (Tolman, 1948). In a strip-map, an animal's position and the goal position are connected by a single path, which is less flexible in the face of changes in the environment. In contrast, in a comprehensive map, the animal would be able to behave correctly with changes made in the environment (see Tolman, 1948 for a review). McNaughton and colleagues (1996) suggested that an abstract mental representation of a two-dimensional environment requires input from place cells and head direction cells. which convey self-motion information for path integration and the landmark information, respectively, could be used to correct for errors that accumulated during path integration, however, without landmark information, the system still works (McNaughton et al., 1996). Trullier and Meyer (2000) proposed a computational model of navigation, in which they consider the hippocampus as a "cognitive graph" (a.k.a hetero-associative network). The temporal sequences of visited places are learned and an environment's topological representation is formed by using a "place-recognition-triggered response" strategy, which is stored by the network. In this case, the model could predict the next position based on the current position by using place cells, goal cells, and sub-goal cells.

If one has a comprehensive cognitive map of multiple environments, how this global map could be formed by combining local maps is unclear. They could be connected hierarchically, in which the multiple maps are combined as a global map in a coarser scale, or they could be chained together via an associative learning process as particularly trajectories are experienced, resulting in a representation of a single connected pathway. In either case, there would be an accuracy cost when inferring locations across maps. When the two environments are always connected by a path, people may combine them as a chain of maps. In contrast, when there are multiple connections between the two environments and with extensive experiences with the multiple connections, people may be able to combine them as a global map.

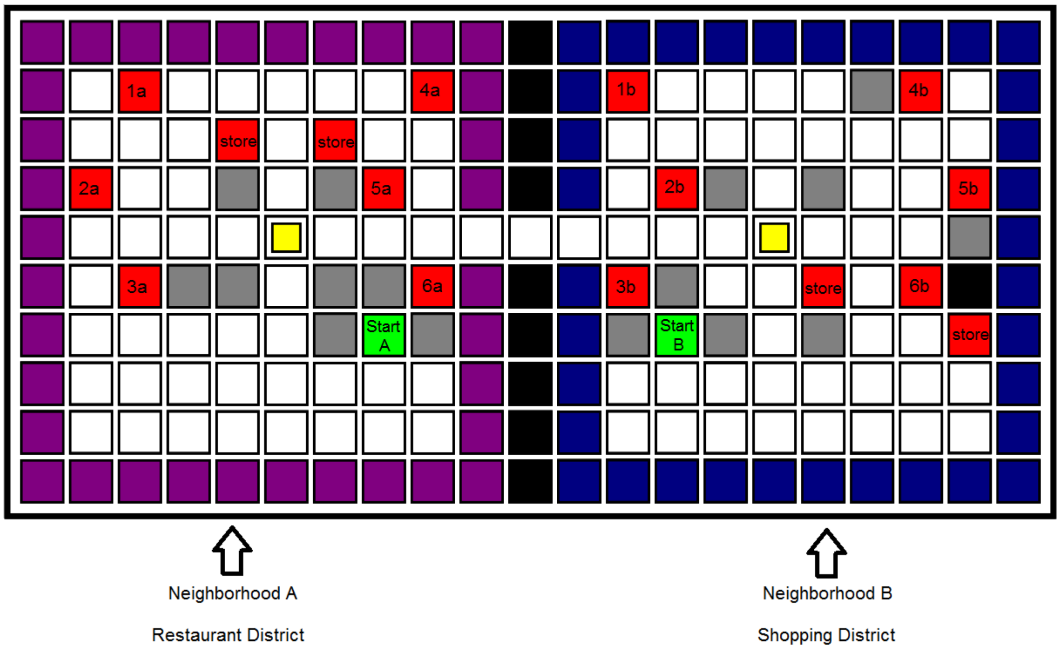
Table 3-1 Correlation r between SBSOD scores and recognition accuracy, pointing latency, and pointing errors in Experiment 2.

	Location	View n=16	Teleport n=17	Whole n=12
Recognition Accuracy	Within	0.102 *p*=0.707	0.201 *p*=0.440	0.119 *p*=0.712
	Between	0.143 *p*=0.598	0.247 *p*=0.338	0.523 *p*=0.081
Pointing Latency	Within	-0.351 *p*=0.183	-0.196 *p*=0.451	-0.572 *p*=0.052
	Between	-0.362 *p*=0.169	-0.083 *p*=0.751	-0.591 *p*=0.043*
Pointing Errors	Within	0.040 *p*=0.883	-0.718** *p*=0.001	0.125 *p*=0.698
	Between	0.093 *p*=0.733	-0.418 *p*=0.733	-0.135 *p*=0.676
Number of store recalled		0.439 *p*=0.089	0.449 *p*=0.071	-0.004 *p*=0.990

Table 3-2 Correlation r between SBSOD scores and recognition accuracy, pointing latency, and pointing errors in Experiments 3 and 4.

	Location	Experiment 3 n=20	Experiment 4 n=20
Recognition Accuracy	Within	-0.076 *p*=0.749	-0.061 *p*=0.779
	Between	-0.018 *p*=0.940	
Pointing Latency	Within	-0.075 *p*=0.755	-0.035 *p*=0.883
	Between	-0.239 *p*=0.311	
Pointing Errors	Within	0.140 *p*=0.557	0.241 *p*=0.306
	Between	-0.217 *p*=0.358	
Number of store recalled		0.159 *p*=0.504	0.022 *p*=0.928

Figure 3.1 Town used in Experiments 1 and 2.



Town's layout (21 by 10 VR units in size) used in Experiments 1 and 2. The grey squares are non-distinctive uniformly textured buildings at locations where the participants are not able to drive into. The black squares are places also where the participants are not able to drive into. The red "store" squares and red numbered squares are the stores that serve as passenger drop-off locations. The two green "Start" squares are the two starting points, locating at the two neighborhoods ("Mike's Restaurant" and "Aaron Chang Gallery"). The smaller yellow squares are the two objects, one in each neighborhood. All the white squares indicate locations along the routes that participants can navigate in the town.

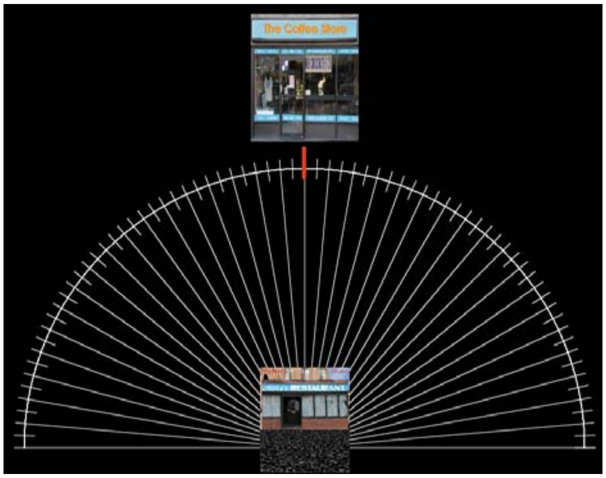
Figure 3.2 Starting points.

(a) Mike's Restaurant starting point

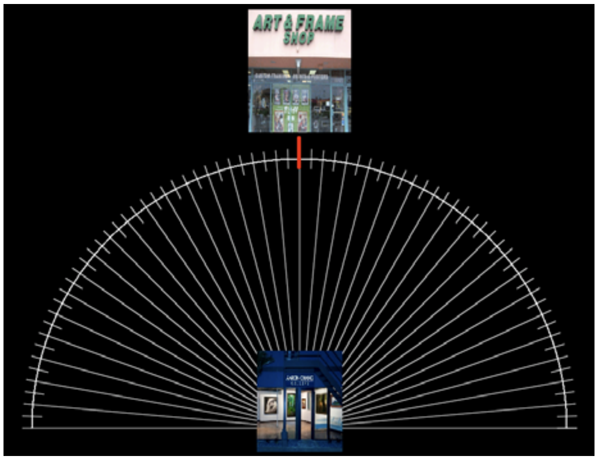
(b) Aaron Chang Gallery starting point

Figure 3.3 Testing viewpoints.

(a) Tested viewpoint: Mike's Restaurant

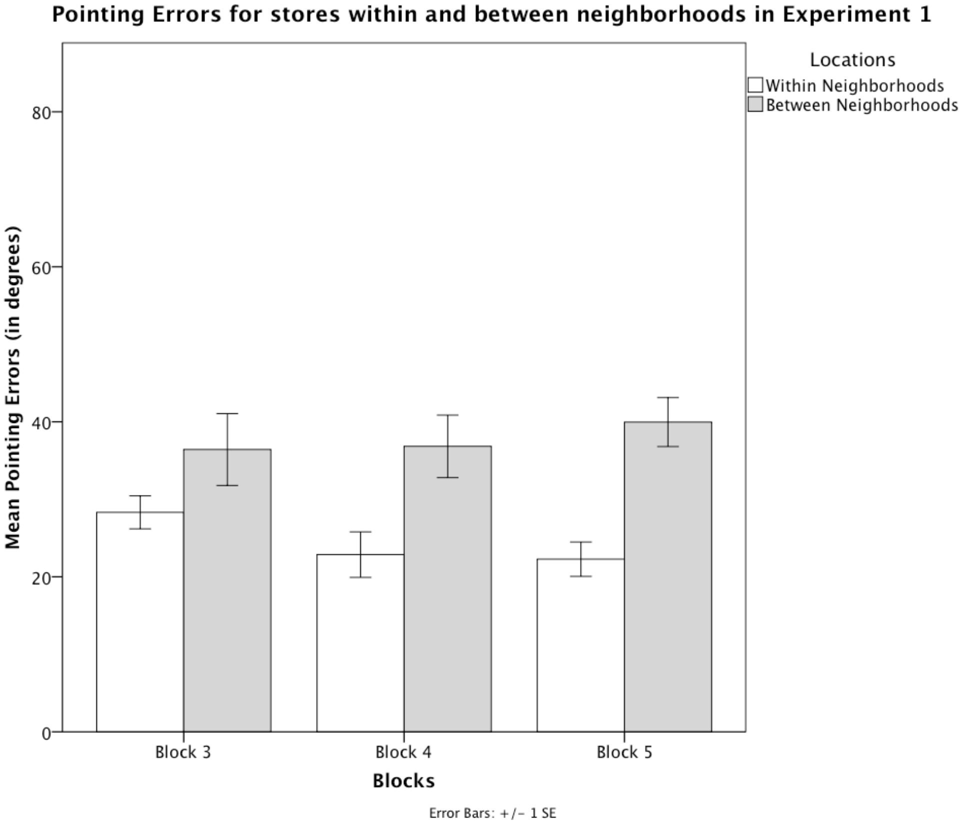


(b) Tested viewpoint: Aaron Chang Gallery



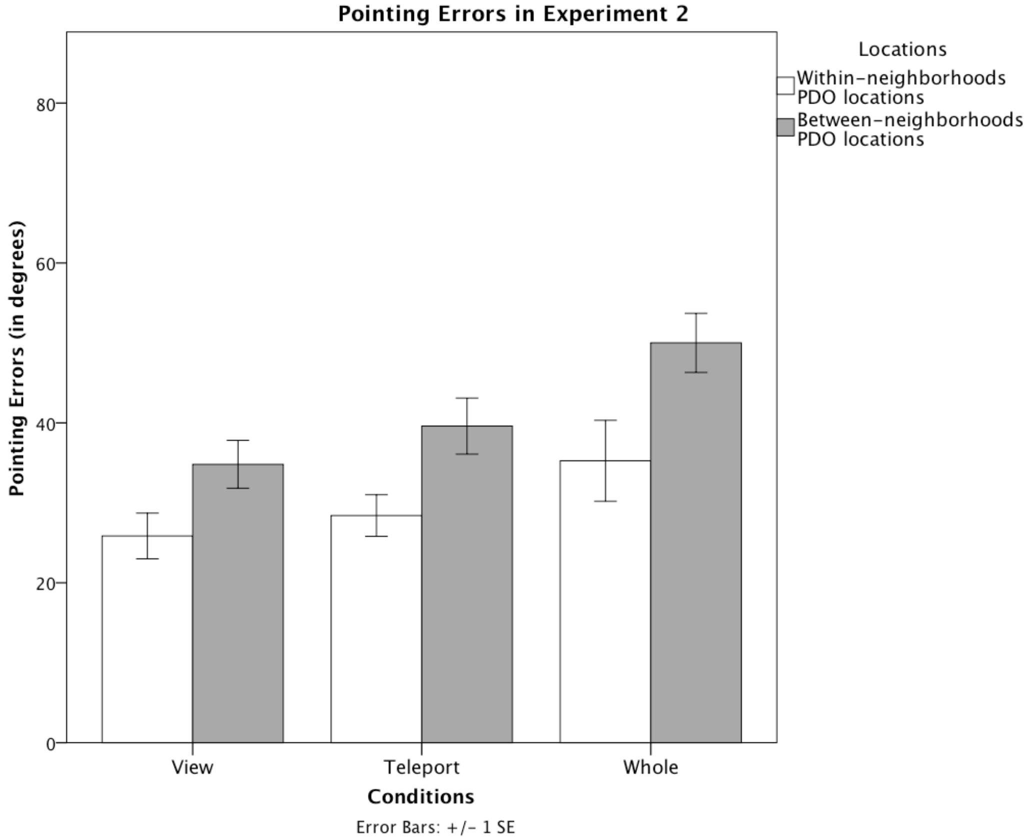
Navigators used in the pointing task. At the tip of the compass pointer (red) an image was shown of the target store for the current trial. It could be moved by moving the pointer. (a) Navigator from "Mike's Restaurant" point of view; (b) Navigator from "Aaron Chang Gallery" point of view.

Figure 3.4 Pointing errors in Experiment 1.



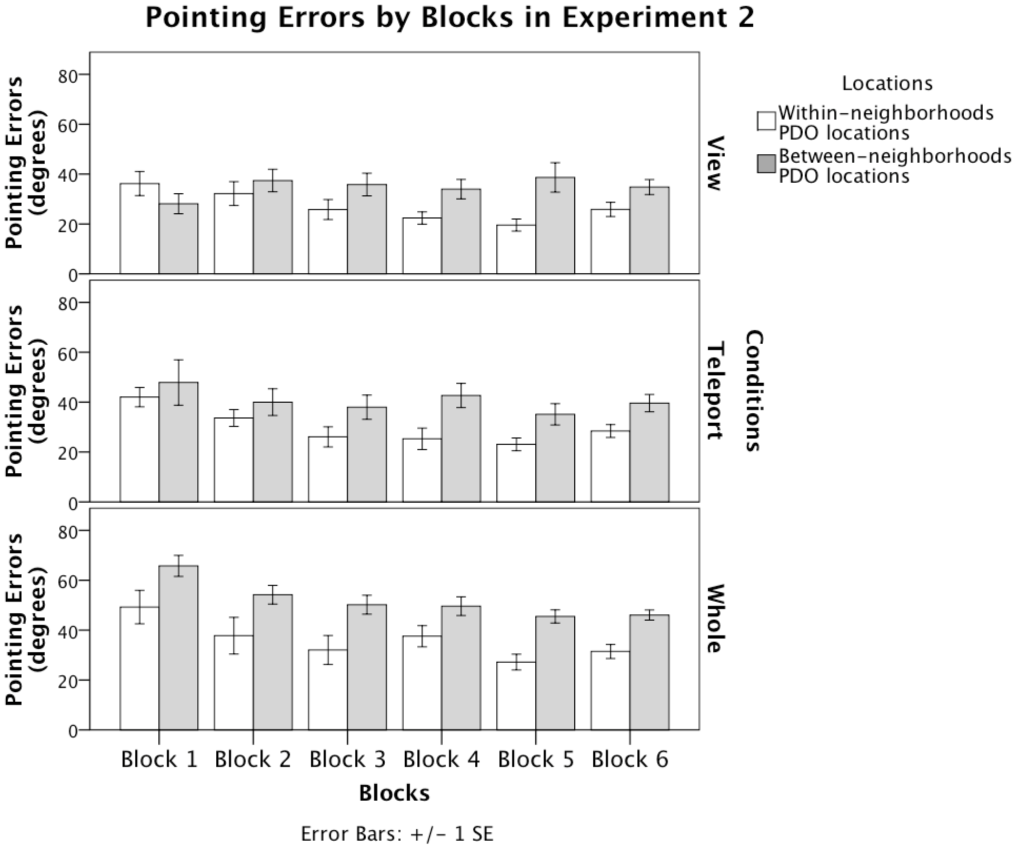
Errors for pointing to PDO locations within vs. between neighborhoods in each block in Experiment 1. White bar is for within neighborhoods stores and grey bar is for between neighborhoods stores. The pointing errors were significantly smaller for stores within the neighborhoods than for those between neighborhoods.

Figure 3.5 Pointing errors in Experiment 2.



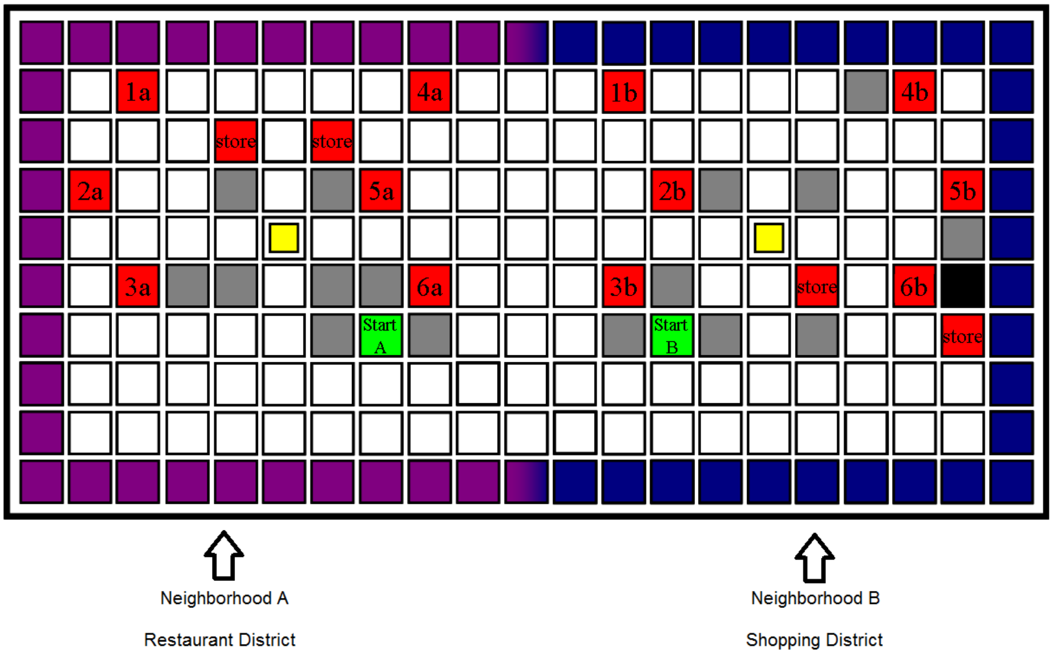
Errors for pointing to PDO locations within vs. between neighborhoods in each condition in Experiment 2. White bar is for within neighborhoods stores and grey bar is for between neighborhoods stores. Left side bars are for View condition, middle bars are for Teleport condition, and right side bars are for Whole condition. The pointing errors were significantly smaller for within-neighborhood PDO locations than for between-neighborhood PDO locations in all three conditions.

Figure 3.6 Pointing errors by blocks in Experiment 2.



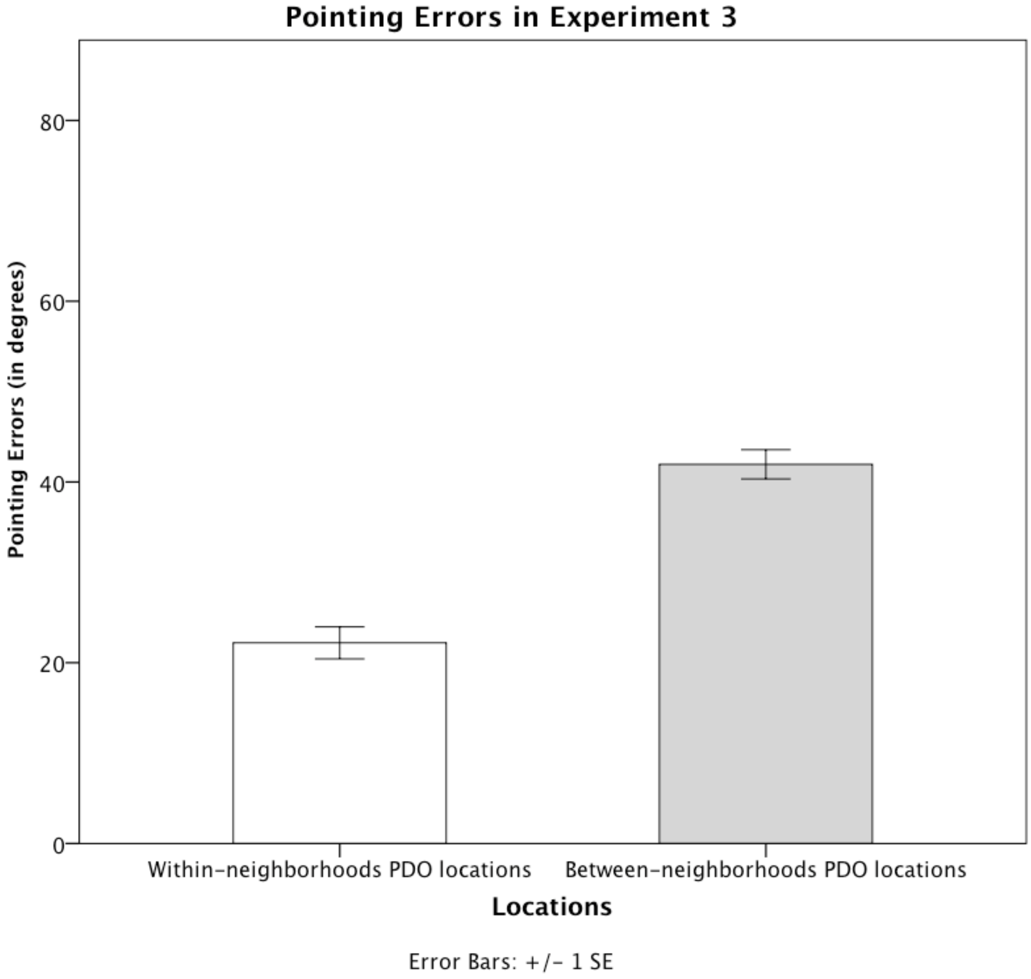
Errors for pointing to PDO locations within vs. between neighborhoods in each block in Experiment 2. White bar is for within neighborhoods stores and grey bar is for between neighborhoods stores. Top row is for View condition, middle row is for Teleport condition, and the bottom row is for Whole condition.

Figure 3.7 Town used in Experiment 3.



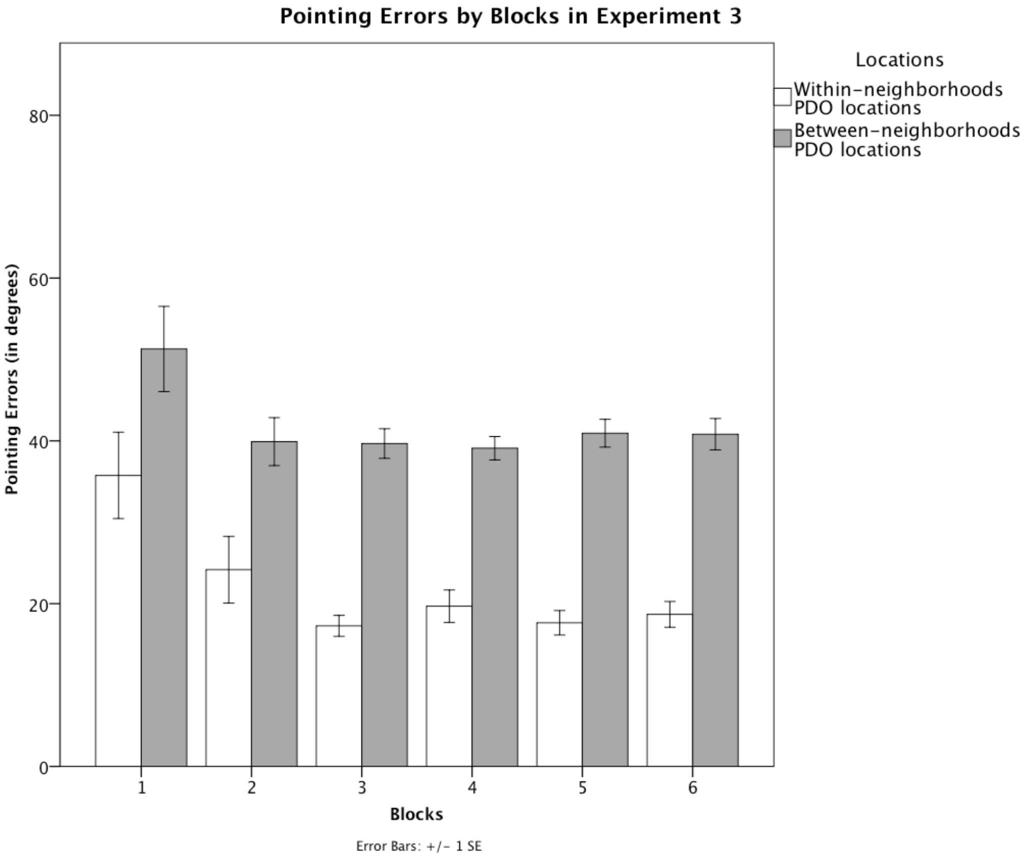
Town's layout (21 by 10 VR units in size) used in Experiment 3, similar to the Town used in Experiments 1 and 2 except the fences between the two neighborhoods were removed.

Figure 3.8 Pointing Errors in Experiment 3.



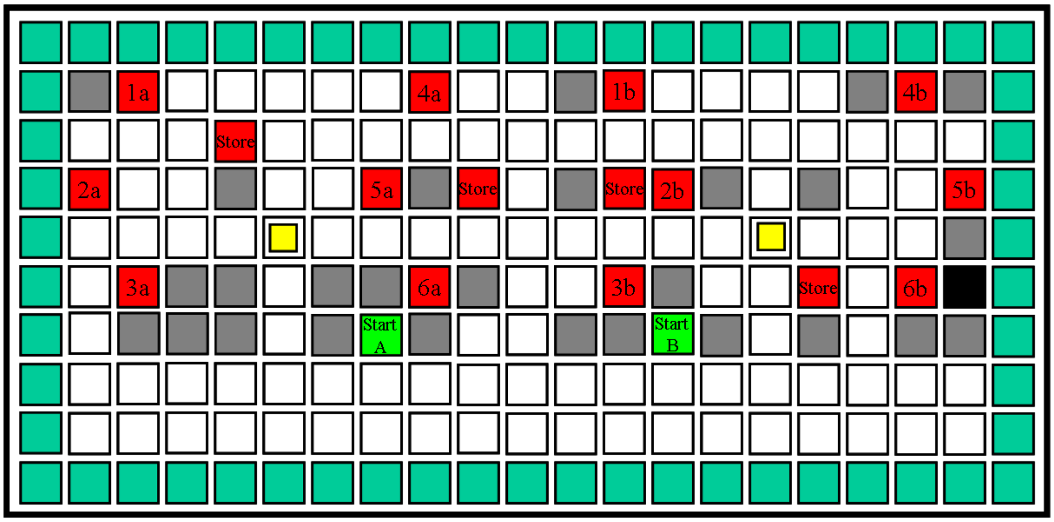
Errors for pointing to PDO locations within vs. between neighborhoods in Experiment 3. White bar is for within neighborhoods stores and grey bar is for between neighborhoods stores. The pointing errors were significantly smaller for within-neighborhood PDO locations than for between-neighborhood PDO locations.

Figure 3.9 Pointing Errors by blocks in Experiment 3.



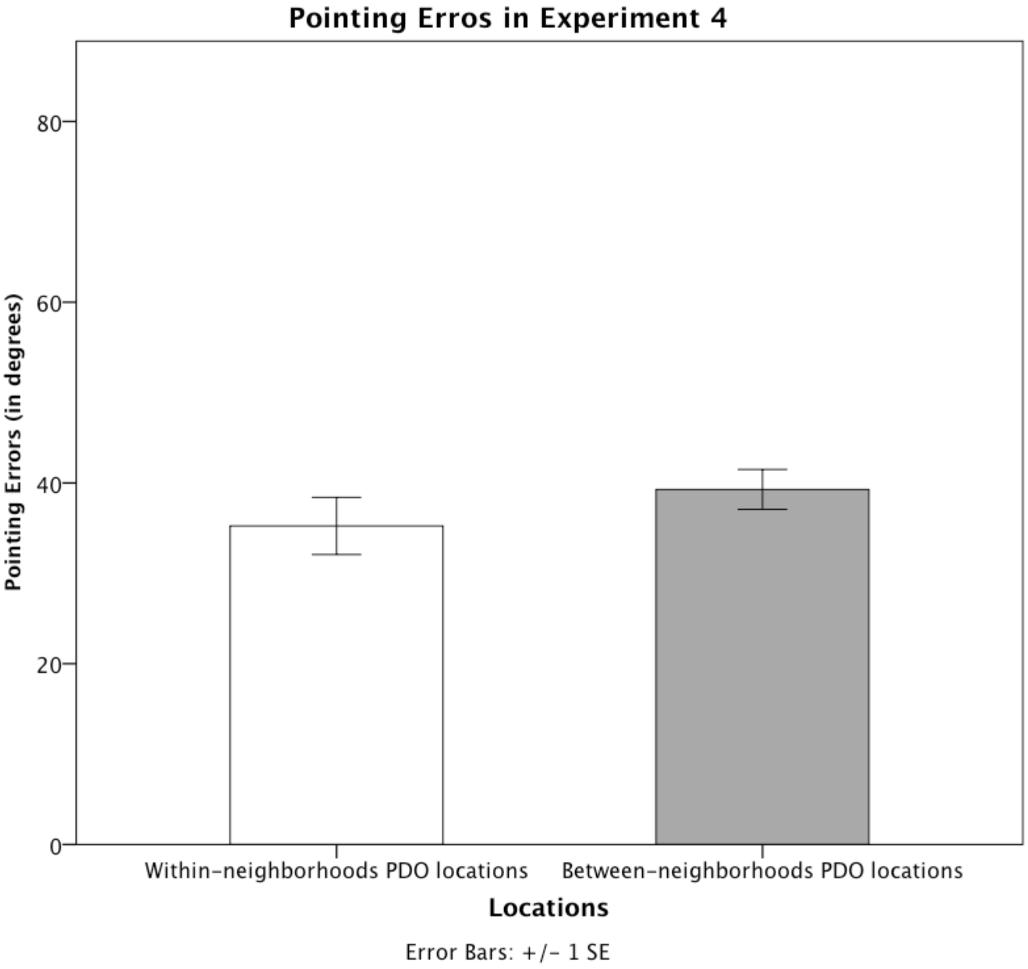
Errors for pointing to PDO locations within vs. between neighborhoods in each block in Experiment 3. White bar is for within neighborhoods stores and grey bar is for between neighborhoods stores.

Figure 3.10 Town used in Experiment 4.



Town's layout (21 by 10 VR units in size) used in Experiment 4.

Figure 3.11 Pointing Errors in Experiment 4.



Errors for pointing to PDO locations within vs. between neighborhoods in Experiment 4. White bar is for within neighborhoods stores and grey bar is for between neighborhoods stores. There was no difference between the two types of PDO locations than for between-neighborhood PDO locations.

3.8 Appendix

Gender: M F Date:

This questionnaire consists of several statements about your spatial and navigational abilities, preferences, and experiences. After each statement, you should circle a number to indicate your level of agreement with the statement. Circle "1" if you strongly agree that the statement applies to you, "7" if you strongly disagree, or some number in between if your agreement is intermediate. Circle "4" if you neither agree nor disagree. For question 17, circle "Yes" or "No" for your answer.

Q1. I am very good at giving directions.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q2. I have a poor memory for where I left things.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q3. I am very good at judging distances.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q4. My "sense of direction" is very good.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q5. I tend to think of my environment in terms of cardinal directions (N, S, E, W).

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q6. I very easily get lost in a new city.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q7. I enjoy reading maps.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q8. I have trouble understanding directions.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q9. I am very good at reading maps.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q10. I don'*t* remember routes very well while riding as a passenger in a car.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q11. I don'*t* enjoy giving directions.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q12. It's not important to me to know where I am.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q13. I usually let someone else do the navigational planning for long trips.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q14. I can usually remember a new route after I have traveled it only once.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q15. I don'*t* have a very good “mental map of my environment.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q16. I use GPS when travel to a new place.

(Strongly agree) 1 2 3 4 5 6 7 (strongly disagree)

Q17. I play video games.

a. Yes

b. No

4 Chapter 4: General Discussion

4.1 Summary and contributions

The current research sheds light on the types of spatial navigation strategies that humans use and what kinds of spatial representations are employed in large complex environments. Our results suggest that people use cognitive map-based navigation in large complex environments, appear to use different spatial representations for different types of objects in the environments based on navigational relevance, and construct multiple maps for the connected environment.

4.1.1 Chapter 2

As described in Chapter 1, a ‘cognitive map’ (a.k.a. allocentric mental representation) consists of path/route and landmark information and relative positions of different places (O’Keefe & Nadel, 1978; Tolman, 1948). However, it is unclear what exactly constitutes a landmark. Landmarks could be distal large orienting cues such as mountains and tall office towers and/or proximal cues such as buildings and trees within an environment. Even smaller objects such as fire hydrants, mailboxes etc could serve as landmarks that cue navigational choices. This raises a major question that has not been resolved in the spatial cognition literature: what factors affect whether an object would be considered as a landmark or simply as an object?

Our results in Chapter 2 filled this gap by showing that the location of objects and number of experienced viewpoints were the primary determinants of whether the objects would be treated as landmarks to serve navigation and incorporated into an allocentric representation of the environment unconsciously. However, when subjects' attention was manipulated to focus on either the location of the objects or the appearance of the objects, whether the objects were located at navigationally relevant points or not was no longer the determinant of landmark encoding, in which all the objects were treated equally. Moreover, our novel pointing consistency measure, consistency of pointing errors across testing viewpoints, proved to be sensitive to this manipulation of navigational relevance, and showed consistent results in both active and passive navigation conditions. Moreover, with this new measurement, it is possible to test allocentric representations from multiple different viewpoints, which is not possible based on correlational analyses of pointing errors or the consistency of signs of pointing errors.

It has been suggested that for the purpose of mental imagery and guiding our actions in space, we retrieve allocentric representations from long-term memory and transform them into egocentric or viewer-centered mental images via the dorsal visual stream or so-called “where” or “how to” pathway, distinguished from the ventral visual stream or so-called "what" pathway; these two visual streams converge at the level of the hippocampus, a structure known to be important for binding together information into long-term episodic memories--memories for events set in particular spatio-temporal contexts (Byrne, Burgess, & Becker, 2007). However, the BBB model does not specify what sorts of features are processed by the visual system to create allocentric representations. We proposed here that an object could be processed through either the dorsal visual stream or the ventral visual stream depending on where it is located or how it is attended. If an object was at a navigationally relevant point, we hypothesized that it would be processed through the dorsal visual pathway and incorporated into an allocentric representation. In contrast, if an object was at a navigationally less relevant point, it was predicted to be processed through the ventral visual pathway and incorporated into egocentric representations. Our results were broadly consistent with these predictions. Moreover, when an object's location was attended, our results were consistent with the notion that it would be processed through the dorsal visual pathway, but whether it would be incoporated into an allocentric representation would depend on whether it was located at a navigationally relevant point.

4.1.2 Chapter 3

Previous research has suggested that humans construct separate mental maps for a multiple-level building (Montello & Pick, 1993), however, it is less clear whether humans create one or multiple maps of an environment consisting of multiple neighborhoods on the same horizontal level during navigation. The current study investigated this question in Chapter 3 and found evidence that people had a very strong tendency to construct multiple maps for an environment with multiple segments. In our experiments, we constructed a virtual town with two distinct neighborhoods (i.e., Restaurant district vs. Shopping district) separated either by fences with different colours and textures but connected by a pathway or by empty spaces in between. If subjects encoded the two neighborhoods into two separate cognitive maps, when pointing to places in another map from a location in the current map, there would be a cost in pointing accuracy. Moreover, not only the physical separation (i.e., fences) between the segments but also the categorical differences (i.e., Restaurant district vs. Shopping district) between the connected segments and perception of separate layouts (i.e., empty space between the two neighborhoods) of the segments were contributing factors to the construction of separate cognitive maps. However, when all distinguishing features between the neighborhoods were removed, people appeared to create one large cognitive map of the environment. Our results suggested that people were more accurate in pointing when they had multiple small cognitive maps than when they had a large cognitive map of the environment.

4.2 Limitations

Although our study provided valuable inputs to the current spatial cognitive research, it has limitations. Our study employed virtual environments, which have been extensively used in spatial cognition research especially in neuroimaging studies. Virtual reality has many advantages over real environments such as better control of the experimental settings, easily varying complexity of environments and constructing large-scaled environments, and it is feasible for use in neuroimaging studies of spatial cognition in humans. However, it has disavantages such as the lack of body-based input, motion sickness, misperception, or/and disorientation. For example, when we navigate in a real environment, the body-based input is part of the input that we use to calculate our location via path-integration. However, in VR, this input is lacking. Moreover, VR can cause motion sickness in some people and may cause disorientation or misperception due to the unnaturally narrow visual field. It would be of interest to use the same procedure in real environments to investigate whether adding body-based input has any effect on allocentric encoding. With body-based input, one may be more accurate in terms of encoding locations of objects, but this may not affect whether an object would be treated as a landmark.

Moreover, the current work showed that video game players (gamers) were more efficient in navigation than non-video game players (non-gamers), which might be an effect of using virtual environments per se. Whether this effect was due to the use of virtual environments or the gamers were actually more efficient in navigation in general should be investigated in the future using real environments.

4.3 Future studies

Although the gamers were more efficient in navigation in the virtual environments, surprisingly, this increase in speed and accuracy in VR navigation did not translate into superior allocentric spatial mapping abilities. The question arises, what are the underlying differences between gamers and non-gamers. Are they using different strategies? For example, gamers may be better at encoding specific routes but poorer at encoding global maps. Future research could investigate this by employing a larger battery of spatial mapping and route-finding tasks. I hypothesize that gamers would out-perform non-gamers on well learned routes, but those with superior alloccentric abilities would tend to be non-gamers and would be superior at dealing with short-cuts and unexpected detours. Further, the specific type of game played most often could bias people to rely more on one strategy or the other.

Although our current study showed that objects at navigationally relevant points were more likely to be encoded into allocentric representations of the environment and experiencing an object from multiple views also contributed to allocentric coding, it has not been clearly established whether experiencing multiple viewpoints is strictly necessary for forming allocentric representations. Further, allocentric representations may be completely invariant, or they may correspond to a range of views. Primate hippocampal cell recordings suggest that a mixture of “place cells” and “view cells” are seen in the hippocampus, with view cells responding to an object being viewed from a range of locations (Rolls & Xiang, 2006). Future research could investigate whether people show behavioural evidence for place versus view-based representations of objects, landmarks, and goal locations. Goal locations are of particular interest as they are maximally relevant to navigation but may be visited less often than for instance choice points in the middle of an environment. Further experiments are required to tease apart the separate influences of navigational relevance, amount of visual exposure and number of viewed perspectives in determining how objects and locations are encoded.

The current study found that people appear to construct multiple cognitive maps of environments consisting of two neighborhoods. Results suggested that a physical separator (e.g. fence) between the neighorhoods was not a necessary determinant of multiple-map encoding, and neither was a distinctively coloured/textured boundary around the neighborhoods. Categorical differences in store types contributed significantly to the multiple-map encoding, in the absence of any other differentiating features between the two neighborhoods. However, the current research did not investigate whether each of the variables separately could contribute to this effect. Future research should investigate which variable(s) could be the determinant(s) of multiple-map encoding, by mixing the types of stores in the two neighborhoods while leaving the buildings untouched, adding some buildings in the middle of the town to make it look like as whole while keeping different types of stores within each half of the environment, or using differently colored buildings or walls in each half of the environment while mixing everything else up.

It is well established that there are individual differences in terms of spatial abilities. For example, the current work showed that people who frequently used GPS devices were less accurate in terms of encoding multiple virtual environments. Moreover, previous research showed people who frequently engage in allocentric navigation tasks have structural brain differences: the posterior part of hippocampus in London taxi drivers was bigger and the anterior part of hippocampus was smaller, relative to hippocampi of bus drivers (Magurie, Woollett, & Spiers, 2006). This evidence leads us to hypothesize that people who rely on a GPS more often would have decreased posterior hippocampal volumes, and correspondingly poorer allocentric spatial abilities, relying more on egocentric representations, while non-GPS users may be more likely to build allocentric maps of the environment(s). Future research should look into this by asking some people to drive a real car with GPS and others without GPS for a period of time to see whether there would be any size differences in their hippocampi.

5 References

Abrahams, S., Pickering, A., Polkey, C. E., & Morris, R. G. (1997). Spatial memory deficits in patients with unilateral damage to the right hippocampal formation. *Neuropsychologia, 35(1),* 11-24.

Aguirre, G. K., & D'Esposito, M. (1997). Environmental knowledge is subserved by separable dorsal/ventral neural areas. *The Journal of Neuroscience, 17(7),* 2512-2518.

Alyan, S., & McNaughton, B. L. (1999). Hippocampectomized rats are capable of homing by path integration. *Behavioral Neuroscience, 113(1),* 1-19.

Alvernhe, A., Cauter, T. V., Save, E., & Poucet, B. (2008). Different CA1 and CA3 representations of novel routes in a shortcut situation. *The Journal of Neuroscience, 28(29),* 7324-7333.

Alvernhe, A., Save, E., & Poucet, B. (2011). Local remapping of place cell firing in the Tolman detour task. *European Journal of Neuroscience, 33,* 1696-1705.

Anderson, M. I., & Jeffery, K. J. (2003). Heterogeneous modulation of place cell firing by changes in context. *The Journal of Neuroscience, 23(26),* 8827-8835.

Barry, C., Hayman, R., Burgess, N., & Jeffery, K. J. (2007). Experience-dependent rescalling of entorhinal grids. *Nature Neuroscience, 10(6),* 682-684.

Becker, S. & Burgess, N. (2001). Modelling spatial recall, mental imagery and neglect. *In Advances in Neural Processing Systems (13, pp.96-102, pp.1065-1071).* Cambridge: MIT Press.

Benhamou, S. (1996). No evidence for cognitive mapping in rats. *Animal Behaviour*, *52,* 201-212.

Bennett, A. T. D. (1996). Do animals have cognitive maps? *The Journal of Experimental Biology, 199,* 219-224.

Blaisdell, A. P., & Cook, R. G. (2005). Integration of spatial maps in pigeons. *Animal Cognition, 8,* 7-16.

Boccara, C., Sargolini, F., Thoresen, V. H., Solstad, T., Witter, M. P., Moser, E. I., & Moser, M. (2010). Grid cells in pre- and parasubiculum. *Nature Neuroscience, 13(8),* 987-994.

Bohbot, V. D., Lerch, J., Thorndycraft, B., Iaria, G., & Zijdenbos, A. P. (2007). Gray matter differences correlate with spontaneous strategies in a human virtual navigation task. *Journal of Neuroscience, 27(38),* 10078-10083.

Booth, M.C.A. & Rolls, E.T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex, 8(6),* 510-523.

Brainard, D. H. (1997). The Psychophysics Toolbox. Spatial Vision, 10, 433-436.

Burgess, N. (2006). Spatial memory: How egocentric and allocentric combine. *Trends in Cognitive Sciences, 10(12),* 551-557.

Brockmole, J. R., & Wang, R. F. (2002). Switching between environmental representations in memory. *Cognition, 83,* 295-316.

Burgess, N., Jackson, A., Hartley, T., & O'Keefe, J. (2000). Predictions derived from modeling the hippocampal role in navigation. *Biological Cybernetics, 83,* 301-312.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory, *Neuron, 35,* 625-641.

Byrne, P., Becker, S. & Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychological Review, 114(2),* 340-375.

Cartwright, B. A., & Collett, T. S. (1982). How honey bees use landmarks to guide their return to a food source. *Nature, 295(5850),* 560-564.

Cartwright, B. A., & Collett, T. S. (1983). Landmark learning in bees-Experiments and models. *Journal of Comparative Physiology A, 151,* 521-543.

Cartwright, B. A., & Collett, T. S. (1987). Landmark maps for honeybees. *Biological Cybernetics, 57,* 85-93.

Chamizo, V. D., Rodrigo, T., & MacKintosh, N. J. (2006). Spatial integration with rats. *Learning & Behavior, 34(4),* 348-354.

Chapuis, N., Durup, M., & Thinus-Blanc, C. (1987). The role of exploratory experience in a shortcut task by golden hamsters (Mesocricetus auratus). *Animal Learning & Behavor, 15(2), 174-178.*

Chapuis, N., & Varlet, C. (1987). Short cuts by dogs in natural surroundings. *The Quarterly Journal of Experimental Psychology, 39B,* 49-64.

Chen, L. L., Lin, L. H., Green, E. J., Barnes, C. A., McNaughton, B. L. (1994). Head-direction cells in the rat posterior cortex. I. Anatomical distribution and behavioral modulation. *Experimental Brain Research, 101(1),* 8-23.

Cheng, K. (1988). Some psychophysics of the pigeon's use of landmarks. *Journal of Comparative Physiology A, 162,* 815-826.

Cheng, K., & Sherry, D. F. (1992). Landmark-based spatial memory in birds (Parus atricapillus and Columba livia): The use of edges and distances to represent spatial positions. *Journal of Comparative Psychology, 106(4),* 331-341.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28-71.

Collett, T. S., Cartwright, B. A., & Smith, B. A. (1986). Landmark learning and visuospatial memories in gerbils. *Journal of Comparative Physiology A: Sensory Neural and Behavioral Physiology, 158(6),* 835-851.

Collett, M., & Collett, T. S. (2006). Insect navigation: No map at the end of the trail? *Current Biology, 16(2),* 48-51.

Collett, T. S., & Land, M. F. (1975). Visual spatial memory in a hoverfly. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology, 100(1),* 59-84.

Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M. B., Moser, E. I. (2009). Fragmentation of grid cell maps in a multicompartment environment. *Nature Neuroscience, 12(10),* 1325-1334.

Doeller, C. *F*., & Burgess, N. (2008). Distinct error-correcting and incidental learning of location relative to landmarks and boundaries. *P.N.A.S., 105,* 5909-5914.

Dudchenko, P. A., & Zinyuk, L. E. (2005). The formation of cognitive maps of adjacent environments: Evidence from the head direction cell system. *Behavioral Neuroscience, 119(6),* 1511-1523.

Duelli, P., & Wehner, R. (1973). The spectral sensitivity of polarized light orientation in cataglyphis bicolor (Formicidae, Hymenoptera). *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology, 86(1),* 37-53.

Dyer, F. C. (1991). Bees acquire route-based memories but not cognitive maps in a familiar landscape. *Animal Behaviour, 41*, 239-246.

Dyer, F. C., & Gould, J. L. (1983). Honey Bee Navigation. *American Scientist, 71(6),* 587-597.

Ekstrom, A., Kahana, M., Caplan, J., Fields, T., Isham, E., Newman, E., et al. (2003). Cellular networks underlying human spatial navigation. *Nature, 425*, 184–187.

Epstein, R. A., Higgins, S., Thompson-Schill, S. L. (2005). Learning places from views: Variation in scene processing as a function of experience and navigational ability. *Journal of Cognitive Neuroscience, 17(1),* 73-83.

Etienne, A. S., & Jeffery, K. J. (2004). Path integration in mammals.

Hippocampus, 14(2). 180-192.

Etienne, A. S., Maurer, R., Berlie, J., Reverdin, B., Rowe, T., Georgakopoulos, J., & Séguinot, V. (1998). Navigation through vector addition. *Nature, 396,* 161-164.

Etienne, A. S., Maurer, R., Boulens, V., Levy, A., & Rowe, T. (2004). Resetting the path integrator: a basic condition for route-based navigation. *Journal of Experimental Biology, 207,* 1491-1508.

Etienne, A. S., Maurer, R., & Séguinot, V. (1996). Path integration in mammals and its interaction with visual landmarks. *The Journal of Experimental Biology, 199,* 201-209.

Etienne, A. S., Teroni, E., Hurni, C., & Portenier, V. (1990). The effect of a single light cue on homing behaviour of the golden hamster. *Animal Behaviour, 39,* 17-41.

Evans, G. W., & Pezdek, K. (1980). Cognitive Mapping: Knowledge of real-world distance and location information. *Journal of Experimental Psychology: Human Learning and Memory, 6(1),* 13-24.

Foo, P., Warren, W. H., Duchon, A., Terr, . (2005). Do humans integrate routes into a cognitive map? Map- versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31(2),* 195-215.

Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., & Moser, M. (2004). Spatial representation in the entorhinal cortex. *Science, 305,* 1258-1264.

Galati, G., Lobel, E., Vallar, G., Berthoz, A., Pizzamiglio, L., & Bihan, D. L. (2000). The neural basis of egocentric and allocentric coding of space in humans: a functional magnetic resonance study. *Experimental Brain Research, 133,* 156-164.

Gibson, B. M., & Kamil, A. C. (2001). Tests for cognitive mapping in Clark's nutcrackers. *Journal of Comparative Psychology, 115(4), 403-417.*

Giurfa, M., & Capaldi, E. A. (1999). Vectors, routes and maps: new discoveries about navigation in insects. *Trends in Neurosciences, 22(6),* 237-242.

Goodale, M., & Haffenden, A. (1998) Frames of reference for perception and action in the human visual system. *Neuroscience and Biobehavioural Reviews 22(2),* 161-172.

Goodale, M., & Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15,* 20–25.

Goodridge, J. P., & Taube, J. S. (1995). Preferential use of the landmark navigational system by head direction cells in rats. *Behavioral Neuroscience, 109(1),* 49-61.

Gould, J. L. (1986). The locale map of honey bees: Do insects have cognitive maps? *Science, 232(4752),* 861-863.

Gramann, K., Müller, H. J., Schönebeck, B. & Debus, G. (2006). The neural basis of ego- and allocentric reference frames in spatial navigation: Evidence from spatio-temporal coupled current density reconstruction. *Brain Research, 1118,* 116-129.

Greenauer, N., & Waller, D. (2010). Micro- and macroreference frames: Specifying the relations between spatial categories in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition. 36(4),* 938-957.

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., & Moser, E. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature, 436(7052),* 801–806.

Hartley, T., Burgess, N., Lever, C., Cacucci, F., O'Keefe, J. (2000). Modeling place fields in terms of the cortical inputs to the hippocampus. *Hippocampus, 10,* 369-379.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron, 37,* 877-888.

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence, 30,* 425-447.

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition, 13,* 208-217.

Hölscher, C. & Schmidt, W., J. (1994). Quinolinic acid lesion of the rat entorhinal cortex pars medialis produces selective amnesia in allocentric working memory (WM), but no in egocentric WM. *Behavioural Brain Research, 63,* 187-194.

Huttenlocher, J., Hedges, L. V., Corrigan, B., & Crawford, L., E. (2004). Spatial categories and the estimation of location. *Cognition, 93*, 75-97.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: prototype effects in estimating spatial location. *Psychological Review, 98(3),* 352-376.

Iaria, G., Petrides, M., Dagher, A., Pike, B., & Bohbot, V. D. (2003). Cognitive strategies dependent on hippocampus and caudate nucleus in human navigation: Variability and change with practice. *The Journal of Neuroscience, 23(13),* 5945-5952.

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology, 52,* 93-129.

Jacobs, J., Kahana, M. J., Ekstrom, A. D., Mollison, M. V., & Fried, I. (2010). A sense of direction in human entorhinal cortex. *PNAS, 107,* 6487-6492.

Janson, C. H. (2007). Experimental evidence for route integration and strategic planning in wild capuchin monkeys. *Animal Cognition, 10,* 341-356.

Janzen, G. & van Turennout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience, 7,* 673-677.

Janzen, G. & Weststeijn, C. G. (2007). Neural representation of object location and route direction: An event-related fMRI study. *Brain Research, 1165,* 116-125.

Jordan, K., Schadow, J., Wuestenberg, T., Heinze, H-J., & Jäncke, L. (2004). Different cortical activations for subjects using allocentric or egocentric strategies in a virtual navigation task. *Neuroreport, 15(1),* 135-140.

Kelly, J. W., & McNamara, T. P. (2008). Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact. *Psychonomic Bulletin & Review, 15(2),* 322-327.

King, J. A., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2002). Human hippocampus and viewpoint dependence in spatial memory. *Hippcampus, 12,* 811-820.

King, J. A., Trinkler, I., Hartley, T., Vargha-Khadem, *F*., & Burgess, N. (2004). The hippocampal role in spatial memory and the familiarity-recollection distinction: A case study. *Neuropsychology, 18(3),* 405-417.

Kjelstrup, K. B., Solstad, T., Brun, V. H., Hafting, T., Leutgeb, S., Witter, M. P., Moser, E. I., & Moser, M. (2008). Finite scale of spatial representation in the hippocampus. *Science, 321,* 140-143.

Klatzky, R. L., Loomis, J. M., Golledge, R. G., Cicinelli, J. G., Doherty, S., & Pellegrino, J. W. (1990). Acquisition of route and survey knowledge in the absence of vision. *The Journal of Motor Behavior, 22(1),* 19-43.

Kohler, M., & Wehner, R. (2005). Idiosyncratic route-based memories in desert ants, Melophorus bagoti: How do they interact with path-integration vectors? *Neurobiology of Learning and Memory, 83,* 1-12.

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience, 12,* 217-230.

Lavenex, P. B., & Lavenex, P. (2010). Spatial relational learning and memory abilities do not differ between men and women in a real-world, open-field environment. *Behaviorual Brain Research, 207,* 125-137.

Leonhard, C. M., Stackman, R. W. & Taube, J. S. (1996). Head direction cells recorded from the lateral mammillary nuclei in rats. *Social Neuroscience Abstract, 22,* 1873.

Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., Burgess, N. (2009). Boundary vector cells in the subiculum of the hippocampal formation. *The Journal of Neuroscience, 29(31),* 9771-9777.

Loftus, G. & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review. 1(4),* 476-490.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.)*, Wayfinding behavior: Cognitive mapping and other spatial processes,* (pp.125-151). Baltimore, MD: Johns Hopkins University Press.

Maaswinkel, H., Jarrard, L. E., & Whishaw, I. Q. (1999). Hippocampectomized rats are impaired in homing by path integration. *Hippocampus, 9(5),* 143-152.

Maaswinkel, H., & Whishaw, I. Q. (1999). Homing with locale, taxon, and dead reckoning strategies by foraging rats: Sensory hierarchy in spatial navigation. *Behavioural Brain Research, 99,* 143-152.

Maguire, E. A., Burke, T., Phillips, J., & Staunton, H. (1996). Topographical disorientation following unilateral temporal lobe lesions in humans, *Neuropsychologia, 34(10),* 993-1001.

Maguire, E. A., Frith, C. D., Burgess, N., Donnett J. G., & O'Keefe, J. (1998). Knowing where things are: Parahippocampal involvement in encoding object locations in virtual large-scale space. *Journal of Cognitive Neuroscience, 10(1),* 61-76.

Mather, J. G., & Baker, R. R. (1981). Magnetic sense of direction in woodmice for route-based navigation. *Nature, 291(14),* 152-155.

Matsumura, N., Nishijo, H., Tamura, R., Eifuku, S., Endo, S., & Ono, T. (1999). Spatial and task-dependent neuronal responses during real and virtual translocation in the monkey hippocampal formation. *The Journal of Neuroscience, 19(6)*, 2381–2393.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective.* (2nd ed.), Mahwah(NJ): Lawrence Erlbaum Associates, Inc.

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology, 18,* 87-121.

McNaughton, B. L., Barnes, C. A., Gerrard, J. L., Gothard, K., Jung, M. W., Knierim, J. J., Kudrimoti, H., Qin, Y., Skaggs, W. E., Suster, M., & Weaver, K. L. (1996). Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *The Journal of Experimental Biology, 199,* 173-185.

McNaughton, B. L., Barnes, C. A., O'Keefe, J. (1983). The contribution of position, direction and velocity to single unit activity in the hippocampus of freely-moving rats. *Experimental Brain Research, 52,* 41–49.

Meilinger, T. (2008). The network of reference frames theory: A synthesis of graphs and cognitive maps. In: *Spatial Cognition VI: Learning, Reasoning, and Talking about Space,* International Conference Spatial Cognition 2008, Springer, Berlin, Germany, 344-360.

Menzel, R., Geiger, K., Joerges, J., Müller, U., & Chittka, L. (1998). Bees travel novel homeward routes by integrating separately acquired vector memories. *Animal Behaviour, 55,* 139-152.

Menzel, R., Greggers, U., Smith, A., Berger, S., Brandt, R., Brunke, S., Bundrock, G., Hülse, S., Plümpe, T., Schaupp, F., Schüttler, E., Stach, S., Stindt, J., Stollhoff, N., & Watzl, S. (2005). Honey bees navigate according to a map-like spatial memory. *P.N.A.S, 102(8),* 3040-3045.

Miller, J., & Carlson, L. (2010). Selecting landmarks in novel environments. *Psychonomic Bulletin & Review*. Online Nov 17, 2010.

Mizumori, S. J. Y., & Williams, J. D. (1993). Directionally selective mnemonic properties of neurons in the lateral dorsal nucleus of the thalamus of rats. *The Journal of Neuroscience, 13(9),* 4012-4028.

Montello, D. R., & Pick, H. L., Jr. (1993). Integrating knowledge of vertically aligned large-scale spaces. *Environment and Behavior, 25(3),* 457-484.

Morgan, L. K., MacEvoy, S. P., Aguirre, G. K., Epstein, R. A. (2011). Distances between real-world locations are represented in the human hippocampus. *The Journal of Neuroscience, 31(4),* 1238-1245.

Moser, E. I., Kropff, E., & Moser, M. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annual Review of Neuroscience, 31,* 69-89.

Mou, W., Liu, X., & McNamara, T. P. (2009). Layout geometry in encoding and retrieval of spatial memory. *Journal of Experimental Psychology: Human Perception and Performance, 35(1),* 83-93.

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(1),* 162-170.

Mou, W., McNamara, T. P., Rump, B. & Xiao, C. (2006). Roles of egocentric and allocentric spatial representations in locomotion and reorientation*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32(6),* 1274-1290.

Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30(1),* 142-157.

Mou, W., Xiao, C., & McNamara, T., P. (2008). Reference directions and reference objects in spatial memory of a briefly viewed layout. *Cognition,* *108,* 136-154.

Mou, W., Zhao, M., & McNamara, T. P. (2007). Layout geometry in the selection of intrinsic frames of reference from multiple viewpoints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33(1),* 145-154.

Muller, R. U., Bostock, E., Taube, J., & Kubie, J. L. (1994). On the directional firing properties of hippocampal place cells. *The Journal of Neuroscience, 14(12),* 7235-7251.

Müller, M., & Wehner, R. (1988). Path integration in desert ants, *Cataglyphis fortis. Proceedings of the National Academy of Sciences, 85,* 5287-5290.

Müller, M., & Wehner, R. (2007). Wind and sky as compass cues in desert ant navigation. *Naturwissenschaften, 94(7),* 589-594.

Müller, R. U., (1996). A quarter century of place cells. *Neuron, 17,* 813-822.

Nesterova, A. P., & Hansen, F. (2009). Simple and integrated detours: Field tests with Columbian ground squirrels. *Animal Cognition, 12,* 655-670.

Normand, E., & Boesch, C. (2009). Sophisticated Euclidean maps in forest chimpanzees. *Animal Behaviour, 77,* 1195-1201.

Noser, R., & Byrne, R. W. (2007). Mental maps in chacma baboons (Papio ursinus): using inter-group encounters as a natural experiment. *Animal Cognition, 10,* 331-340.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental Neurology, 51,* 78-109.

O’Keefe, J., & Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons. *Nature, 381,* 425-428.

O’Keefe, J., & Dostrovsky, (1971). The hippocampus as a spatial map. *Brain Research , 34,* 171-175.

O’Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map.* Oxford: Oxford University Press.

O’Keefe, J., & Recce, M. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus, 3,* 317–330.

Ono, T., Nakamura, K., Nishijo, H., & Eifuku, S. (1993). Monkey hippocampal neurons related to spatial and nonspatial functions. *Journal of Neurophysiology, 70(4),* 1516-1529.

Paz-Villagrán, V., Save, E., & Poucet, B. (2004). Independent coding of connected environments by place cells. *European Journal of Neuroscience, 20,* 1379-1390.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437-442.

Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception, 23,* 1447-1455.

Quirk, G. J., Muller, R. U., & Kubie, J. L. (1990). The firing of hippocampal place cells in the dark depends on the rat’s recent experience. *The Journal of Neuroscience, 10(6),* 2008-2017.

Redish, A. D., McNaughton, B. L., & Barnes, C. A. (2000). Place cell firing shows an inertia-like process. *Neurocomputing, 32-33,* 235-241.

Ranck, J. B. Jr. (1984). Head direction cells in the dep layer of dorsal presubiculum in freely moving rats. *Social Neuroscience Abstract, 10,* 599.

Roberts, W. A., Cruz, C., & Tremblay, J. (2007). Rats take correct novel routes and shortcuts in an enclosed maze. *Journal of Experimental Psychology: Animal Behavior Processes, 33(2),* 79-91.

Rolls, E. T., & O'Mara, S. M. (1995). View-responsive neurons in the primate hippocampal complex. *Hippocampus, 5(5),* 409–424.

Roskos-Ewoldsen, B., McNamara, T. P., Shelton, A. L., & Carr, W. (1998). Mental representations of large and small spatial layouts are orientation dependent. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24(1),* 215-226.

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M. B., et al. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science, 312(5774),* 758–762.

Schenk, T. (2006). An allocentric rather than perceptual deficit in patient D.*F. Nature Neuroscience, 9(11),* 1369.

Schenk, F., Grobéty, M., & Gafner, M. (1997). Spatial learning by rats across visually disconnected environments. *The Quarterly Journal of Experimental Psychology, 50B(1),* 54-78.

Schinazi, V. R., & Epstein, R. A. (2010). Neural correlates of real-world route learning. *NeuroImage, 53,* 725-735.

Séguinot, V., Maurer, R., & Etienne, A. S. (1993). Dead reckoning in a small mammal: The evaluation of distance. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology, 173*, 103-113.

Shelton, A. L., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin & Review, 4(1),* 102-106.

Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology, 43,* 274-310.

Shelton, A. L., & McNamara, T. (2004). Orientation and perspective dependence in route and survey learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition. 30(1),* 158-170.

Smith, A. D., Gilchrist, I. D., Cater, K., Ikram, N., Nott, K., & Hood, B. M. (2008). Reorientation in the real world: The development of landmark use and integration in a natural environment. *Cognition, 107,* 1102-1111.

Stevens, A., & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10,* 422-437.

Sutherland, R. J., Kolb, R., & Whishaw, J. Q. (1982). Spatial mapping: Definitive disruption by hippocampal or medial frontal cortical damage in the rat. *Neuroscience Letters, 31,* 271-276.

Suzuki, S., Augerinos, G., & Black, A. H. (1980). Stimulus-control of spatial-behavor on the 8-arm maze in rats. *Learning and Motivation, 11(1),* 1-18.

Takahashi, N., Kawamura, M., Shiota, J., Kasahata, N., & Hirayama, K. (1997).

Pure topographic disorientation due to right retrosplenial lesion. *Neurology, 49(2),* 464-469.

Tanila, H. (1999). Hippocampal place cells can develop distinct representations of two visually identical environments. *Hippocampus,9,* 235-246.

Taube, J. S. (1995). Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *The Journal of Neuroscience, 15(1),* 70-86.

Taube, J. S., Muller, R. U., & Ranck, J. B. Jr. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *The Journal of Neuroscience, 10(2),* 420-435.

Teroni, E., Portenier, V., & Etienne, A. S. (1987). Spatial orientation of the golden-hamster in conditions of conflicting location-based and route-based information. *Behavioral Ecology and sociobiology, 20(6),* 389-397.

Tolman, E. C. (1948). Cognitive maps in rats and men. *The Psychological Review, 55(4),* 198-208.

Trullier, O., & Meyer, J. (2000). Animat navigation using a cognitive graph. *Biological Cybernetics, 83,* 271-285.

Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale & R.J.W. Mansfield (Eds.). *Analysis of visual behavior* (pp.549-586). Cambridge, MA: MIT Press.

Valiquette, C., & McNamara, T. P. (2007). Different mental representations for place recognition and goal localization. *Psychonomic Bulletin & Review, 14(4),* 676-680.

Valiquette, C. M., McNamara, T. P., & Labrecque, J. S. (2007). Biased representations of the spatial structure of navigable environments. *Psychological Research, 71,* 288-297.

Valiquette, C. M., & McNamara, T. P., & Smith, K., (2003). Locomotion, incidental learning, and the selection of spatial reference systems. *Memory & Cognition, 31(3),* 479-489.

Wang, R. *F*., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition, 77,* 215-250.

Werner, S., & Schmidt, K. (1999). Environmental reference systems for large-scale spaces. *Spatial Cognition and Computation, 1,* 447-473.

Wehner, R. (1972). Dorsoventral asymmetry in the visual field of the bee, Apis mellifica. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology, 77(3),* 256-277.

Wehner, R., Boyer, M., Loertscher, F., Sommer, S., Menzi, U. (2006). Ant navigation: One-way routes rather than maps. *Current Biology, 16,* 75-79.

Wehner, R., & Menzel, R. (1990). Do insects have cognitive maps? *Annual Review of Neuroscience, 13,* 403-414.

Wehner, R., & Müller, M. (2010). Piloting in desert ants: pinpointing the goal by discrete landmarks. *The Journal of Experimental Biology, 213,* 4174-4179.

Wehner, R., & Räber, F. (1979). Visual spatial memory in desert ants, Cataglyphis bicolor (Hymenopters: Formicidae). *Experientia, 35,* 1569-1571.

Wehner, R., & Srinivasan, M. V. (1981). Searching behaviour of desert ants, Genus Cataglyphis. *Journal of Comparative Physiology, 142*, 315-338.

Wehner, R., & Wehner, S. (1986). Path integration in desert ants-Approaching a long-standing puzzle in insect navigation. *Italian Journal of Zoology, 20(3),* 309-331.

Whishaw, I. Q., Hines, D. J., & Wallace, D. G. (2001). Dead reckoning (path integration) requires the hippocampal formation: Evidence from spontaneous exploration and spatial learning tasks in light (allothetic) and dark (idiothetic) tests. *Behavioural Brain Research, 127,* 49-69.

Whishaw, I. Q., & Maaswinkel, H. (1998). Rats with fimbria-fornix lesions are impaired in path integration: A role for the hippocampus in “Sense of Direction”. *The Journal of Neuroscience, 18(8),* 143-152.

Wolbers, T., Hegarty, M., Buchel, C., & Loomis, J. M. (2008). Spatial updating: how the brain keeps track of changing object locations during observer motion. *Nature Neuroscience, 11(10),* 1223-1230.

Worden, R. (1992). Navigation by fragment fitting: A theory of hippocampal function. *Hippocampus, 2(2),* 165-188.

Wystrach, A., Schwarz, S., Schultheiss, P., Beugnon, G., & Cheng, K. (2011). View, landmarks, and routes: How do desert ants negotiate an obstacle course? *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology, 197,* 167-179.