

Predicting One's Memory and Remembering Landmarks on a Map

A dissertation submitted by

Lauren A. Mason

in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Psychology

Tufts University

February 2025

© 2025, Lauren A. Mason

Advisors: Holly A. Taylor; Ayanna K. Thomas

Abstract

Map displays are rich with information. They contain landmark names that hold semantic meaning, such as ‘coffee shop’ referring to a place one can purchase a caffeinated beverage. One can also identify the coffee shop’s location in reference to surrounding landmarks, otherwise referred to as relative landmark locations. Anecdotal evidence suggests that as we interact with maps, we engage in metacognitive judgment formation. For instance, when we reflect on our likelihood to remember where a coffee shop is located relative to our hotel, we are essentially making a metacognitive judgment of learning (JOL). Research highlights various factors that influence map memory. Yet, how learners predict what they will successfully remember from maps remains a distinct and understudied question. Recent work explored cues that contributed to metacognitive predictions of map memory and found that map complexity and information accumulation across the learning span both predicted JOL rating magnitude (Mason et al., 2024). In the present study, I expanded on this finding by explicitly manipulating landmark complexity as defined by the number of landmarks present on the map (Experiment 1) and two other display features: perceptual landmark salience, defined by the visual distinctiveness of the goal-relevant landmarks (Experiment 2) and semantic accessibility of landmark name meaning, operationalized by map language (Experiment 3). Learners studied landmark pairs from maps, made JOLs about their likelihood to remember relative landmark locations, and were tested on landmark location memory. Landmark complexity and salience impacted JOLs, and access to landmark name meaning benefitted memory. Components such as self-rated spatial ability and time across the learning span also contributed to JOLs. These results collectively motivate the need for an integrative cue-utility model to demonstrate the dynamic information learners consider when forming metacognitive monitoring judgments during map-based learning.

To every educator who helped fill the gaps in my life,
and to my dad, who spent over 40 years navigating American highways.

TABLE OF CONTENTS

List Of Tables	5
List Of Figures	6
Predicting One's Memory And Remembering Landmarks On A Map	7
Predicting One's Memory For Map Components, And Why It Matters	8
Map Complexity	11
Perceptual Salience	15
Semantic Accessibility	17
Individual Differences: Sense Of Direction And Gender	20
Summary Of Purpose	21
Experiment 1	22
Method	22
Results	33
Discussion	46
Experiment 2	48
Method	49
Results	52
Discussion	64
Experiment 3	66
Method	67
Results	70
Discussion	81
General Discussion	82
References	91
Appendix A.	101
Appendix B.	102
Appendix C.	103
Appendix D.	104
Appendix E.	105

LIST OF TABLES

Table 1 23

Table 2 39

Table 3 45

Table 4 58

LIST OF FIGURES

Figure 1.	22
Figure 2.	27
Figure 3.	36
Figure 4.	40
Figure 5.	43
Figure 6.	46
Figure 7.	49
Figure 8.	54
Figure 9.	56
Figure 10.	59
Figure 11.	62
Figure 12.	64
Figure 13.	67
Figure 14.	73
Figure 15.	74
Figure 16.	76
Figure 17.	77
Figure 18.	78
Figure 19.	79

Predicting One's Memory and Remembering Landmarks on a Map

Imagine being in a new city. You pull up a map on your phone to see if there is a coffee shop located between your hotel and the museum you are going to later. Here, you are looking at the relative landmark locations, or where a landmark is in reference to other landmarks. In addition to the hotel, coffee shop, and museum, there are many other landmarks scattered across the map, such as a university, bakery, and river. These landmarks offer further locational context to help us orient ourselves in an environment (Caduff & Timpf, 2008; Cartwright et al., 1999; Keil et al., 2020; Klippel, 2003; Michon & Denis, 2001). For instance, you might say that the coffee shop is close to the river or north of the bakery. These landmark names also carry semantic information, or meaning (Hirtle & Mascolo, 1986; Merrill & Baird, 1987; Thomas et al., 2012; Wessel et al., 2010). When we see 'coffee shop' on the map, we immediately know that we can purchase a caffeinated beverage there. In sum, map displays are information-rich (Bestgen et al., 2017; Bunch & Lloyd, 2006; Keil et al., 2020), and anecdotal evidence suggests that as we interact with this information, we engage in metacognitive judgment formation. When we reflect on our likelihood to remember where the coffee shop is relative to our hotel tomorrow morning, we are essentially making a metacognitive monitoring judgment, also referred to as a judgment of learning (JOL) (Hirtle & Mascolo, 1986; Merrill & Baird, 1987; Thomas et al., 2012; Wessel et al., 2010). Though research highlights components that contribute to map memory, it is not understood how map features influence one's prediction of their own memory during map learning. Elucidating this metacognitive process is important because accurate memory predictions can prevent people from becoming spatially disoriented or lost by prompting metacognitive control processes such as restudying a map, recreating it in one's mind, or encouraging the utilization of a navigation aid during travel.

The aims of the present work were threefold: First, I assessed whether metacognitive judgments of learning (JOLs), or one's prediction of their memory, were sensitive to three map display features: landmark complexity, as defined by the number of landmarks present on the map (Experiment 1), perceptual landmark salience, defined by the visual distinctiveness of the goal-relevant landmarks (Experiment 2), and semantic accessibility of landmark name meaning, operationalized by map language (Experiment 3). Here, *JOL sensitivity* addresses whether JOL ratings varied as a function of map display features. Next, I assessed whether variations in these map features impacted memory. Lastly, I examined the relationship between JOLs and memory to determine whether JOLs were predictive of memory and whether predictive accuracy varied as a function of landmark complexity, perceptual landmark salience, and semantic accessibility, or knowledge of landmark name meaning.

Predicting one's memory for map components, and why it matters

Little is known about the cues people refer to when deciding whether they'll remember map information, specifically. Prior verbal learning research with word pairs offers insight into the metacognitive judgment formation process. Namely, an eminent theory suggests that learners do not have direct access to the memory strength formed at study (Koriat, 1997). Rather, learners collectively weigh information from several cues when forming memory predictions, such as characteristics of the items themselves, the level of cognitive processing, and the ease at which the items are learned (Koriat, 1997; Mason et al., 2022). Learners may also arrive with pre-existing beliefs or theories about their own task-relevant abilities (Koriat & Bjork, 2006; Mason et al., 2022). In other words, JOLs are sensitive to various factors during word-pair learning. Given that maps are comprised of several components, such as landmark locations, names, and routes between landmarks, to name a few, there are many map elements to which JOLs may be

sensitive. Therefore, the present study investigated whether metacognitive judgments were sensitive to three different variations in map display features.

The first display feature assessed in the present study was landmark complexity. Exploratory findings from recent research suggested that map complexity contributed to metacognitive judgment formation (Mason et al., 2024). Learners studied landmark pairs on low and high complexity maps, operationalized by the number of turns in routes between landmarks. Then, they rated their likelihood to remember map components and were tested. Findings suggested that learners rated themselves more likely to remember less complex maps. Experiment 1 of the present study expanded on this finding by investigating the role of map complexity, now defined by the number of landmarks present on the map (Cheng et al., 2022), for JOL rating magnitude, memory, and metacognitive accuracy.

Next, given that it is difficult to determine the cues that learners attend to while forming metacognitive monitoring judgments during spatial learning (Mason et al., 2022), and judgment accuracy depends upon referring to cues that are relevant later at test (Dunlosky et al., 2005; Koriat, 1997), the present study also employed a salience manipulation to guide learners' attention to goal-relevant information during learning and subsequent metacognitive judgment formation. Specifically, Experiment 2 investigated the role of perceptual landmark salience for JOL magnitude, memory, and metacognitive accuracy. In half of the maps, the goal-relevant information (i.e., the critical landmark pair) was made perceptually salient by altering the transparency of goal-relevant map elements.

Lastly, preliminary evidence (Mason et al., in progress) suggested that learners referred to the semantic association of landmark names when deciding whether they would remember map information in the future. For instance, some learners reported trying to remember

landmarks from maps based on their personal association or experience with them, such as ‘bus stop’ being a place they go often. This strategy reportedly helped them remember other nearby landmark locations from the map. However, it is unknown whether knowledge of landmark name meaning during map learning explicitly contributes to metacognitive judgment formation or memory for landmark locations. Therefore, Experiment 3 assessed whether JOLs were sensitive to semantic accessibility, or knowledge of landmark name meaning, and whether such knowledge aided 1) memory for landmark locations, and 2) one’s ability to accurately predict their future memory. In the final experiment, learners studied landmark pairs on maps in English, or in a language foreign to them, such that they were unaware of landmark name meaning.

In the present work, a positive relationship between one’s JOL and later memory indicates that the metacognitive judgment is generally accurate, or predictive of memory. In verbal learning contexts, the judgments one ultimately forms are generally accurate as long as 1) the cues used at the time of the judgment are consistent with factors that affect subsequent test performance, and 2) the cues used carry predictive validity (Dunlosky et al., 2005; Koriat, 1997). However, because maps are visually complex, there are many cues learners can attend to or fail to attend to and encode, and some cues may be more predictive of later recall than others (Mason et al., 2022). Additionally, little is known about which map features may be more or less predictive of memory.

Predictive accuracy has ecological relevance. In the real world, if we predict that we are unlikely to remember the relative locations of two landmarks in the future, we can metacognitively control learning, such as by restudying the map before navigating (Mason et al., 2022), or even reproducing the map from memory (Brunyé et al., 2020). In other words, metacognition consists of monitoring and control, and these components interact to support

learning (Mason et al., 2022; Nelson & Narens, 1994; Nelson, 1990). Restudying the map, as an example, may decrease one's chances of becoming spatially disoriented. In contrast, if we inaccurately predict that we are likely to remember the relative landmark locations but later cannot, we may become lost. Therefore, the present study also investigated how landmark complexity, salience, and knowledge of name meaning influenced memory and metacognitive accuracy, or accurate predictions of future memory.

Map Complexity

Experiment 1 investigated the role of landmark complexity for JOL magnitude, memory, and metacognitive accuracy. Landmark complexity was defined by the number of landmarks on the map, as motivated by recent work that found the number of landmarks on maps can interact with users' cognitive capacity and impact spatial learning (Cheng et al., 2022). This complexity categorization was quantifiably confirmed via JPEG file sizes, an established approach to defining the complexity of visual displays in cognitive psychology (Bestgen et al., 2017; Donderi & McFadden, 2005; Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014). Learners studied landmark pairs on low and high complexity maps.

Predicting one's memory for complex maps. The first outcome of interest in this study was JOL magnitude, or how likely learners deemed themselves to remember map information in the future. Map complexity could influence subjective ease of encoding. In one study (Rhodes & Castel, 2008), participants made JOLs after studying words printed in small and large fonts. Those printed in a large font size produced a higher JOL magnitude as compared to small font-sized words, meaning people rated themselves more likely to remember the words in larger font. One explanation for this finding is that learners subjectively processed large words with more ease and therefore assigned them a higher JOL (Miller & Geraci, 2016; Rhodes & Castel, 2008).

Actual memory did not differ between word sizes. These findings suggested that stimulus characteristics served as contributory cues during JOL formation, but that these cues do not always reliably predict later memory. Given that maps with more visual information (i.e., visual clutter) threaten perceptual ease by potentially resulting in extraneous cognitive load and limiting available resources to effectively process goal-relevant map information (Dobson, 1980; Klippel et al., 2006; Mayer, 2001; Phillips & Noyes, 1977; van Merriënboer et al., 2002), learners in the present study could assign higher ratings to landmark pairs studied on less complex, and therefore more easily processed maps.

In line with verbal learning findings, map learning research also found that JOLs were sensitive to perceptual aspects of stimuli. In one study (Stevens & Carlson, 2019), participants learned from low detail (schematic) or high-detail (perceptually rich, satellite) maps, made JOLs, and performed a spatial pointing task where the deviation between their imagined viewpoint and the map north was manipulated. Learners provided higher JOLs for high-detail maps, suggesting that judgments were sensitive to the perceptual detail of maps. Further, only those in the high-detail condition produced JOLs that accurately predicted their later pointing performance. Based on these findings, JOLs for map materials were sensitive to visual display features, revealed a preference for high-detail maps, and appeared to be based on diagnostic cues, or cues that are predictive of later memory (namely, variations in perceptual detail). This finding suggests learners may assign higher JOLs to high complexity maps.

Additional research on geospatial display design also determined that learners' JOLs indicated a preference for geographically complex displays, but more complex displays produced slower, less accurate comprehension compared to simpler maps (Hegarty et al., 2012). Taken together, JOLs across various studies were sensitive to perceptual aspects of visual displays, with

two cases revealing higher JOLs for high-detail, complex visual displays. Yet, the predictive value of visual display cues varied. Whereas variations in map detail may serve as an informative indicator of spatial memory (Stevens & Carlson, 2019), a preference for more complex displays does not always translate to optimal comprehension (Gardony et al., 2022; Hegarty et al., 2012).

In partial contrast to the aforementioned findings, recent research again found that JOLs were sensitive to variations in map complexity, but instead suggested a preference for less rather than more complex maps (Mason et al., 2024). Participants studied, made JOLs, and were tested on their memory for routes or relative landmark locations. Across the board, learners assigned higher JOL magnitudes to less complex maps. It is possible that these findings contrasted with prior work due to variations in the methodology, such as the fact that learners in this study (Mason et al., 2024) were instructed to learn routes or relative landmark locations and completed a two alternative forced choice test where they had to select the map containing the correct landmark locations or route that they had seen during learning. In contrast, the work from Stevens and Carlson (2019) had learners perform a pointing task as a measure of spatial knowledge. In sum, high-detail maps may produce higher JOLs under some circumstances, but not all.

Results from prior research produced competing JOL magnitude hypotheses. First, learners in the present study could assign higher JOLs to pairs studied on low complexity maps (Mason et al., 2024), given that visual clutter might impair perceptual ease (Dobson, 1980; Klippel et al., 2006; Phillips & Noyes, 1977). Alternatively, learners could assign higher JOLs to pairs studied on high complexity maps, because high-detail displays could be perceived as more accurate or informative (Hegarty et al., 2012; Stevens & Carlson, 2019).

Remembering complex maps. The second outcome of interest in the present study was whether memory varied as a function of complexity. Map complexity contributes to memory success (Bestgen et al., 2017; Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014; Keil et al., 2020; Mason et al., 2024). In prior work, object location memory improved with increasing map complexity, such that adding grid elements (e.g., continuous lines, dashed lines, crosses), density (e.g., lines, areas, and points) and urban landscape helped learners group information into a more hierarchical organization and improved location memory (Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014; Keil et al., 2020). These outcomes were likely due to a greater number of reference points supporting a mental representation of the map, also referred to as cognitive map (Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014; Keil et al., 2020; Montello, 2002; Siegel & White, 1975; Tversky, 1993). However, too many map components can prove problematic. The relationship between map complexity and memory performance is not linear, but rather is characterized by an inverted U-shaped function. In other words, memory performance is best with intermediate complexity, but suffers with both low and high complexity (Bestgen et al., 2017; Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014). This assumption aligns with claims that learning from complex materials, such as multimedia displays, can result in extraneous cognitive load, limit the range of attentional resources available, and impair encoding (Bunch & Lloyd, 2006; Caduff & Timpf, 2008; Garden et al., 2002; Kirschner et al., 2006; Knight & Tlauka, 2017; Mayer, 2001; Scott & Schwartz, 2007; Valcke, 2002; van Merriënboer et al., 2002).

Mixed findings on the role of complexity for map memory again produced competing memory hypotheses in the present experiment. Namely, learners could better remember pairs studied on low complexity maps, since visually complex geographic displays could impede

comprehension (Hegarty et al., 2012). Conversely, more complex maps could provide a greater number of reference points, support the spatial representation, and subsequently bolster memory (Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014; Keil et al., 2020; Mason et al., 2024; Stevens & Carlson, 2019).

JOL-memory relationship. The third outcome of interest was metacognitive accuracy, or the relationship between JOL predictions and memory. Two possibilities were considered by which JOLs were sensitive to variations in complexity, and a positive relationship between predictions and performance could be observed. First, if fewer landmarks on a map benefitted perceptual and encoding ease (Bjork, 1994), memory for less complex maps might be advantaged, and a positive relationship between predictions and performance might be observed. In contrast, if more complex maps were perceived as more informative and offered additional reference points to support spatial mental representation development, then greater memory and a positive JOL-memory relationship might be observed for pairs studied on high complexity maps. These considerations aside, it was also possible that JOLs would be sensitive to complexity, but memory did not vary, suggesting that complexity was a poor indicator to base one's memory predictions on.

Perceptual Salience

Aside from landmark complexity, an alternative approach to altering the perceptual experience while learning from maps is by modifying the visual distinctiveness of goal-relevant map regions (Britton et al., 1979; Hegarty et al., 2010; LaBerge, 1995; Mason et al., 2024; Pichert & Anderson, 1977; Taylor et al., 1999). In Experiment 2, I considered how perceptual salience influenced JOL magnitude, memory, and metacognitive accuracy. Here, perceptual salience was defined as the extent to which goal-relevant map regions (i.e., regions containing

critical landmark pairs) were made visually distinct as compared to the rest of the map. Learners studied landmark pairs on perceptually salient and standard maps, made JOLs, and were tested.

Predicting one's memory for perceptually salient maps. As previously described, learners may be sensitive to perceptual aspects of stimuli (Klippel et al., 2006; Phillips & Noyes, 1977, 1982; Schwarz, 2018), and such sensitivity should be reflected by JOLs (Hegarty et al., 2012; Miller & Geraci, 2016; Stevens & Carlson, 2019). If emphasizing the goal-relevant map information (namely, the critical landmark pair), readily directed attention to the critical landmark pair and thereby promoted perceptual ease, then higher JOLs could be observed for landmark pairs studied on perceptually salient maps. Based on findings from the literature on map design and navigation (Caduff & Timpf, 2008; Hegarty et al., 2010; Keil et al., 2020; Yeh & Wickens, 2001), I hypothesized that learners would assign higher JOLs to pairs studied on perceptually salient maps.

Remembering perceptually salient maps. In addition to their role in judgment formation, contrast and perceptual distinction also contribute to map memory. Memory for map components depends upon attentional processing, and attentional processing is influenced by salience (Keil et al., 2020). For a map attribute to be perceptually salient, it must contrast with the background environment somehow, such that it is perceptually distinct (Caduff & Timpf, 2008; Itti & Koch, 2001). In prior work, variations in intensity differences, or foreground-background color contrast, made map display variables more or less visually distinct (Yeh & Wickens, 2001). This visual distinctness captured attention, helping map users locate and focus on the salient information (Yeh & Wickens, 2001). In sum, more perceptually salient environment features likely capture and direct attention toward those features in a bottom-up manner (Caduff & Timpf, 2008; Hegarty et al., 2010). Therefore, if goal-relevant map

information is made perceptually salient, attention might be preferably directed toward this information to facilitate task-relevant processing and benefit memory (Hegarty et al., 2010). In sum, I hypothesized that greater memory would be observed for landmarks studied on perceptually salient as compared to standard maps.

JOL-memory relationship. If JOLs were sensitive to perceptual salience, such that higher JOLs were assigned to landmark pairs studied on perceptually salient as compared to standard maps, and if perceptually salient maps benefitted landmark memory, then a positive JOL-memory relationship would be observed for perceptually salient maps. Alternatively, if learners assigned lower JOLs to standard maps and had worse memory for standard maps, then a positive JOL-memory relationship might also be observed for standard maps. Notably however, if perceptually salient maps guided learners to better-attend to goal-relevant, diagnostic cues, then a stronger JOL-memory relationship might be observed for pairs on perceptually salient maps, specifically. Said differently, making JOLs for standard maps arguably lends to a less guided encoding experience as compared to that of perceptually salient maps. This experience could result in learners attending to less diagnostic cues during judgment formation for standard maps. Consequently, the JOL-memory relationship might be weaker for pairs studied on standard maps, specifically.

Semantic Accessibility

I discussed the proposed role of landmark complexity (Experiment 1) and landmark salience (Experiment 2) for JOL magnitude, memory, and metacognitive accuracy. In Experiment 3, I investigated semantic accessibility, or knowledge of the landmark name meaning. For instance, when we see ‘coffee shop’ on a map, we immediately know it refers to a place where we can purchase a caffeinated beverage. In contrast, when we travel to a distant

place and view a map with landmark names written in a language foreign to us, we no longer have access to landmark name meaning.

Predicting one's memory for English and foreign language maps. Research suggests that familiar material is easier to process than novel material (Schwarz, 2018), and ease of processing contributes to metacognitive judgment formation (Begg et al., 1989; Koriat, 1997; Miller & Geraci, 2016; Rhodes & Castel, 2008). Notably, in Experiment 3 of the present study, landmark names presented in a language foreign to learners were not only novel, but they were meaningfully uninterpretable. Based on this notion, it was assumed that rating magnitude would vary as a function of semantic accessibility (namely, studying landmark names in English or in a language foreign to the learner). Specifically, I hypothesized that learners who studied maps in English would assign higher JOLs to landmark pairs as compared to those who studied maps in a language foreign to them. A potential alternative hypothesis was that differences in JOL magnitude would not be observed because learners in this between-participants experiment studied pairs in English or in a language foreign to them. Therefore, they could not compare their processing experience between stimulus types within the experiment itself. If null results were observed for JOL magnitudes, they could be attributed to the between-subjects design employed in this experiment (Dechêne et al., 2010; Schwarz, 2018).

Remembering English and foreign language maps. Next, I assessed how semantic accessibility affected map memory. Findings from the perceptual sciences suggest navigators process this type of information through a specialized “what” pathway dedicated to object recognition and semantic understanding, whereas the “where” pathway responds to spatial aspects of stimuli, such as direction (Ungerleider, 1994; Wessel et al., 2010). Having a strong spatial understanding of an environment can provide a solid structure for organizing semantic

components (Wessel et al., 2010). In one study (Wessel et al., 2010), participants navigated an urban environment, drew a sketch map, recalled semantic elements (i.e., objects) by name, and then placed the names on the map. The results suggested that a better spatial understanding of the environment aided learners' placement of semantic components. Additionally, in prior work, landmarks from within-cluster pairs, such as cities within states (i.e., within a semantic category), were rated as being closer together than equally distant between-cluster pairs (i.e., across semantic categories) (Allen, 1981; Hirtle & Mascolo, 1986; Thomas et al., 2012). Taken together, these findings suggest that spatial and semantic information interact to form environment representations.

Whether access to semantic information facilitated spatial learning was a separate question relevant to the present study. Prior work investigating the relationship between object identity memory and spatial locations found that landmark buildings recognized by name were relocated more accurately in space (Pezdek & Evans, 1979). Participants studied a city model containing 16 buildings. The buildings were either represented as wooden blocks with a color photograph of a real building, photographs with a printed label, or only verbal labels. Spatial memory was better for items recognized by name as compared to those not recognized by name. In sum, processing semantic information associated with the landmark supported memory for its location. However, the conditions in their study either contained landmark labels or they did not. To my knowledge, what remained unknown was whether meaningful semantic labels (i.e., known landmark names) support location memory when compared to landmark labels that do not carry meaning to learners (i.e., names in a language foreign to them). This question is particularly relevant to people navigating in new countries with different written languages, yet this discussion is notably neglected throughout the literature. In Experiment 3, I investigated

whether maps with landmark names in English were better remembered as compared to those in Estonian, because learners were aware of the associative meaning of the name. Here, “semantic accessibility” referred to maps with landmark names in the learners’ native language. I compared this to a condition where learners studied maps in a language foreign to them. In this case, the locations still had a written label, but learners did not have access to the semantic meaning associated with the landmark. I hypothesized that maps that leveraged semantic accessibility by presenting landmark names in English would lend themselves to greater memory for landmark locations.

JOL-memory relationship. Learners in the English condition had access to additional, meaningful information (i.e., known landmark names) during judgment formation than learners who studied maps in a language foreign to them. Such knowledge could result in greater processing ease and support encoding. If so, higher JOL magnitudes (Begg et al., 1989; Koriat, 1997; Miller & Geraci, 2016; Rhodes & Castel, 2008) and memory (Pezdek & Evans, 1979) would be observed, lending to a positive JOL-memory relationship in the English condition. It was also possible that learners without semantically meaningful info would perceive the task as more difficult and therefore produce lower JOLs. If the task was in fact more difficult, such that lower JOLs and lower memory scores were observed, then a positive JOL-memory relationship could also be found in the Estonian map condition. Finally, it was possible that differences in JOL magnitudes and/or memory would not be found, since learners could not compare their perceptual experience across stimuli types, and the stimuli and paradigm employed were novel.

Individual Differences: Sense of Direction and Gender

In addition to the primary variables of interest, I also conducted exploratory analyses across all three experiments to consider the role of two individual difference variables: self-

reported sense of direction (SOD) ability and gender. Individuals differ significantly in their ability to perform spatial tasks (Weisberg & Newcombe, 2016), and self-reported measures can be a promising way to predict spatial performance outcomes (Hegarty et al., 2002; Kozlowski & Bryant, 1977). Given that a motivation of the present work was to investigate JOL rating and memory predictors, I considered the relationship between learners' self-reported SOD scores, JOL rating magnitudes, and memory. Prior research also found that women were less confident in spatial self-assessments, but just as capable at accurately monitoring their performance as men (Ariel et al., 2018). Therefore, I also considered the role of gender for memory predictions and spatial performance.

Summary of Purpose

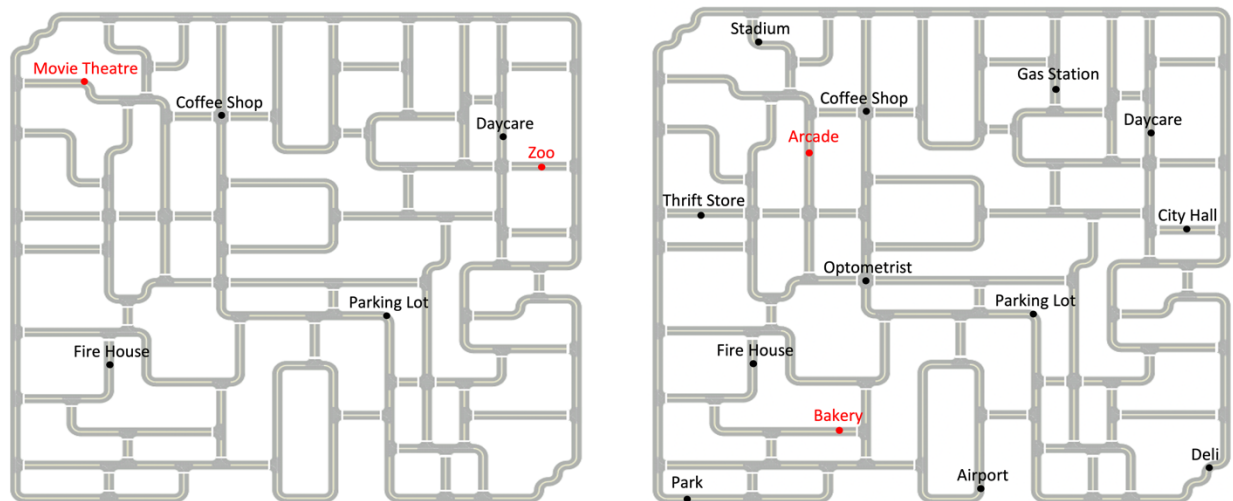
Research highlights various factors that influence map memory, but how learners predict what they will successfully remember from maps remains a distinct and understudied question. In the present study, I assessed how variations in map features influenced JOL rating magnitude, map memory, and the relationship between JOLs and memory. In Experiment 1, I manipulated landmark complexity, defined by the number of landmarks present on the map. In Experiment 2, I manipulated perceptual landmark salience, defined by the visual distinctiveness of the goal-relevant landmarks. In Experiment 3, I manipulated the semantic accessibility of landmark name meaning, operationalized by map display language. Learners studied maps containing a critical landmark pair and additional context landmarks, made JOLs about their likelihood to remember relative landmark locations, and were tested on their landmark location memory. Findings from this work offer insights on the role of map features for spatial memory predictions and performance.

Experiment 1

In Experiment 1, I investigated the role of landmark complexity for JOL rating magnitude, memory, and metacognitive accuracy. Learners were instructed to learn the relative locations of two critical landmarks. The number of additional landmarks on the map was either low (4) or high (12). After studying each map, learners provided a JOL about their likelihood to remember the relative landmark locations on a later test. After studying all maps, their memory for the relative landmark locations was tested. An example of the complexity manipulation is provided (Figure 1).

Figure 1.

Examples of a Low (Left) and High Complexity (Right) Learning Map in Experiment 1



Method

Participants

This experiment implemented novel spatial stimuli in a novel paradigm loosely modeled after one previously employed (Mason et al., 2024). Due to the exploratory nature of this work, the

desired sample sizes for the present study were determined by holistic consideration of previous work from the verbal, spatial, and metacognitive literatures, broadly. A summary of this consideration is provided in Table 1.

Table 1

Participant Number Determination: Overview of Literature Considered

Reference	Keywords	Design(s)	Stimuli	Participants
Bestgen et al. (2017)	Cartography, spatial memory, complexity	(2) Two-factorial, 3x2 mixed (1) Two-factorial, 3x2 within-subjects	12-15 maps	40, 46, 32
Pezdek et al. (1979)	Visual memory, verbal memory, spatial location	One-way independent group	Scaled graphic city map	45
Hegarty et al. (2012)	Metacognition, visual display, complexity	3x2x3 within-subjects	3 weather maps	38, 26, 13
Hegarty et al. (2010)	Graphics comprehension, salience	(2) 2x3x3 mixed (1) 2x2x3x3	10 maps	40, 40, 80
Stevens et al. (2019)	Spatial cognition, metacognition	2x5x2 mixed	2 maps	60
Koriat (1997)	JOLs, cue-utilization	Between-subjects	70-100 word-pairs	32, 24, 24, 30
Rhodes et al. (2008)	Memory, metacognition, JOLs	Within-subjects	42-84 words	20, 28, 28, 54
Mason et al. (2023)	Metacognition, spatial memory, map learning	(2) 2x2 between-subjects	16 maps	102, 94

From the experiments outlined, the total number of participants ranged from 13 at the low end to 102 at the high end. In collective consideration of the aforementioned participant numbers and designs, in addition to the 11-15% data loss rate observed in prior work (Mason et al., 2024), the present study aimed to analyze data from 52 participants in Experiments 1 and 2 (within-subjects) and 72 participants in Experiment 3 (between-subjects).

In Experiment 1, 66 participants ages 18-35 were recruited through Prolific, an online participant recruitment platform. All participants were within the United States and were required to use a computer to complete the experiment. They were compensated with the standard Prolific rate. Data from two participants was excluded from analyses due to technical difficulties. An additional 12 participants were excluded due to redundant JOL responses, such that they only responded with one to two numbers on the scale or responded with the same number for nine or more consecutive trials. From the remaining 52 participants, 56% identified as women and 44% identified as men. The sample was majority White (62%) and non-Hispanic or Latino (90%). Most participants completed education beyond high school (75%).

Design

Map complexity was a within-participant variable. All participants studied half of the landmark pairs on low complexity maps and half on high complexity maps.

Materials

Landmarks. There were 52 total landmarks in the study: 20 landmark pairs and 12 additional context landmarks.

Landmark pairs. There were two lists of 10 landmark pairs (20 total landmark pairs, Appendix A). Landmark names were general, based on function, and inspired by those typically found in cities, such as ‘Movie Theatre’ and ‘Grocery Store’. Each pair consisted first of a “cue”

landmark, and next a “target” landmark. Landmark name pairings and assignments as cues versus targets were determined randomly but held consistent throughout the study.

Context landmarks. In addition to the 20 landmark pairs, there were 12 additional context landmarks present on the maps that were never part of a critical pair. On low complexity maps, only four additional context landmarks were shown. The four shown remained constant throughout the experiment. On high complexity maps, the same four context landmarks were shown along with eight more context landmarks.

Maps. The first author developed the map stimuli by adding landmarks to a modified road network image from a virtual reality (VR) environment similar to one employed in prior research (Brunyé et al., 2018). Maps were approximately 5 in x 5.5 in. Each map stimulus contained two critical landmarks labeled in red font and context landmarks labeled in black font. During the practice phase, the maps contained two critical landmarks and either two (low complexity practice map) or six (high complexity practice map) additional context landmarks. The context landmarks never appeared in a critical pair.

During the learning phase of the experiment, half of the participants studied 10 landmark pairs from list (A) on low complexity maps and 10 from list (B) on high complexity maps, while the other half of participants studied list (A) on high complexity maps and list (B) on low complexity maps. The order of map presentation was randomized, such that it was possible for a participant to study a low complexity map, followed by a high complexity map, and so on. On average, learners saw between 1-3 maps of the same type consecutively. Map complexity was confirmed by JPEG file sizes (Bestgen et al., 2017; Donderi & McFadden, 2005; Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014). The average file size was 142KB for low complexity

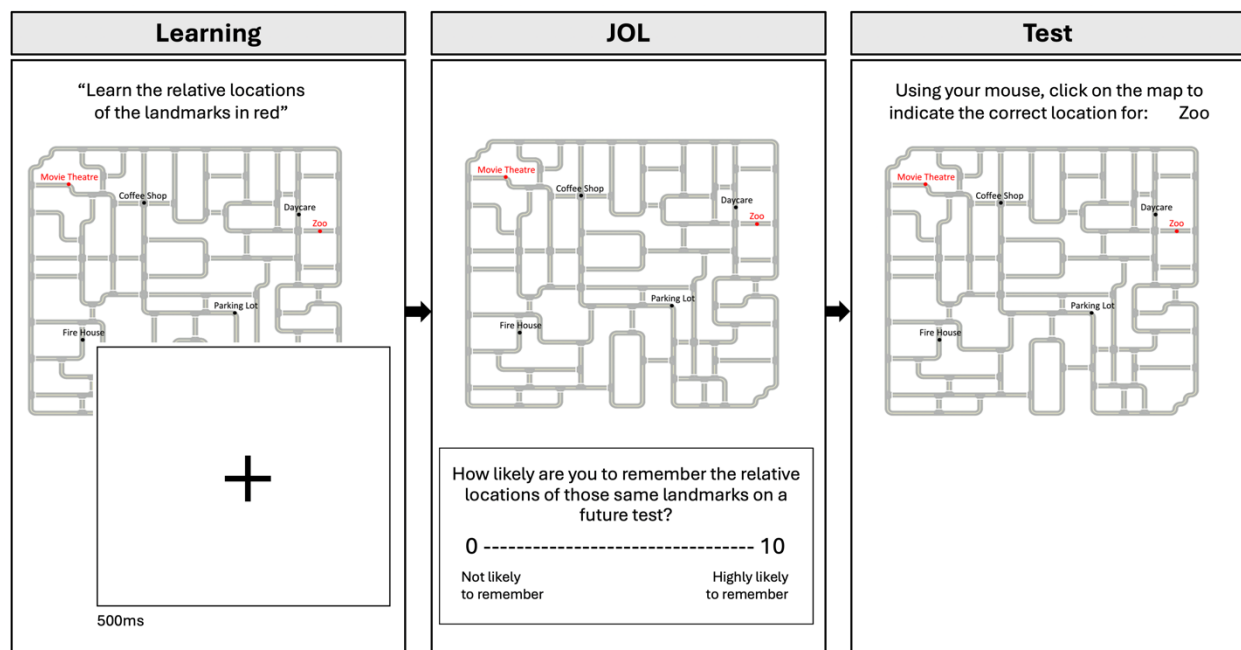
maps and 158KB for high complexity maps. This difference between map complexity types was statistically significant ($t(18) = 31.66, p < .0001$).

JOLs. After each learning trial, participants viewed the same map again with the cue (first landmark of a pair) and context landmarks present, but the target (second landmark of a pair) absent. They were prompted to make a JOL, where they rated their likelihood to remember the relative locations of those same landmarks on a future memory test. Participants provided a rating using a sliding scale from 0 (not at all likely to remember) to 10 (very likely to remember) (Son & Metcalfe, 2005).

Target landmark location memory test. Participants were shown the same maps that they made JOLs on (cue and context landmarks present, target landmark absent). The order of test stimuli was randomized. They were provided with the target landmark name and they clicked on the map to indicate the correct target landmark location. For instance, if the landmark pair was ‘Movie Theatre – Zoo,’ they were shown a map with ‘Movie Theatre’ in red and the context landmarks in black, just as they saw while making a JOL during learning. They were instructed to click on the map to indicate the correct location for ‘Zoo.’ Coordinates and response times were recorded. The dependent variable was distance from the studied location, calculated from these coordinates. An example learning, JOL, and target map test trial is provided in Figure 2.

Figure 2.

Examples of a Learning, JOL, and Test Trial for a Low Complexity Map in Experiment 1



After completing the target landmark location memory test, two exploratory tests were given: the free recall test of landmark names and context landmark location memory test. The implementation of exploratory map memory measures was supported by the notion that, "one memory paradigm does not cover the entire spectrum of options for examining the retrieval of map information" (Edler et al., 2017, p. 80).

Free recall of landmark names. This test prompted participants to recall as many landmark names as they could from memory and type them into a response box.

Context landmark location memory test. This test consisted of a screen containing a blank map each context landmark name next to the map. Each context landmark name was associated with a color, and the instructions stated that participants should click on the color that

represented the landmark location they wished to place and then click on the map to indicate the correct location for all 12 context landmarks. Coordinates and response times were recorded.

Attention checks. Given that remote participants may not conform to the same expectations in-person participants are subject to (Shamon & Berning, 2020), participants saw two attention checks during the experiment. The first attention check was drawn from the Prolific Research Help Centre and stated, “The color test is simple, when asked for your favorite color you must enter the word 'green' in the text box below.” Participants were expected to enter ‘green’ in the answer box provided. The next attention check was pulled from the literature and read, “I was born before 1920” with response options as follows, “Does not apply/ Somewhat applies/ Likely applies/ Applies completely” (Shamon & Berning, 2020). These attention checks can be valuable to screen for careless responses among participants and have been shown to exert a motivational influence (Shamon & Berning, 2020).

Retention interval. Participants completed a 2-minute retention interval after the learning phase. Retention intervals serve multiple purposes, including heightening the likelihood that studied materials were not actively held in working memory during testing and promoting consolidation before testing occurs. Findings suggest even a 1-minute retention interval is satisfactory (Chan et al., 2018). Participants were told that they would complete a visual search task involving a drawing with hidden items in it. The list of hidden items was provided at the bottom of the photo. They were given 2 minutes to study the photo and find the hidden items. Then, they reported how many items they found. The hidden picture printable for adults was pulled from the internet (<https://www.printablee.com>).

Questionnaires. Participants completed three questionnaires: demographics, strategies, and sense-of-direction. They also answered a question at the end of the study about whether they

had difficulty loading the displays. Demographics questions asked for participant gender, ethnicity, highest degree completed, and age. The remaining strategies and SOD questionnaires were included for exploratory purposes. The strategies questionnaire inquired on the participant experience learning and being tested on the maps. Participants responded to some questions using a Likert scale and had the option to describe their responses in an answer box. One of the strategies questions probed at whether they took a photo of the map(s) with their phone as a strategy (to be used for screening purposes). Finally, participants rated the overall difficulty of remembering the information needed at test on a scale from 1 (not difficult) to 7 (very difficult) and were asked whether it was more difficult learning from maps with fewer or more landmarks. The SOD questionnaire included the Santa Barbara Sense of Direction (SBSOD) (Hegarty et al., 2002) and a single question about their SOD (“When traveling somewhere in a new city, please rate your overall sense of direction”). Participants responded using a Likert scale ranging from 0 or “terrible” to 10 or “excellent”.

Procedures

The experiment was carried out over the internet using the Gorilla experiment building system (gorilla.sc). The map task was split into three phases: practice, learning, and test.

Practice phase. At the start of the experiment, participants were told, “For this task, you will study maps. Your goal is to learn where certain landmarks are relative to other landmarks on the maps. As you learn, you will rate your likelihood to remember the landmark locations on a future test. Then, your memory will be tested.” Next, they were taught about what relative landmark locations are. They were told, “Let’s start by learning a little bit about relative landmark locations. Then, you will get to practice before the actual task begins.” Participants saw a blank practice map, and then the same map again with sample landmarks highlighted on it.

They were told, “The maps you will learn from contain landmarks. Some sample landmarks are highlighted below.” Then they saw a map without the landmarks highlighted and the instruction stated, “From maps, we can learn about relative landmark locations.” Next, they saw a map with two of the landmarks in red font. The instruction read, “As an example, we can remember the location of twig relative to where grass is located. In other words, we can remember the relative locations of these landmarks.” They were reminded, “In this task, you will study and learn where landmarks on maps are located relative to each other.” They, they were told, “After you study a map, you will automatically be prompted to rate your likelihood on a scale of 0-10 to remember the relative landmark locations on a future memory test. If you are highly likely to remember the landmark locations, pick 10. If you are highly unlikely to remember, pick 0. If you are somewhere in-between, pick your best estimate accordingly. Try to use the entire scale when responding throughout the experiment. Your rating is called a judgment of learning.”

Practice learning. Participants completed two practice learning trials. They were reminded that their goal is to learn where certain landmarks are relative to others on the maps and were given five seconds to study each practice map. On one trial, they studied a landmark pair on a low complexity map. On another trial, they studied a different pair on a high complexity map. The landmark name pairs studied were leaf and rock (one trial) and tree and flower (another trial). Half of the participants saw leaf/rock on a low complexity map, while the other half saw it on a high complexity map. The same was true for tree/flower. The stimuli order was randomized. Participants saw a map showing two critical landmarks in red font and either a low (two) or high (six) number of additional context landmarks scattered across the map. An instruction at the top of the screen reminded them of their goal, namely, “learn the relative locations of the landmarks in red.”

Practice JOLs. After five seconds, participants were automatically forwarded to the next screen. Here, the same map was shown, except the target landmark (rock) was missing. Participants were prompted with the question, “How likely are you to remember the relative locations of those same landmarks on a future memory test?” The response scale ranged from 0 (not likely to remember) to 10 (highly likely to remember). Participants were encouraged to use the entire response scale throughout the duration of the experiment. After making their JOL, a 500ms fixation cross appeared and the next learning trial began.

Practice target landmark location memory test. Next, participants completed a cued-recall memory test. They were told, “Now, let’s take a practice memory test. Here, your memory for the relative landmark locations will be tested. You will be shown the same maps that you saw when rating your likelihood to remember the relative landmark locations. Your task is to click on the map to indicate the correct location of the missing landmark.” On each test trial, they were shown a map that they saw while making their JOL. Participants were instructed to use their mouse to click on the map to indicate the correct target landmark location. For instance, they were told, “Using your mouse, click on the map to indicate the correct location for: Rock.” Once they clicked on the map, they automatically advanced to the next trial.

Practice free recall of landmark names. After the target landmark location memory test, they completed a free recall test of landmark names. They were told, “Now, you will take a practice landmark name test. Here, your memory for landmark names will be tested. You will be asked to remember as many landmark names as you can from the maps. Type them into the response box provided. If you can’t remember all of the names, simply do your best.”

Practice context landmark memory test. Next, they completed the context landmark memory test. They were told, “Now, you will complete a practice memory context landmark

memory test. Here, your memory for context landmarks will be tested. There were several landmark names that always appeared on the maps but were never part of a pair. These landmark names will be provided for you. Your task is to try and remember where these landmarks were located. You will be shown a blank map and asked to indicate the correct location on the map for each of the landmark names.” Participants were provided with a blank map and a list of the eight, practice context landmark names, each associated with its own unique color. Their task was to mark on the map to indicate the correct location for each context landmark.

Learning. During the learning phase, participants were reminded that their goal was to learn where certain landmarks were relative to others on the maps. They were told that they would be given 10 seconds to study each map. Participants studied 20 map stimuli. Half were low complexity maps (4 context landmarks) and half were high complexity (12 context landmarks). Just like the practice phase, they were instructed to learn the relative locations of the landmarks in red for every trial. The map stimulus was presented below the instruction. After 10 seconds, they were automatically forwarded to the next screen where they make a JOL.

The first attention check was placed at the end of the learning phase. Participants entered a specified color name in the box provided. Then, they completed the two-minute retention interval before beginning the test phase.

Test. The test phase largely mirrored the practice test implemented at the start of the experiment. Participants saw the same 20 maps that were used to make JOLs. They were instructed to click on the map to indicate the correct target landmark location. Once they clicked on the map, they automatically advanced to the next trial. After the target landmark location memory test, they completed a free recall test of landmark names and a context landmark

memory test, where they marked on the map to indicate the locations for each of the 12 context landmarks.

The second attention check was placed after the test phase. Participants selected the extent to which the statement “I was born before 1920” applied to them.

Questionnaires. After testing, they completed the demographics, strategies, and SOD questionnaires in addition to the technical difficulty question.

Results

Data Screening

The data were screened to ensure that only participants who passed attention checks, completed the experiment, reported that they did not take a photo of the map(s) as a strategy used in the strategies questionnaire, and who did not report technical difficulties on the displays questionnaire were included in the analyses.

Analysis Plan

I used RStudio Version 2023.12.1+402 (2023.12.1+402) to complete all analyses. Prior to model estimation, outliers in the dependent variables in question were detected using the Interquartile Range (IQR) method and subsequently excluded from analyses. The IQR is computed by calculating the difference between the first quartile, or 25th percentile of the data, and the third quartile, or 75th percentile of the data. The lower bound is the first quartile - 1.5 * IQR, and the upper bound is the third quartile + 1.5 * IQR. Any data points that fall below the lower bound or above the upper bound are considered outliers. An advantage of the IQR method is that it is commonly used when the dataset is slightly asymmetric (Wada, 2020).

I estimated multilevel models (MLMs) to assess the research questions of interest. For this work, multilevel modeling is advantageous because one can determine the amount of variability that is accounted for by variation across participants and pairs (Murayama et al.,

2014). For instance, some individuals may be more skilled at map learning than others, and some pairs may be more memorable than others. I estimated multilevel models (MLMs) with 1) the continuous outcome variable, JOL magnitude, predicted by complexity level (low and high), a binary within-subjects predictor, and 2) the continuous outcome variable, memory accuracy, predicted by complexity level, JOL magnitude, and their interaction. Entering JOL magnitude as a predictor for memory accuracy allowed me to assess the predictive strength of JOLs (i.e., metacognitive accuracy). The interaction between JOLs and complexity level allowed me to investigate whether that predictive strength varied by complexity. Upon model estimation, I assessed model fit compared to a previously estimated model based on theoretical motivation. I present statistical test results comparing log likelihood values to support conclusions regarding model fit throughout the results section. I also present all model formulas and corresponding AIC values in Appendix B.

Model Estimation: JOL Magnitude

First, I investigated whether complexity predicted JOL magnitude. JOL ratings ranged from 0 (not likely to remember) to 10 (highly likely to remember). Learners provided a rating of 4 on average.

How much of the variation in JOL magnitude is between pairs and participants?

To evaluate the variation in JOL magnitude by pair and participant, I estimated a series of empty models. I evaluated the ICC and conducted likelihood ratio tests after estimating each subsequent model to see whether addition of the random effect in question decreased the model deviance.

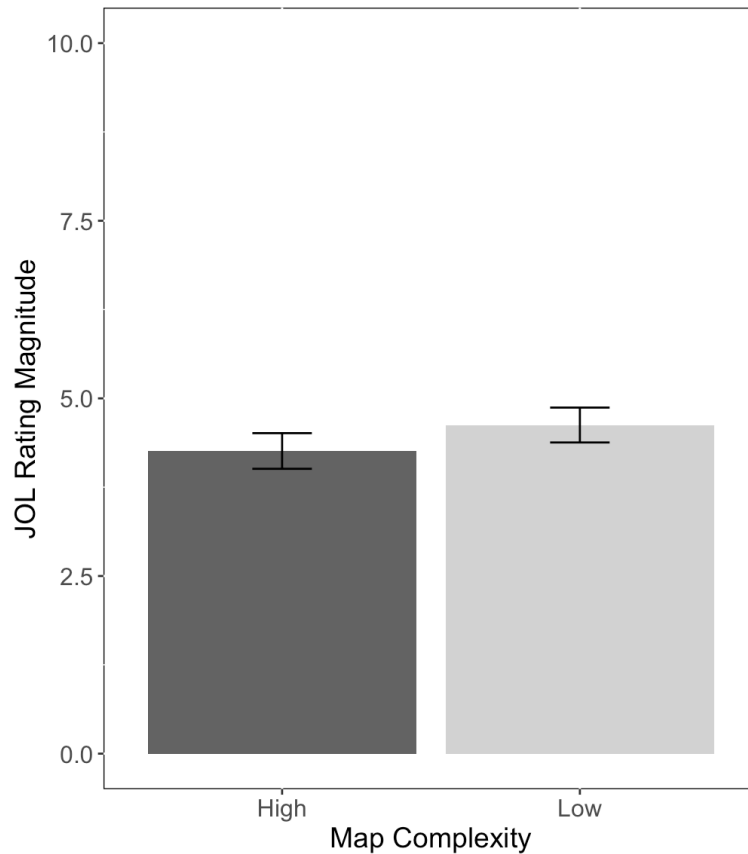
The first empty model included a random effect for the intercept for pair and revealed that pair did not account for substantial variation in JOL magnitude. The second included a

random effect for the intercept for participant, revealing that 39% of the variation in JOL magnitude was accounted for by this random effect. A random intercept for participant was therefore retained in the model predicting JOL magnitude.

Does complexity predict JOL magnitude?

Next, I added a fixed effect of complexity to the model to test whether JOL magnitude differed between levels of complexity. This model (E1J1, Appendix B) was a better fit for the data than the empty model based on AIC and log likelihood value comparisons ($\chi^2(1) = 8.78, p < .005$).

I proceeded by estimating a model with a fixed effect of map complexity and a random effect for 1) the intercept for participant and 2) the slope for map complexity. The decision to allow for a random slope for map complexity was theoretically driven due to competing JOL magnitude hypotheses. Namely, not only can JOL magnitude vary by map complexity, but the rate at which JOL magnitudes rise or fall may also vary across levels of complexity. I quantifiably confirmed this decision with model fit comparisons. This model with a random intercept for participant and slope for map complexity (E1J2, Appendix B) was a better fit for the data than the model with a random intercept for participant alone ($\chi^2(2) = 25.18, p < .001$). I observed a main effect of map complexity, such that learners rated themselves more likely to remember locations of landmarks studied on less complex maps ($\beta = -0.18, SE = 0.09, z(1040) = -1.98, p < .05$) (Figure 3).

Figure 3.*JOL Rating Magnitude by Map Complexity***Planned Model Estimation: Memory & Metacognitive Accuracy**

Next, I investigated whether map complexity, JOL, and their interaction (i.e., metacognitive accuracy) predicted map memory. Map memory was represented by standardized distance scores, or how far a participant's remembered target landmark location at test varied from the studied target landmark location at study. Higher values reflected worse memory. I detected and excluded four outlier distance values from analyses. On average, a remembered location was around 256 distance units away, or around 1.5 inches, from the studied location.

How much of the variation in memory accuracy is between pairs and participants?

To evaluate the variation in memory accuracy by pairs and participants, I estimated an empty model with a random effect for the intercept for pair. This model revealed that pairs accounted for a minimal amount of variation in memory ($ICC = .05$). The second empty model included a random effect for the intercept for participant, revealing that 25% of the variation in memory scores was accounted for by participant variation. A random intercept for participant was retained in the models predicting memory.

Does complexity predict memory accuracy?

Next, I added a fixed effect of complexity to the model to test whether memory accuracy differed between low and high complexity maps (E1M1, Appendix B). This model was not a better fit for the data than the empty model based on AIC and log likelihood value comparisons, meaning complexity did not significantly impact landmark location memory.

Do JOLs predict memory accuracy above and beyond complexity?

To determine whether JOLs predicted memory and whether predictive accuracy varied between levels of complexity, I estimated a model with a fixed effect of complexity and its interaction with JOL magnitude (E1M2, Appendix B). This model was not a better fit for the data than the empty model based on AIC and log likelihood value comparisons, meaning JOLs did not significantly predict landmark location memory, nor did the predictive value of JOLs significantly vary by complexity.

Exploratory Analyses***JOL Magnitude Across the Learning Span***

Prior research with similar map stimuli to the present study found that JOLs decreased over time (i.e., across learning trials) (Mason et al., 2023), likely reflecting the cognitive load of

information accumulation. I estimated exploratory models to assess whether the finding replicated and whether it varied as a function of map complexity.

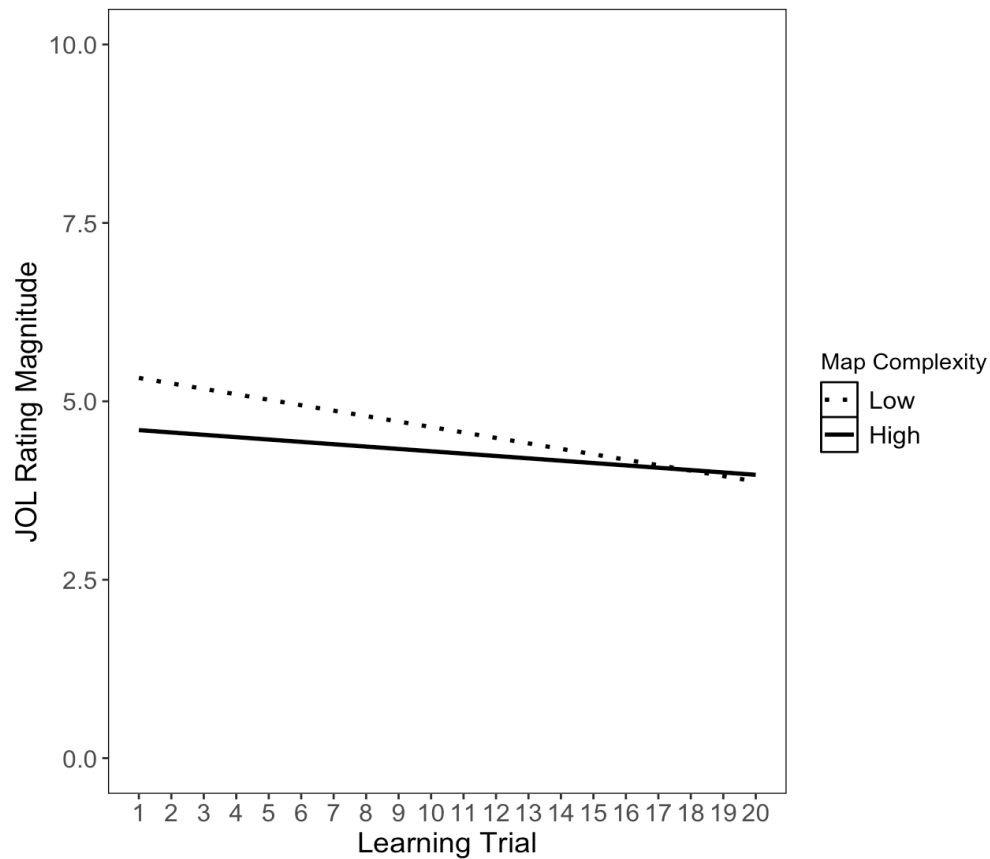
I allowed for a random effect for 1) the intercept for participant, and 2) the slope for learning trial. The decision to allow for a random effect for the slope for learning trial was theoretically driven. Namely, JOL ratings might decrease as a function of information accumulation and this effect might be stronger for some participants than others. To quantifiably confirm this decision, I also estimated a model that only allowed for a random effect for the intercept for participant (E1J3, Appendix B). The model with a random intercept and slope for learning trial (E1J4, Appendix B) was a better fit for the data ($\chi^2(2) = 64.81, p < 0.001$).

I observed main effects of learning trial, map complexity, and a learning trial*map complexity interaction (Table 2). Learners rated themselves increasingly less likely to remember locations of landmarks as the learning span progressed, particularly for low complexity maps (Figure 4).

Table 2*Parameter Estimates for Model E1J4*

Fixed effects	Estimate	<i>SE</i>	<i>No. Obs</i>	<i>Z</i>	<i>p</i>	95% CI [LL, UL]
Parameter						
Intercept	5.02	0.29	1040	17.25	< .001****	[4.45, 5.59]
Learning Trial	-0.05	0.02	1040	-2.87	<.005***	[-0.09, -0.02]
Map Complexity	-0.39	0.12	1040	-3.21	< .001****	[-0.62, -0.15]
Trial*Complexity	0.02	0.01	1040	2.14	< .05*	[0.00, 0.04]

Note. No. Obs = Number of observations. * $p < .05$, ** $p < .01$, *** $p < .005$, **** $p < .001$.

Figure 4.*JOL Rating Magnitude by Learning Trial and Map Complexity****Reaction Time***

I conducted exploratory analyses on reaction times (RTs) to assess whether a significant amount of variation in RT was accounted for by pair and participant and whether levels of complexity, JOL magnitude, and their interaction predicted RT on the target landmark location memory test. I identified and excluded nine outlier RT values from analyses via boxplot visualization. On average, participants took around eight seconds to indicate their response for each test map.

The first empty model included a random effect for the intercept for pair and revealed that pair did not account for substantial variation in RT. The second included a random effect for the intercept for participant, revealing that 43% of the variation in RT was accounted for by this random effect. A random intercept for participant was therefore retained in the model predicting RT. I added a fixed effect of complexity and its interaction with JOL magnitude (E1RT1, Appendix B). This model was not a better fit for the data compared to the null model according to AIC value comparison and likelihood ratio testing, suggesting that JOLs, map complexity, and their interaction did not meaningfully predict response time at test.

Free Recall Test of Landmark Names

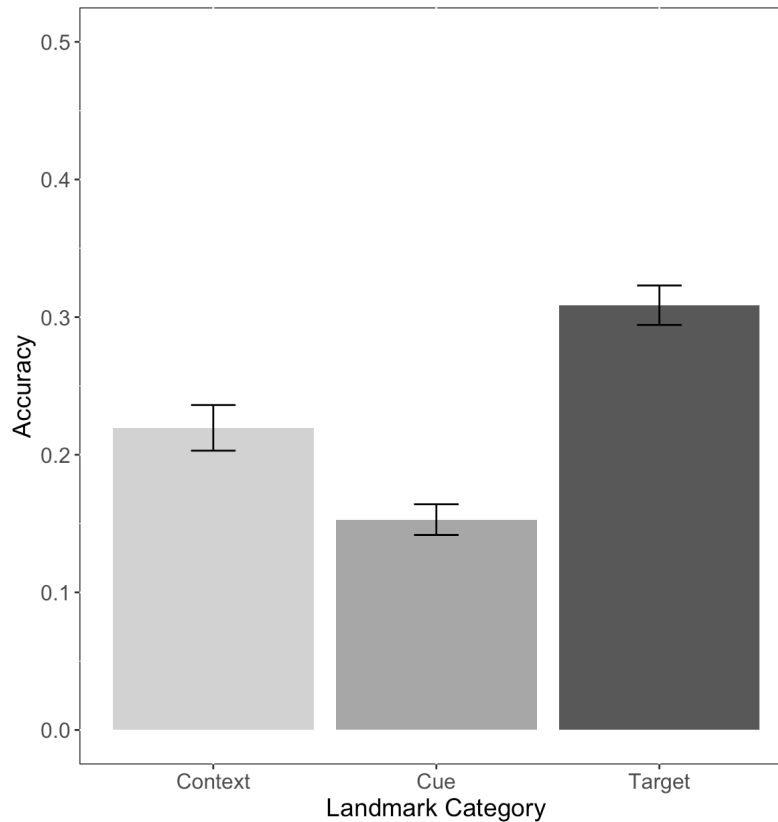
I conducted exploratory analyses on data from the free recall test of landmark names. Participants received a score of 1 for each landmark name they correctly recalled out of the total 52 landmarks and a score of 0 for those they did not recall. The landmarks were further coded by category (context, target, or cue). On average, around 23% of the landmark names were recalled.

How much of the variation in free recall memory performance is between landmarks and participants?

To evaluate the variation in free recall performance by landmarks and participants, I estimated a series of empty models. The first empty model included a random effect for the intercept for landmark and revealed that landmark accounted for 14% of the variability in free recall scores. The second empty model included a random effect for the intercept for participant, revealing that 12% of the variation in free recall scores was accounted for. Random intercepts for landmark and participant were therefore retained in the model predicting free recall scores.

Does category predict free recall scores?

I proceeded by adding a fixed effect of landmark category to the empty model with random intercepts for landmark and participant (E1FR1, Appendix B). Specifically, I assessed whether learners better remembered context landmark names as compared to other types of landmark names (cue or target), and whether their memory for cue versus context landmark names differed. This model was a better fit for the data as compared to the empty model ($\chi^2(2) = 16.78, p < .001$). The planned contrasts compared recall for context landmarks vs. other types of landmarks (i.e., cue or target), and recall for cue vs. target landmarks. I observed a main effect of category, such that learners better remembered target as compared to cue landmark names ($\beta = 0.52, SE = 0.12, z(2704) = 0.54, p < .001$) (Figure 5). The graph also suggests that learners better remembered context compared to cue landmark names during test.

Figure 5.*Free Recall of Landmark Names by Landmark Category****Context Landmark Memory Test***

Due to poor data quality, the context landmark memory test data was not analyzed.

Santa Barbara Sense of Direction

Next, I investigated whether scores on the Santa Barbara Sense of Direction (SBSOD) scale were correlated with two primary outcome measures of interest: memory accuracy or JOL rating magnitude, in addition to an exploratory, individual difference measure: the type of map (i.e., low or high complexity) that the learner deemed as more difficult to learn from during the experiment, described here as perceived map difficulty. Pearson correlations revealed that

SBSOD scores were correlated with JOLs ($r(50) = .33, p < .05$), but not memory accuracy or gender. A majority of learners (60%) reported learning from high complexity maps as more difficult, while only 11% deemed them less difficult, and 29% reported that the difficulty learning from each map type was the same. There was not a significant correlation between perceived map difficulty and SBSOD scores.

JOL Magnitude: SBSOD*Map Complexity. I proceeded by adding a fixed effect of SBSOD score into the empty JOL model (E1J5, Appendix B). This model was a better fit for the data than the empty model ($\chi^2(1) = 4.87, p < 0.05$) and produced a main effect of SBSOD score ($\beta = 0.42, SE = 0.18, z(1040) = 2.28, p < .05$). I added a fixed effect of map complexity and its interaction with SBSOD to test whether the relationship between SBSOD scores and JOL magnitude varied by map complexity (E1J6, Appendix B). This model was a better fit for the data than the previous model ($\chi^2(2) = 15.65, p < .001$). I observed a main effect of SBSOD, map complexity, and a SBSOD*complexity interaction (Table 3). Learners with higher SBSOD scores also produced higher JOLs, and this trend was most extreme for high complexity maps (Figure 6). Those who rated themselves as having a high SOD ability assigned higher JOLs to high as compared to low complexity maps.

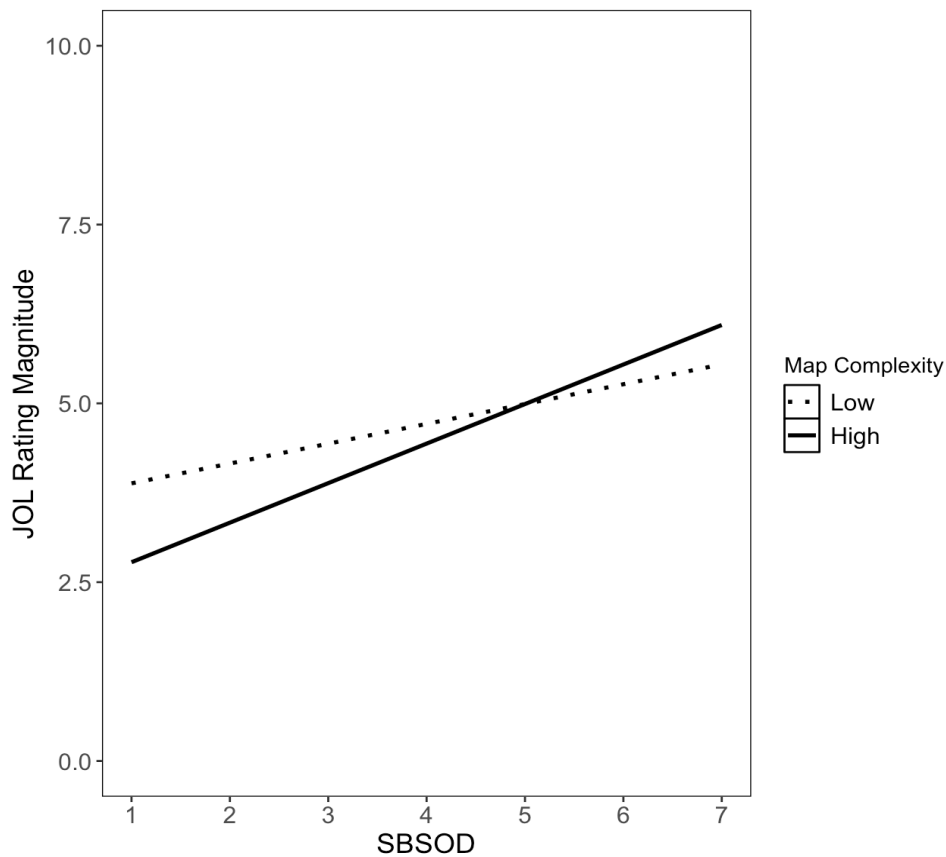
Table 3*Parameter Estimates for Model E1J6*

Fixed effects	Estimate	SE	No. Obs	Z	p	95% CI [LL, UL]
Parameter						
Intercept	2.91	0.71	1040	4.10	<.001****	[1.52, 4.31]
SBSOD	0.42	0.18	1040	2.26	< .05*	[0.05, 0.78]
Map Complexity	-0.69	0.20	1040	-3.42	<.001****	[-1.08, -0.29]
SBSOD*Complexity	0.14	0.05	1040	2.64	< .01**	[0.04, 0.24]

Note. No. Obs = Number of observations. *p < .05, **p < .01, ***p < .005, ****p < .001.

Figure 6.

*JOL Rating Magnitude: Map Complexity*SBSOD*



Discussion

In Experiment 1, I investigated the role of map complexity, as operationalized by the number of landmarks present on the map, for metacognitive judgment formation, memory, and metacognitive or predictive accuracy. In alignment with prior research (Mason et al., 2024), learners rated themselves more likely to remember less complex maps, suggesting that they were sensitive to perceptual aspects of the map stimuli (Stevens & Carlson, 2019). Contrary to my hypotheses, multilevel modeling results suggested no differences in memory or metacognitive

accuracy between map types. In other words, though learners deemed themselves more likely to remember less complex maps, their memory did not vary, nor did their ability to accurately predict their memory.

Consistent with prior work (Mason et al., 2024), learners rated themselves less likely to remember landmark locations as the learning span progressed. Such a pattern may be representative of information accumulation across the learning span resulting in cognitive load, which was subsequently reflected in JOL magnitudes. Notably, this trend was steeper for low complexity maps. Given that JOLs for low complexity maps were highest at the start of learning (Figure 4), one potential explanation for this pattern is that learners over-estimated their memory predictions for low complexity maps at the start of learning, leading to a steeper decline in ratings as learning progressed.

Though JOL ratings were not predictive of memory differences across map types, learners with higher self-reported SOD ability produced higher JOLs, particularly for high complexity maps. One interpretation for this finding is that skilled spatial learners may have subjectively benefitted from additional reference points offered by high complexity maps, resulting in higher JOLs for those maps specifically. In contrast, those who reported lower spatial ability may have not appreciated the additional complexity, resulting in lower ratings for high complexity maps among these individuals. In sum, learners' judgments may reflect a preference for one type of map versus the other, and this preference interacts with self-reported spatial ability.

In addition to investigating memory for landmark locations, I also conducted analyses to determine whether memory for landmark names varied across landmark type (cue, target, or context). Context landmark names appeared on all maps across learning, thereby motivating

inquiry around whether these names were better remembered as compared to cue or target landmarks, which were only studied once. Interestingly, learners showed superior memory for target landmarks, followed by context, and then cue. This finding aligns with evidence from the memory literature broadly, suggesting that prompting retrieval for materials at study may promote memory at test (Ariel et al., 2021; Karpicke & Roediger, 2008; Smith et al., 2016; Thomas & McDaniel, 2007). Here, the act of making target-absent JOLs may have initiated retrieval for target landmark locations and names, resulting in optimal memory for target names when tested. In sum, directing learners to monitor one type of landmark information (location) subsequently resulted in better memory for a different form of information (landmark names).

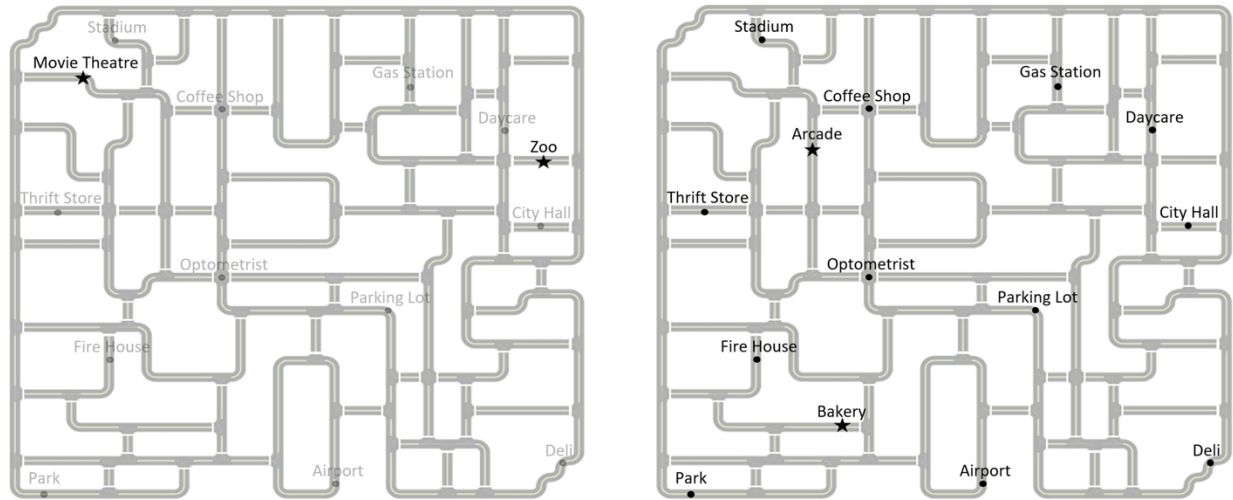
In sum, in Experiment 1, variations in map complexity impacted JOLs but did not result in memory or metacognitive accuracy differences, and learners high in spatial ability provided higher JOLs for high complexity maps. In Experiment 2, I held map complexity constant and investigated the role of a second map display manipulation: landmark salience.

Experiment 2

In Experiment 2, learners were again instructed to learn the relative locations of two critical landmarks from maps. On half of the trials, the landmark pair was displayed in 100% opacity and the salience of the remaining landmark information was reduced by 70%. See Figure 7 for an example of this manipulation. After studying each map, learners provided a JOL about their likelihood to remember the relative landmark locations on a later test. Then, their memory was tested. I hypothesized that JOL magnitudes and memory would be greater for pairs studied on maps with perceptually salient critical landmark pairs as compared to standard maps.

Figure 7.

Examples of a Perceptually Salient (Left) and Standard (right) Learning Map in Experiment 2



Note. The critical pairing is indicated by stars for landmark points. Context landmark points are indicated by circles.

Method

Participants

In Experiment 2, 60 participants ages 18-35, within the United States were recruited. Data from two participants was excluded from analyses due to technical difficulties. The remaining six participants were excluded due to redundant JOL responses, such that they only responded with one to two numbers on the scale or responded with the same number for nine or more consecutive trials. From the remaining 52 participants, 48% identified as women, 46% identified as men and 6% identified as nonbinary. The sample was majority White (78%) and non-Hispanic or Latino (90%). Most participants completed education beyond high school (83%).

Design

Map salience was a within-participants variable. All participants studied half of the landmark pairs on perceptually salient maps and half on standard maps. They made JOLs after each study trial and were tested on their memory for relative landmark locations.

Materials

Landmark pairs. The 20 landmark pairs were the same as those used in Experiment 1.

Context landmarks. The same 12 context landmarks from Experiment 1 appeared on all the maps but again were never part of a critical pair.

Maps. Each map stimulus contained two critical landmarks labeled with stars and 12 additional context landmarks labeled with circles. All landmarks appeared in black font. Therefore, every learning map stimulus contained 14 total landmarks. On half of the maps, the critical landmark pair was made salient, and the transparency of the remaining map information was reduced by 70%.

During the practice phase, the maps contained two critical landmarks and six additional context landmarks. During the learning phase, they contained two critical landmarks and 12 additional context landmarks. Order of landmark pair presentation on maps was randomized. Half of the participants studied 10 landmark pairs from list (A) on perceptually salient maps and 10 from list (B) on standard maps, while the other half studied list (A) on salient maps and list (B) on standard maps. The order of map presentation was randomized, such that it was possible for a participant to study a salient map, followed by a standard map, and so on. As in Experiment 1, learners saw between 1-3 maps of the same type consecutively on average.

JOLs. After each learning trial, participants saw the same map again with the cue and additional landmarks present but the target landmark absent. If they studied a salient map, they

saw a salient map during the JOL. If it was standard, it was standard during the JOL.

Immediately after study, they were prompted to make a JOL, where they rated their likelihood to remember the relative locations of the starred landmarks on a future memory test. They provided a rating using a sliding scale from 0 (not at all likely to remember) to 10 (very likely to remember). After making each JOL, participants saw a brief interval screen before automatically proceeding to the next trial.

Target landmark location memory test. If a particular pair was studied on a salient map, it was also tested on a salient map. Pairs studied on standard maps were tested on standard maps. The test stimuli were randomized. Participants were shown the same maps that they made JOLs on (cue and additional landmarks present, target landmark absent). They were reminded of the target landmark name and told to click on the map to indicate the correct target landmark location. There were 20 test trials. Response location coordinates and response times were recorded.

Free recall of landmark names. After the target landmark location memory test, they completed the free recall test of landmark names.

Attention checks. The same attention checks from Experiment 1 were used.

Retention interval. The same retention interval from Experiment 1 was used.

Questionnaires. The same questionnaires from Experiment 1 were used. The strategies questionnaire was modified slightly to probe at the subjective experience of the participant when learning from perceptually salient and standard maps. Participants were asked to report whether it was more difficult learning from maps with or without “faded” landmarks.

Procedures

Just as in Experiment 1, the experiment was carried out over Gorilla and split into practice, learning, and test phases.

Practice phase. The practice phase largely modeled that from Experiment 1. Now, participants practiced learning pairs on perceptually salient and standard maps. They were told to learn the relative locations of the starred landmarks.

Learning phase. During the experimental learning phase, participants studied 20 map stimuli with half of the pairs on perceptually salient maps and half on standard maps. Just like the practice phase, they were instructed to learn the relative locations of the starred landmarks. Then, they make a JOL. The ‘color’ attention check and retention interval are placed at the end of the learning phase.

Test Phase. The test phase largely mirrored the practice test implemented at the start of the experiment. Participants saw the same 20 maps that they made JOLs about during the learning phase. They clicked on the map to indicate the remembered target landmark location. After the target landmark location memory test, they completed the free recall test of landmark names. The ‘1920’ attention check was placed after the test phase.

Questionnaires. After the test phase, they completed the same questionnaires from Experiment 1. The strategies questionnaire was slightly modified.

Results

Data Screening

I conducted the same data screening as Experiment 1. I estimated multilevel models (MLMs) with 1) the continuous outcome variable, JOL magnitude, predicted by salience level (salient and standard), a binary within-subjects predictor, and 2) the continuous outcome variable, memory accuracy, predicted by salience level, JOL magnitude, and their interaction (i.e., metacognitive accuracy).

Model Estimation: JOL Magnitude

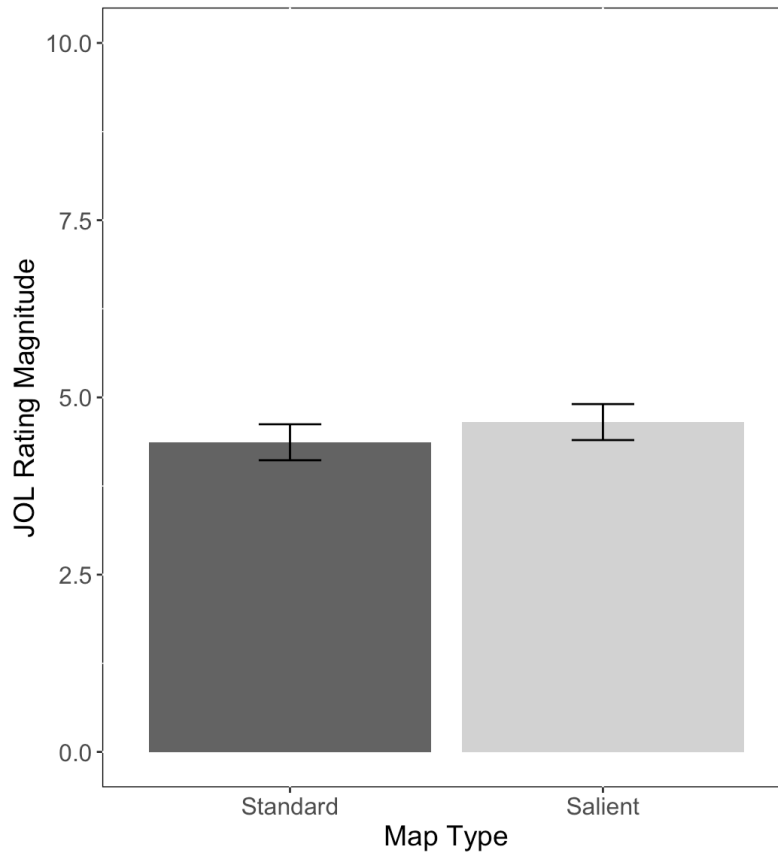
First, I investigated whether map salience predicted JOL magnitude. JOL ratings ranged from 0 (not likely to remember) to 10 (highly likely to remember). Learners provided a rating of 4.5 on average.

How much of the variation in JOL magnitude is between pairs and participants?

The first empty model included a random effect for the intercept for pair and revealed that pair did not account for variation in JOL magnitude. The second included a random effect for the intercept for participant, revealing that 43% of the variation in JOL magnitude was accounted for by this random effect. A random intercept for participant was therefore retained in the model predicting JOL magnitude.

Does perceptual salience predict JOL magnitude?

Next, I added a fixed effect of salience to the model to test whether JOL magnitude differed by landmark salience. This model (E2J1, Appendix C) was a better fit for the data than the empty model based on AIC and log likelihood value comparisons ($\chi^2(1) = 5.33, p < .05$). I also estimated a model with a fixed effect of map salience and a random effect for 1) the intercept for participant, and 2) the slope for salience; however, it was not a better fit for the data than the model with a random intercept for participant alone. I observed a main effect of map salience, such that learners rated themselves more likely to remember locations of landmarks studied on salient maps ($\beta = 0.14, SE = 0.06, z(1040) = 2.31, p < .05$) (Figure 8).

Figure 8.*JOL Rating Magnitude by Map Salience***Planned Model Estimation: Memory & Metacognitive Accuracy**

I investigated whether map salience, JOL, and their interaction (i.e., metacognitive accuracy) predicted map memory. I detected and excluded six outlier distance values from analyses. On average, a remembered location was around 254 distance units away from the studied location.

How much of the variation in memory accuracy is between pairs and participants?

To evaluate the variation in memory accuracy by pairs and participants, I estimated an empty model with a random intercept for pair. This model revealed that pair accounted for a minimal amount of variation in memory ($ICC = .06$). The second empty model included a random effect for the intercept for participant, revealing that 24% of the variation in memory scores was accounted for. A random intercept for participant was retained in the models predicting memory.

Does salience predict memory accuracy?

Next, I added a fixed effect of salience to test whether memory accuracy differed between standard and salient maps (E2M1, Appendix C). Though an effect of salience was not observed, this model was a better fit for the data than the empty model based on AIC and log likelihood value comparisons ($\chi^2(1) = 85.03, p < .001$). One interpretation for this statistical result is that the effect of salience on memory did not reach significance, but incorporating the effect into the model predicting memory accounted for more variability in the memory outcome than no predictor at all.

Do JOLs predict memory accuracy above and beyond salience?

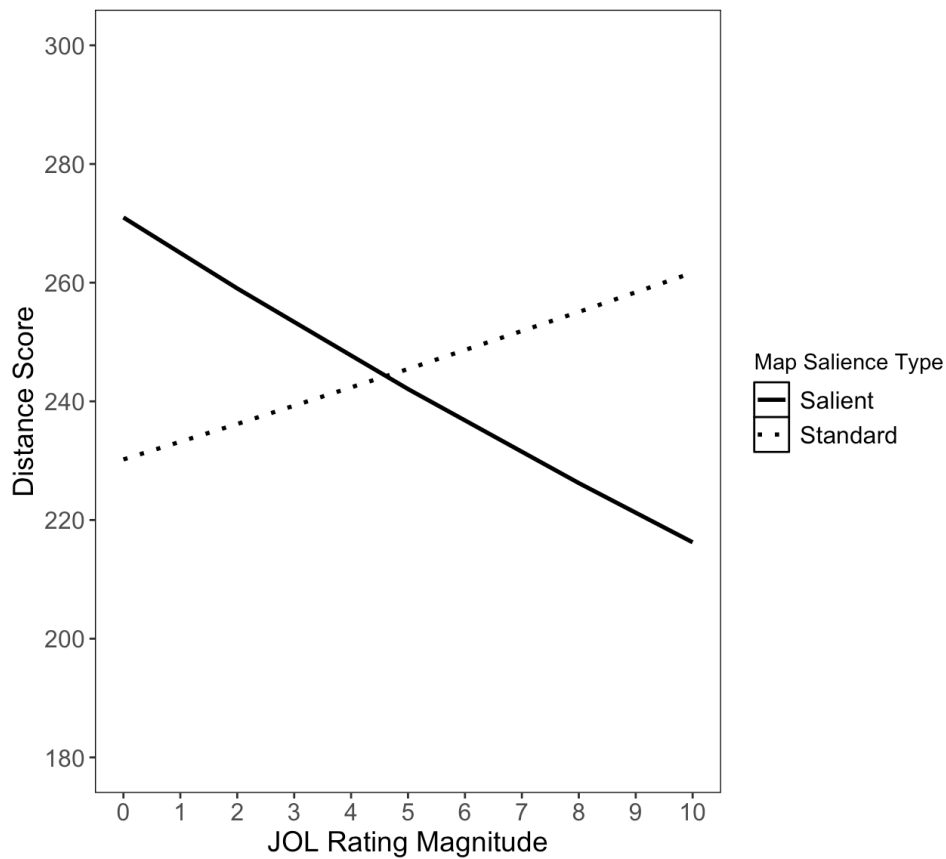
To determine whether JOLs predicted memory and whether predictive accuracy varied by salience, I estimated a model with a fixed effect of salience and its interaction with JOL magnitude (E2M2, Appendix C). This model was a better fit for the data than the model with salience alone ($\chi^2(2) = 6.29, p < .05$).

I observed a salience*JOL interaction. When studying salient maps, a positive JOL-memory relationship was observed, such that higher JOL ratings were associated with lower distance scores and thereby better memory for landmark locations. The opposite pattern was

observed for standard maps: Higher JOL rating magnitudes were associated with greater distance scores and thereby worse memory ($\beta = -0.02$, $SE = 0.007$, $z(1034) = -2.43$, $p < .05$) (Figure 9).

Figure 9.

*Distance Score: Map Type*JOL Rating Magnitude*



Exploratory Analyses

JOL Magnitude Across the Learning Span

I estimated exploratory models to assess whether JOL magnitudes decreased across learning trials and whether this pattern varied as a function of map salience. I allowed for a

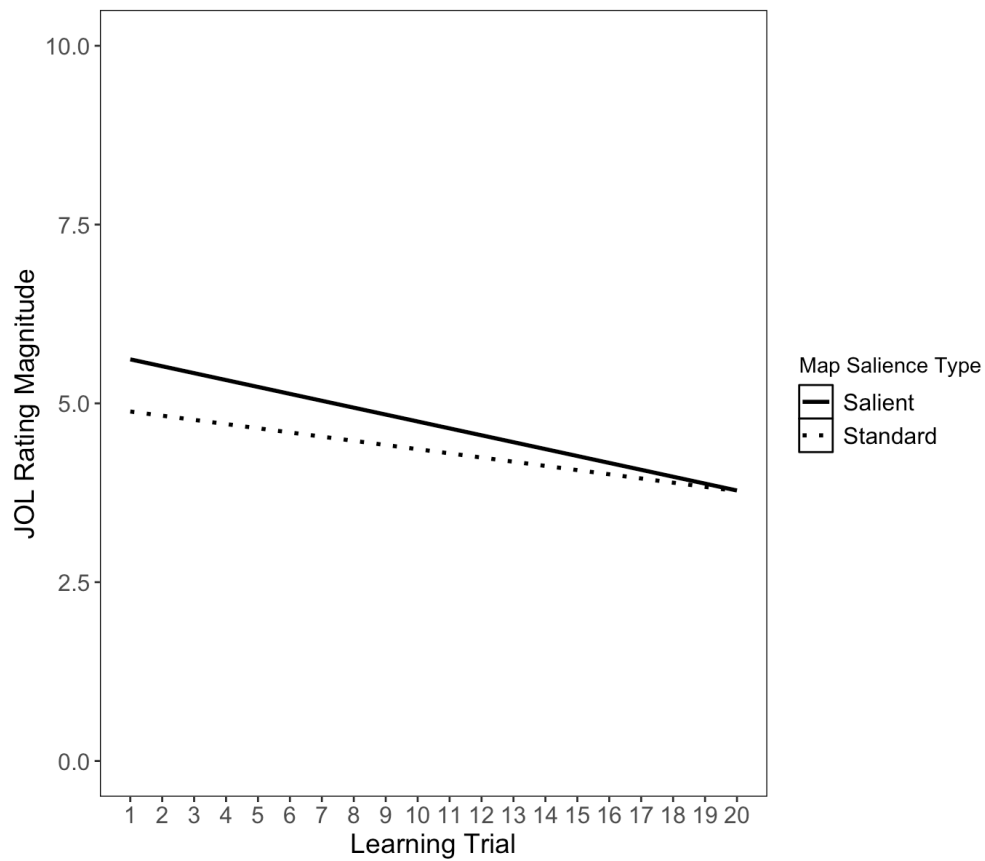
random effect for 1) the intercept for participant, and 2) the slope for learning trial. This model (E2J3, Appendix C) was a better fit for the data as compared to a model with only a random intercept for participant ($\chi^2 (2) = 48.30, p < .001$) or the empty model ($\chi^2 (5) = 111.6, p < .001$).

I observed main effects of learning trial and map salience (Table 4). Learners rated themselves increasingly less likely to remember landmark locations as the learning span progressed, independent of map type (Figure 10). The interaction was not significant.

Table 4*Parameter Estimates for Model E2MX*

Fixed effects	Estimate	SE	No. Obs	Z	p	95% CI [LL, UL]
Parameter						
Intercept	5.33	0.28	1040	18.8	< .001****	[4.77, 5.88]
Learning Trial	-0.08	0.02	1040	-4.43	<.001****	[-0.11, -0.04]
Map Salience	0.38	0.12	1040	-3.21	< .005***	[0.15, 0.62]
Trial*Salience	-0.02	0.01	1040	-1.86	.06	[-0.04, 0.00]

Note. No. Obs = Number of observations. *p < .05, **p < .01, ***p < .005, ****p < .001.

Figure 10.*JOL Rating Magnitude by Learning Trial and Map Type****Reaction Time***

I conducted exploratory analyses on reaction times (RTs) to assess whether a significant amount of variation in RT was accounted for by pair and participant and whether map salience, JOL magnitude, and their interaction predicted RT on the target landmark location memory test. I identified and excluded nine outlier RT values from analyses. On average, participants took around 6 seconds to indicate their response for each test map.

The first empty model included a random effect for the intercept for pair and revealed that pair did not account for variation in RT. The second included a random effect for the intercept for participant, revealing that 43% of the variation in RT was accounted for by this random effect. A random intercept for participant was therefore retained in the model predicting RT. I added a fixed effect of salience and its interaction with JOL magnitude. The results of this model (E2RT1, Appendix C) were non-significant, and it was not a better fit for the data as compared to the empty model based on log-likelihood value comparison, meaning that map salience type and JOL rating magnitude did not significantly predict response time at test.

Free Recall Test of Landmark Names

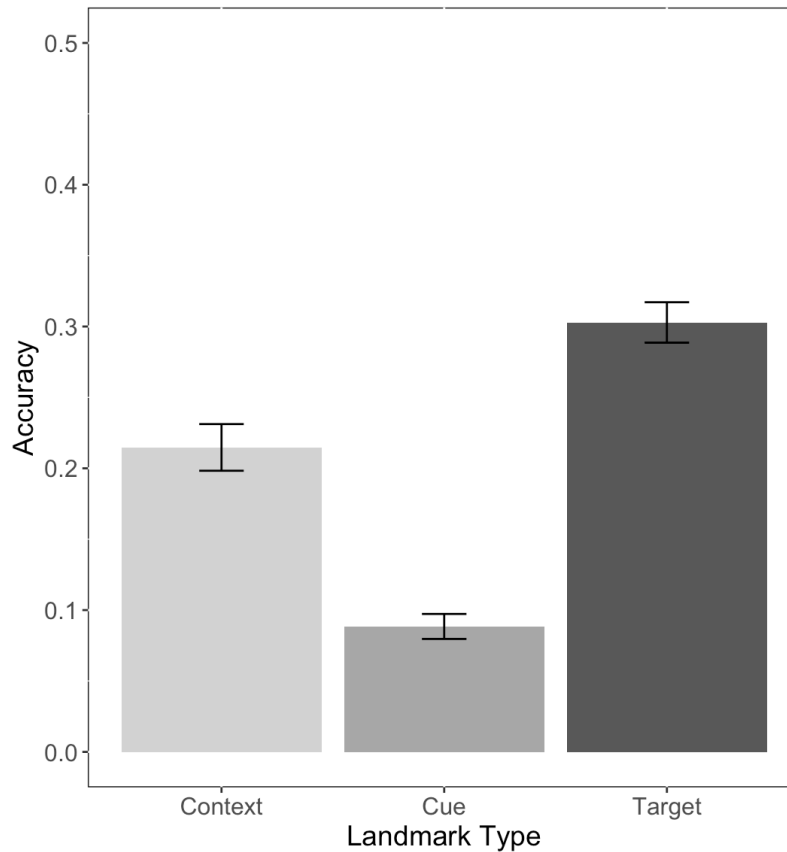
I conducted exploratory analyses on data from the free recall test of landmark names. As in Experiment 1, participants received a score of 1 for each landmark name they correctly recalled out of the total 52 landmarks and a score of 0 for those they did not recall. The landmarks were further coded by type (context, target, or cue).

How much of the variation in free recall memory performance is between landmarks and participants?

To evaluate the variation in free recall performance by landmarks and participants, I estimated a series of empty models. The first empty model included a random effect for the intercept for landmark and revealed that landmark accounted for 21% of the variability in free recall scores. The second empty model included a random effect for the intercept for participant, revealing that only 0.05% of the variation in free recall scores was accounted for. A random intercept for landmark was therefore retained in the model predicting free recall scores.

Does category predict free recall scores?

I proceeded by adding a fixed effect of landmark category to the empty model. Specifically, I assessed whether learners better remembered context landmark names as compared to other types of landmark names (cue or target), and whether their memory for cue versus target landmark names differed. This model (E2FR1, Appendix C) was a better fit for the data as compared to the empty model ($\chi^2(2) = 36.14, p < .001$). I observed a main effect of category, such that learners again better remembered target as compared to cue landmark names ($\beta = 0.80, SE = 0.12, z(2704) = 6.84, p < .001$) (Figure 11). The graph also suggests that learners again better remembered context compared to cue landmark names during test.

Figure 11.*Free Recall of Landmark Names by Landmark Category*

Santa Barbara Sense of Direction

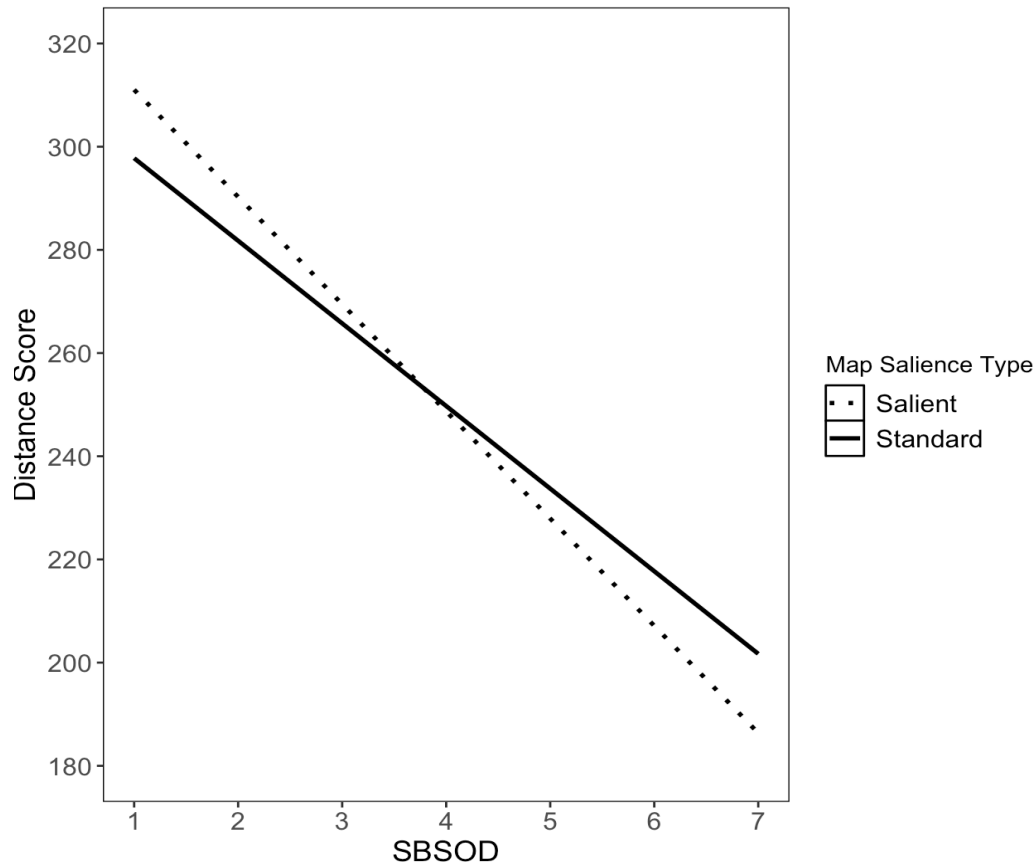
Next, I investigated whether scores on the Santa Barbara Sense of Direction (SBSOD) scale were correlated with two primary outcome measures of interest: memory accuracy or JOL rating magnitude, in addition to two exploratory measures of interest: gender and the type of map (i.e., standard or salient) that the learner deemed as more difficult to learn from during the experiment. Pearson correlations revealed that SBSOD scores were moderately correlated with

memory accuracy ($r(50) = .33, p < .05$), but not JOLs or gender. Further, JOL magnitudes and distance scores among men did not differ from those of women. Nonbinary participants were not included in the correlational analyses as there were too few datapoints for these individuals. When asked about the perceived map difficulty, 31% of learners reported salient maps were more difficult, 40% reported standard maps, and 29% deemed them the same difficulty. There was not a significant correlation between perceived map difficulty and SBSOD scores.

I proceeded by adding a fixed effect of SBSOD score into the empty memory accuracy model (E2M3, Appendix C). This model was a better fit for the data as compared to the empty model ($\chi^2(1) = 52.03, p < .001$) and produced a main effect of SBSOD score ($\beta = -18.4, SE = 7.05, z(1034) = -2.61, p < .01$). I proceeded by adding a fixed effect of map salience and its interaction with SBSOD to test whether the relationship between SBSOD scores and memory accuracy varied by map salience. This model (E2M4, Appendix C) was not a better fit for the data based on log likelihood comparison values. Independent of map salience type, learners with higher SBSOD scores performed better on the target landmark location memory test, as indicated by lower distance scores (Figure 12).

Figure 12.

*Distance Scores: Map Saliency*SBSOD*



Discussion

In Experiment 2, I investigated the role of landmark saliency for metacognitive judgment formation, memory, and metacognitive or predictive accuracy. Learners rated themselves more likely to remember pairs studied on salient maps. This finding extends observations from Experiment 1 and prior research (Mason et al., 2024), again suggesting that learners were sensitive to perceptual aspects of the stimuli (Klippel et al., 2006; Phillips & Noyes, 1977, 1982),

and such sensitivity was reflected by JOLs (Hegarty et al., 2012; Miller & Geraci, 2016; Stevens & Carlson, 2019). Though learners deemed themselves more likely to remember salient maps, multilevel modeling results suggested no difference in memory between map salience types. This finding raises questions around whether the perceptual salience manipulation actually honed learners' attention to the goal-relevant landmark information, or if variations in JOL magnitudes simply resulted from a perceptual ease of processing effect (Caduff & Timpf, 2008; Hegarty et al., 2010; Keil et al., 2020; Koriat, 1997; Yeh & Wickens, 2001).

Despite the lack of memory differences, differences in metacognitive accuracy were observed. A positive JOL-memory relationship was observed for salient maps, whereas a negative JOL-memory relationship was observed for standard maps. Said differently, higher JOLs were associated with lower distance scores and therefore better landmark location memory for salient maps, specifically. This finding suggests that manipulating the salience of goal-relevant map information supports a stronger relationship between metacognitive monitoring predictions and later memory. The map learning process is subject to many cues. Therefore, one interpretation for the present result is that making goal-relevant map cues more salient reduced the number of cues considered and subsequently encouraged learners to base their metacognitive judgments on those cues, specifically. Considering this result in tandem with the null memory finding could suggest that the perceptual salience manipulation successfully honed attention during judgment formation, resulting in a stronger JOL-memory relationship, but this honing effect did not translate to better memory at test. Future research may investigate whether more accurate monitoring judgments translate to effective metacognitive control processes for similar map-based materials, such as by giving learners an opportunity to self-select restudy opportunities (DeCaro & Thomas, 2020).

Consistent with findings from Experiment 1 and prior work (Mason et al., 2024), learners rated themselves less likely to remember landmark locations as the learning span progressed, now independent of map type. Furthermore, memory was again greatest for target landmark names, followed by context, and cue. Whereas SBSOD scores and JOLs were positively associated in Experiment 1, a positive relationship SBSOD scores and memory accuracy was observed in Experiment 2, such that learners higher in spatial skills showed better landmark location memory at test, independent of map type.

In sum, in Experiment 2, variations in map salience impacted JOLs and metacognitive accuracy, such that a positive JOL-memory relationship was observed for salient maps, specifically. Further, learners higher in spatial ability also had better memory accuracy for target landmark locations. Consistent with Experiment 1, optimal memory was observed for target landmark names. In Experiment 3, I shifted to address the role of semantic accessibility, or access to landmark name meaning, for metacognitive judgment formation, memory, and metacognitive accuracy.

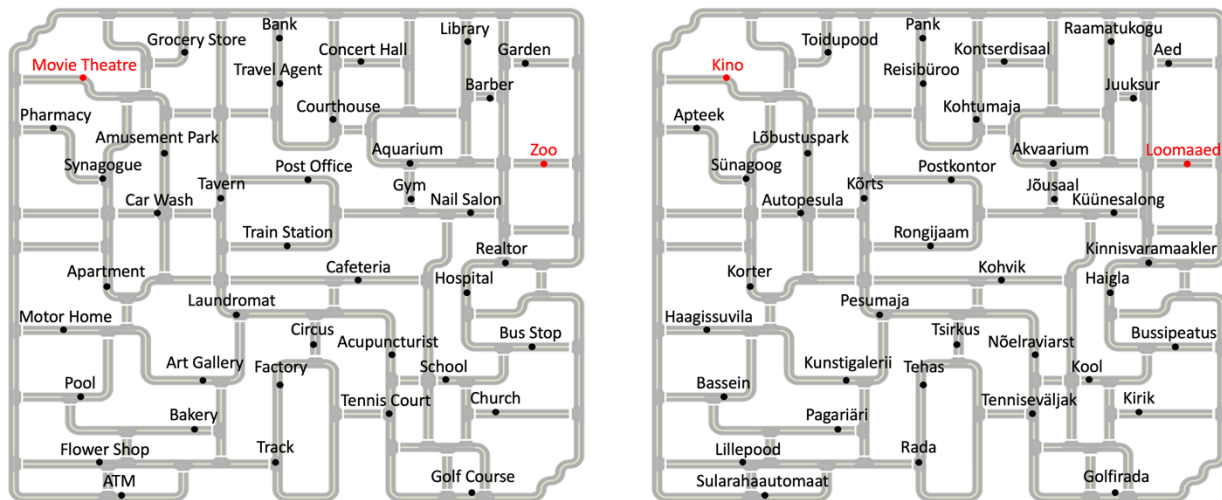
Experiment 3

In Experiment 3, I assessed whether map displays with landmark names in English produced higher JOL ratings, memory accuracy, and predictive accuracy among native English speakers, as compared to maps with landmark names in Estonian. Here, I manipulated semantic accessibility by presenting landmark names in the participants' native language (English) or in a language foreign to them (Estonian). An example of this manipulation is provided in Figure 13. As in Experiments 1 and 2, learners were instructed to learn the relative locations of two critical landmarks. After studying each map, they provided a JOL about their likelihood to remember the relative landmark locations on a later test, and their memory was tested. I hypothesized that JOL

magnitudes and memory would be greater for maps with landmark names in English as compared Estonian.

Figure 13.

Examples of a Map with Landmark Names in the Participants' Native Language (English, Left), and a Language Foreign to Them (Estonian, Right)



Method

Participants

In Experiment 3, 142 participants ages 18-35, within the United States were recruited. Participants were screened to ensure that English was their first language, and a post-survey probed at whether they had exposure to the language Estonian, Finno-Ugric languages, Slavic language, Scandinavian languages, or West Germanic languages. Data from 30 participants who were 1) in the Estonian map condition, and 2) familiar with one or more of the above languages was sorted into a “language exposed” group and later analyzed for exploratory purposes. Data

from 41 participants in the English and 44 participants in the Estonian condition was analyzed to answer the primary questions of interest.

Across all participants, data from four was excluded because they reported taking a photo of the maps as a strategy, data from three was excluded because they reported map loading troubles, and data from 20 was excluded because they had redundant JOLs such that they only responded with one to two numbers on the scale or responded with the same number for nine or more consecutive trials.

Data from the remaining 115 (30 Estonian-exposed) participants was analyzed. Among them, 42% identified as women, 57% identified as men, and 1% as nonbinary. The sample was majority White (69%) and non-Hispanic or Latino (89%). Most participants completed education beyond high school (84%).

Design

Semantic accessibility, as operationalized by landmark language, was a between-participants variable. Half of the participants studied landmark pairs on maps in English and half studied pairs on maps in Estonian. They all made JOLs after each study trial and were tested on their memory for relative landmark locations.

Materials

Landmark pairs. For this experiment, all pairs were translated into Estonian to develop the condition 2 materials. Estonian was selected because it fit the following language constraints: 1) it used the same alphabet as English, a constraint intended to help maintain a desirable level of difficulty, 2) it has few overlapping cognates with English, and 3) the landmark name translations were not too long and could therefore fit on the map. The 20 landmark pairs were the

same as those used in Experiment 1, with the exception of six landmark names that were altered to prevent the presence of cognates when translated into Estonian (Appendix D).

Maps. Each map stimulus contained two critical landmarks labeled in red font and all the remaining landmarks that made up the critical pairs in black font. Therefore, every map stimulus contained 40 total landmarks (2 critical landmarks, and 38 additional landmarks that were part of a pair), meaning all landmarks were present on the map for every trial. No additional context landmarks were included on maps in this experiment in an effort to maintain a desirable level of difficulty. During the practice phase, the maps contained two critical landmarks and six additional landmarks. Participants either learned from maps with landmark names in English (condition 1) or Estonian (condition 2). The order of map presentation was randomized within participants.

JOLs. The same JOL prompt from Experiments 1 and 2 was used. Now, they were shown the same map as they saw for the learning trial with the cue and additional landmarks present, but the target landmark absent.

Target landmark location memory test. The test phase modeled that from Experiments 1 and 2. Participants were shown the same maps that they made JOLs on (cue and additional landmarks present, target landmark absent), were reminded of the target landmark name, and were told to click on the map to indicate the correct target landmark location for 20 test trials.

Attention checks. The same attention checks from Experiment 1 and 2 were used.

Retention interval. The same retention interval from Experiment 1 and 2 was used.

Questionnaires. The same questionnaires from Experiment 1 and 2 were used. The strategies questionnaire was slightly modified.

Procedures

Practice phase. The practice phase largely modeled that from Experiments 1 and 2. Now, participants learned pairs in English or Estonian.

Learning Phase. During the experimental learning phase, participants studied 20 map stimuli in English or Estonian. Just like practice, they were instructed to learn the relative locations of the landmarks in red. Then, they made a JOL.

The ‘color’ attention check and retention interval were placed at the end of the learning phase.

Test Phase. The test phase largely mirrored the practice test implemented at the start of the experiment. Participants were shown the same 20 maps that they made JOLs about during the learning phase. They clicked on the map to indicate the correct target landmark location. The free recall of landmark names and context landmark memory tests were not included in this experiment.

The ‘1920’ attention check was placed after the test phase.

Questionnaires. After the test phase, they completed the same questionnaires from Experiments 1 and 2. At the end of the strategies questionnaire, participants completed a screening question where they reported their exposure to Estonian and neighboring languages.

Results

Data Screening

I conducted the same data screening as Experiments 1 and 2. Data from 41 participants in the English language condition and 44 participants in the Estonian condition was analyzed.

As in Experiments 1 and 2, I estimated multilevel models (MLMs) with 1) the continuous outcome variable, JOL magnitude, predicted by semantic accessibility condition, a binary

between-subjects predictor, 2) the continuous outcome variable, memory accuracy, predicted by semantic accessibility, JOL magnitude, and their interaction (i.e., metacognitive accuracy).

Model Estimation: JOL Magnitude

First, I investigated whether semantic accessibility predicted JOL magnitude. JOL ratings ranged from 0 (not likely to remember) to 10 (highly likely to remember). Learners provided an average JOL rating of 4.9 in the English condition and 4.7 in the Estonian condition. No outlier values were detected in the JOL magnitude variable.

How much of the variation in JOL magnitude is between pair locations and participants?

The first empty model included a random effect for the intercept for pair location. Note that across the two conditions, the landmark pair language was different, but the location of the pairs was the same. Therefore, pair location was entered as a random effect to assess whether the location of pairs contributed for a significant amount of variation in the outcome. This model found that pair location accounted for minimal variation in JOL magnitude and was not retained in the model. A random effect for the intercept for participant accounted for 50% of the variation in JOL magnitude and was therefore retained in the model predicting JOL magnitude.

Does semantic accessibility predict JOL magnitude?

I added a fixed effect of condition to the model to test whether JOL magnitude differed by semantic accessibility (E3J1, Appendix E). The results of this model were not significant, nor was it a better fit for the data than the empty model, meaning that knowledge of landmark name meaning did not significantly impact JOL ratings.

Planned Model Estimation: Memory & Metacognitive Accuracy

I investigated whether semantic accessibility, JOL, and their interaction (i.e., metacognitive accuracy) predicted map memory. I detected and excluded three outlier distance

values from analyses. On average, a remembered location was around 216 distance units away from the studied location in the English condition and around 258 distance units away in the Estonian condition.

How much of the variation in memory accuracy is between pairs and participants?

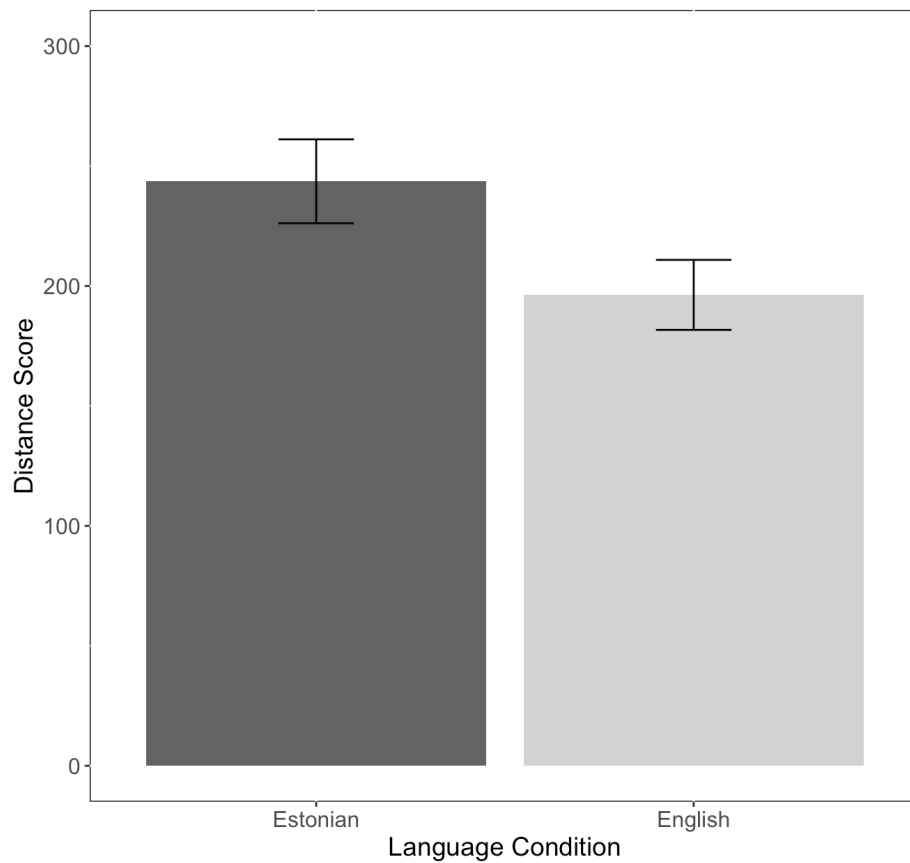
To evaluate the variation in memory accuracy by pairs and participants, I estimated an empty model with a random intercept for pair location. This model revealed that pair accounted for a minimal amount of variation in memory ($ICC = .08$). The second empty model included a random effect for the intercept for participant, revealing that 33% of the variation in memory scores was accounted for. A random intercept for participant was retained in the models predicting memory.

Does semantic accessibility predict memory accuracy?

Next, I added a fixed effect of condition to test whether memory accuracy differed by semantic accessibility (E3M1, Appendix E). Learners who studied semantically accessible maps with landmark names in English showed better memory for landmark locations, as indicated by lower distance scores compared to the Estonian condition ($\beta = -0.11$, $SE = 0.052$, $z(1697) = -2.09$, $p < .05$) (Figure 14). This model was a better fit for the data than the empty model based on AIC and log likelihood value comparisons ($(\chi^2(1) = 4.26, p < .05)$).

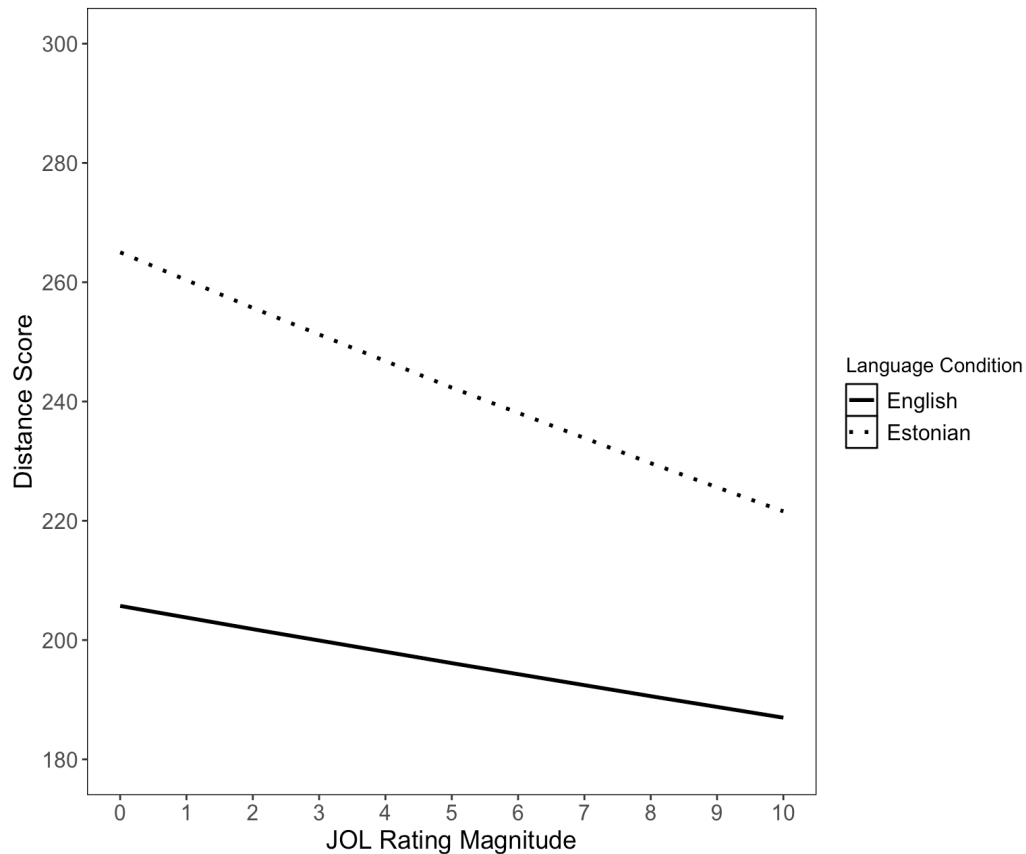
Figure 14.

Memory by Semantic Accessibility (Language) Condition



Do JOLs predict memory accuracy above and beyond semantic accessibility?

To determine whether JOLs predicted memory and whether predictive accuracy varied by condition, I estimated a model with a fixed effect of condition and its interaction with JOL magnitude (E3M2, Appendix E). This model was not a better fit for the data than the model with condition alone. Visual inspection of the plot revealed that across both conditions, higher JOL ratings were associated lower distance scores (Figure 15).

Figure 15.*Distance Scores: Language Condition*JOL Rating Magnitude*

Exploratory Analyses

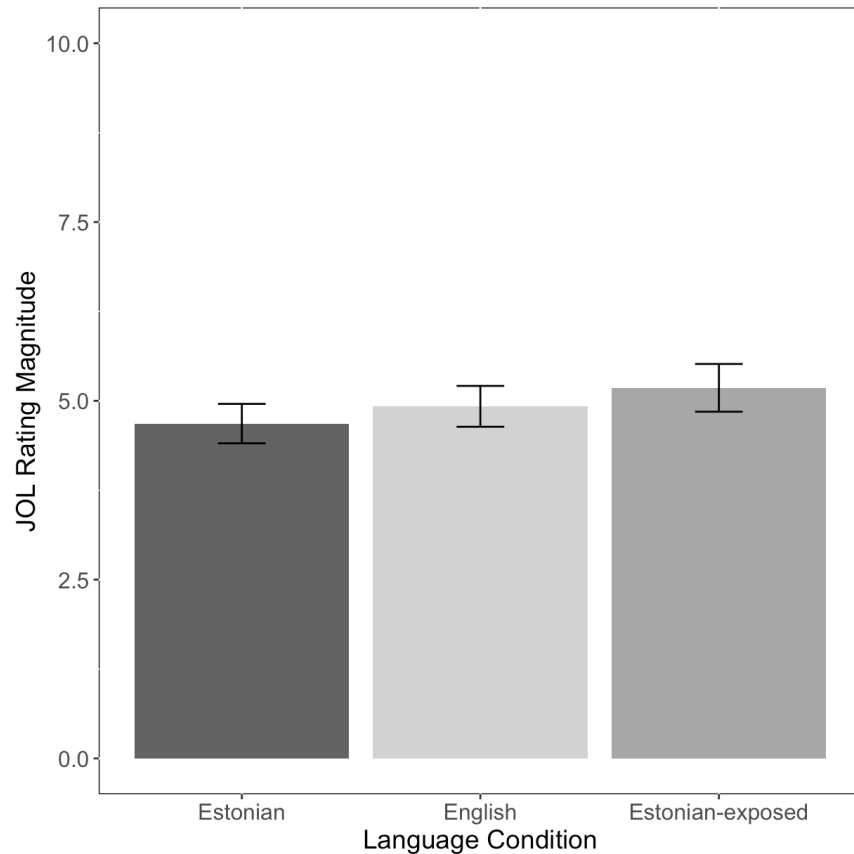
Data from 30 participants who were 1) in the Estonian map condition, and 2) reported being familiar with Estonian and/or one or more of the following languages: Estonian, Finno-Ugric languages, Slavic language, Scandinavian languages, or West Germanic languages, was retained, sorted into a “language exposed” group, and analyzed for exploratory purposes. Specifically, I was interested in whether any familiarity with Estonian or neighboring languages resulted in observable differences in metacognitive predictions, memory, or the relationship

between the two. Given the familiarity, I predicted that participants in this language exposed condition would show higher JOLs, memory, and metacognitive accuracy compared to the Estonian condition.

I re-estimated the model predicting JOL magnitude, now using a three-group condition variable (English, Estonian, and Estonian-exposed). Learners in the Estonian-exposed condition provided a JOL rating of 5.2 on average. No outlier values were detected. A random effect for the intercept for participant accounted for 50% of the variation in JOL magnitude. Adding a fixed effect of condition resulted in a non-significant model (E3JE1, Appendix E, Figure 16), meaning that knowledge of landmark name meaning did not significantly impact JOL ratings, even when a rough measure of Estonian language exposure was taken into consideration.

Figure 16.

JOL Rating Magnitude by Semantic Accessibility (Language) Conditions

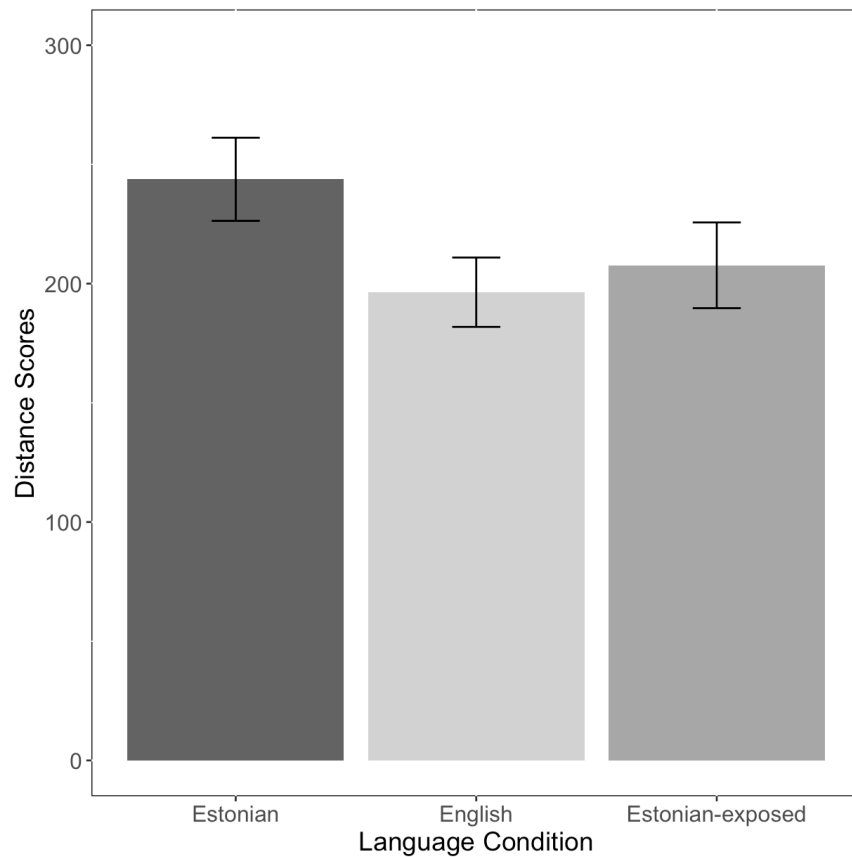


Next, I investigated memory. A total of six outlier observations were detected and excluded from analyses. On average, a remembered location was around 225 distance units away from the studied location in the Estonian-exposed condition. A random effect for the intercept for participant accounted for 31% of the variation in distance scores. I entered condition into a model predicting distance scores (E3ME1, Appendix E). In alignment with the two-group memory model, the English condition performed better on the memory task than the Estonian

condition ($\beta = 0.22$, $SE = 0.103$, $z(2294) = 2.10$, $p < .05$) (Figure 17), but the model was not a better fit for the data as compared to the empty model.

Figure 17.

Memory by All Semantic Accessibility (Language) Conditions

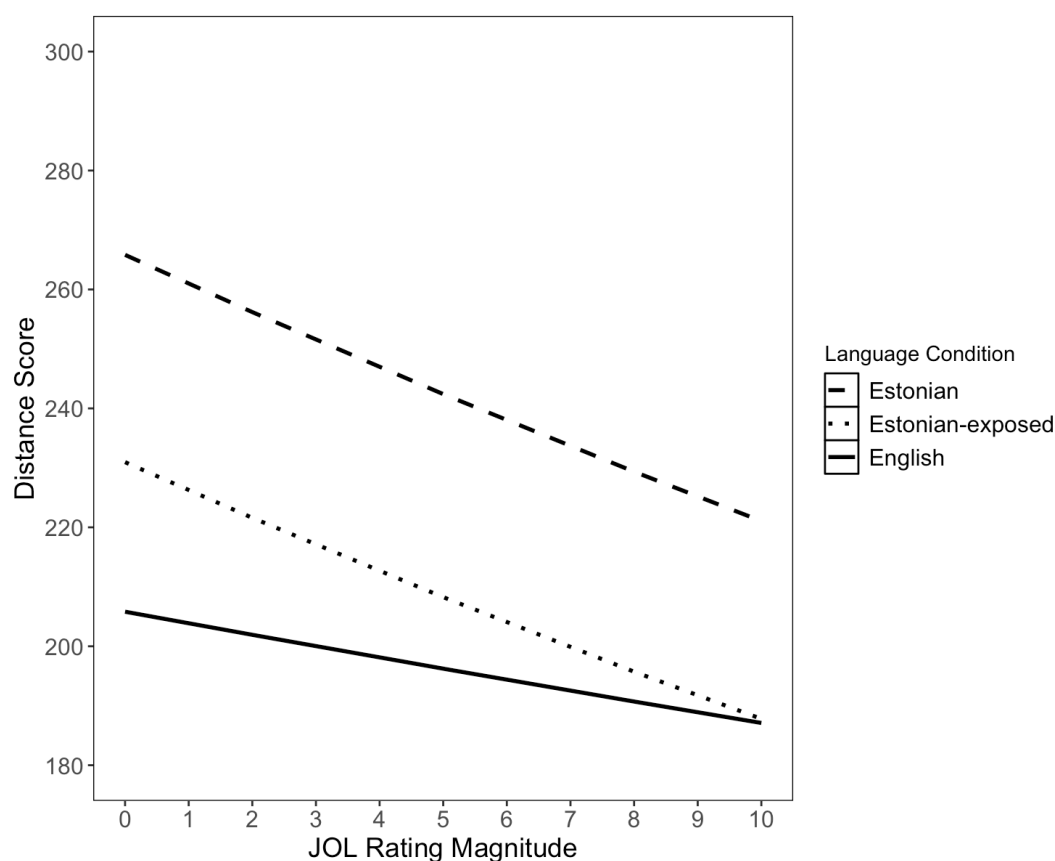


I proceeded by adding JOL and its interaction with language condition into the model predicting memory to explore metacognitive accuracy across the three conditions. The results of this model (E3ME2, Appendix E) were non-significant (Figure 18), meaning that knowledge of

landmark name meaning did not significantly predict landmark location memory, nor did the predictive value of JOLs significantly vary across language conditions.

Figure 18.

*Distance Scores: JOL Rating Magnitude*Semantic Accessibility (Language) Conditions*



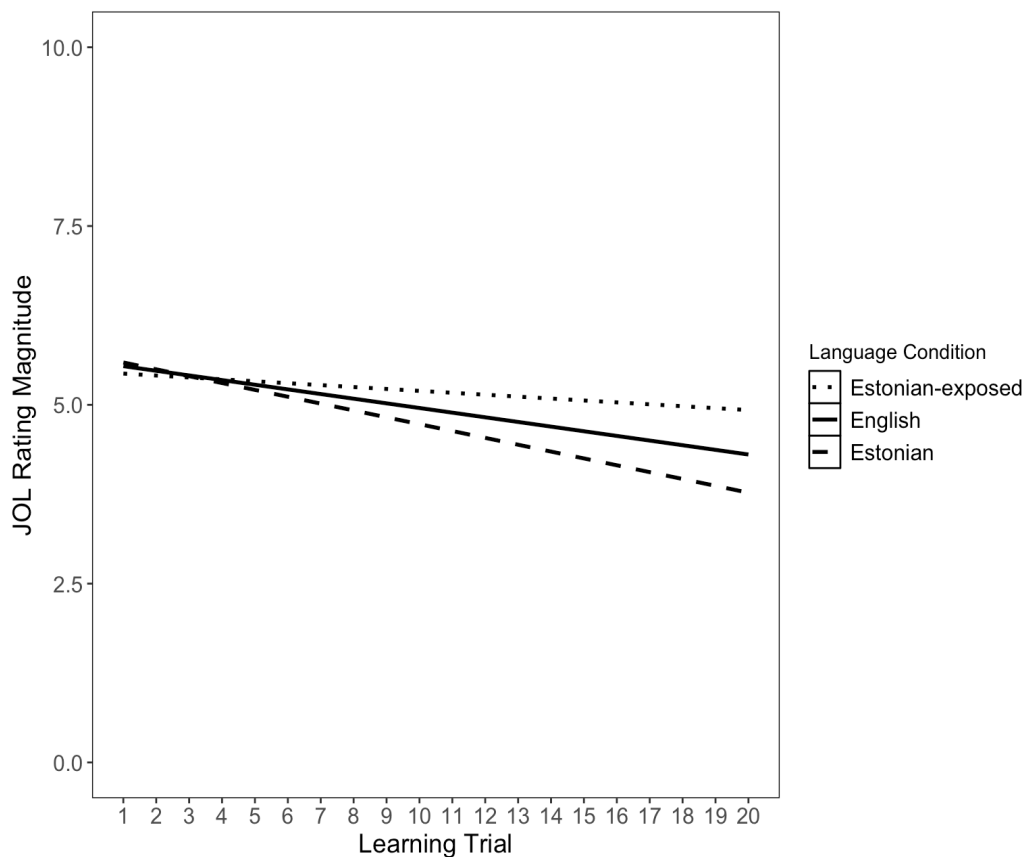
JOL Magnitude Across the Learning Span

I estimated exploratory models to assess whether JOL magnitudes decreased across learning trials and whether this pattern varied as a function of condition. I allowed for a random effect for 1) the intercept for participant and 2) the slope for learning trial. This model (E3JE3, Appendix E) was a better fit for the data as compared to a model with only a random intercept

for participant ($\chi^2(2) = 122.85, p < .001$) or the empty model ($\chi^2(5) = 129.86, p < .001$). I observed a significant learning trial*condition interaction, such that JOL rating magnitudes decreased at the steepest rate for the Estonian condition ($\beta = -0.07, SE = 0.03, z(2300) = -2.48, p < .05$). Across all conditions, learners rated themselves less likely to remember landmark locations as the learning span progressed (Figure 19).

Figure 19.

JOL Rating Magnitude by Learning Trial and Semantic Accessibility (Language) Conditions



Reaction Time

I conducted exploratory analyses on reaction times (RTs) to assess whether a significant amount of variation in RT was accounted for by pair and participant and whether condition, JOL magnitude, and their interaction predicted RT on the target landmark location memory test. I identified and excluded eight outlier RT values from analyses. On average, participants took around 7 seconds to indicate their response for each test map.

The first empty model included a random effect for the intercept for pair location, which did not account for variation in RT. The second included a random effect for the intercept for participant, revealing that 48% of the variation in RT was accounted for by this random effect. A random intercept for participant was therefore retained in the model predicting RT. I added a fixed effect of condition and its interaction with JOL magnitude. This model (E3RT1, Appendix E) was a better fit for the data as compared to the empty model ($\chi^2(5) = 13.58, p < .05$) and produced a main effect of JOL ($\beta = -0.14, SE = 0.007, z(2300) = -2.02, p < .05$). Higher JOLs at learning were associated with lower response times at test, independent of condition.

Santa Barbara Sense of Direction

Next, I investigated whether scores on the Santa Barbara Sense of Direction (SBSOD) scale were correlated with two primary outcome measures of interest: memory accuracy or JOL rating magnitude, in addition to one exploratory measure of interest: gender.

Pearson correlations revealed that SBSOD scores were correlated with JOLs ($r(113) = .26, p = .005$), but not memory accuracy or gender. I proceeded by adding a fixed effect of SBSOD score into the empty JOL model (E3JE4, Appendix E). This model was a better fit for the data as compared to the empty model ($\chi^2(1) = 8.00, p < .005$) and produced a main effect of SBSOD score ($\beta = 0.35, SE = 0.12, z(2300) = 2.88, p < .005$). I proceeded by adding a fixed effect of condition

and its interaction with SBSOD to test whether the relationship between SBSOD scores and JOL rating magnitude varied by semantic accessibility. This model (E3JE5, Appendix E) was not a better fit for the data based on log likelihood value comparisons. Independent of condition, learners with higher SBSOD scores produced higher JOL rating magnitudes.

Discussion

In Experiment 3, I investigated the role of semantic accessibility, or knowledge of landmark name meaning for metacognitive judgment formation, memory, and metacognitive or predictive accuracy. Learners' metacognitive ratings did not vary by language condition. This finding aligns with prior claims (Dechêne et al., 2010; Schwarz, 2018), suggesting that to derive differences in metacognitive rating magnitude, it is necessary for learners to compare their encoding experience between stimulus types. Given that they studied landmark names in English or Estonian, such a within-task comparison was not possible. In contrast, memory accuracy did vary, with better memory among learners who studied English maps. This result corresponds to prior research that found English participants had worse memory for Italian words, which were unfamiliar to them, as compared to English words (Hulme et al., 1991). Critically, when participants were taught the translation of the Italian words, their memory span was improved. In other words, knowledge of word meaning helped bolster memory for the words. Results from the present experiment expand on this discovery from Hulme and colleagues (1991) by adding a spatial component.

Earlier work also found that those with a better spatial understanding of an environment could more accurately place semantic components, or objects by name, in their assigned locations (Wessel et al., 2010). The memory result from the present experiment offers further evidence for a link between semantic and spatial knowledge but finds it in the opposite direction.

Namely, access to semantic information, or landmark name meaning, supported memory for landmark locations. To further disentangle the role of semantic accessibility for the outcome variables of interest, I also explored the impact of rough language exposure. Learners who 1) studied Estonian maps, and 2) reported having exposure to Estonian or a neighboring language, were sorted into a third group and analyzed for exploratory purposes. Model trends suggested that the worst memory was observed within the Estonian condition, followed by Estonian-exposed, and then the English condition. This exploratory observation supports the notion that any knowledge of landmark name meaning, even when operationalized by rough exposure, may support landmark location memory.

In sum, in Experiment 3, learners who studied maps with landmark names in English showed better memory for landmark locations as compared to those who studied names in Estonian. Data visualizations suggested that all learners were relatively accurate at predicting their memory for landmark locations, as evident from positive JOL-memory relationships, though the metacognitive accuracy models did not produce significance. In alignment with Experiments 1 and 2, learners rated themselves less likely to remember landmark locations across the learning span, independent of language condition. When a rough measure of language exposure was included in exploratory models, the observed trends offered support for the beneficial role of landmark name meaning accessibility during map-based learning.

General Discussion

The present study assessed the role of map display feature manipulations for metacognitive judgment formation, memory, and metacognitive accuracy, or the relationship between judgments and memory. Map complexity, as operationalized by the number of landmarks present on the map, and salience, as determined by the visual transparency of the

goal-relevant landmarks, each impacted JOL rating magnitude. Learners deemed themselves more likely to remember landmark pairs studied on less complex maps and salient maps. These findings align with verbal learning results, now suggesting that learners are sensitive to perceptual aspects of map-based stimuli, specifically. The complexity result also corresponds to prior work which found that learners assigned higher JOLs to maps with less complex routes between landmarks (Mason et al., 2024). However, it comes in contrast to prior spatial metacognition work with map displays, which found that learners attributed higher metacognitive ratings to more detailed maps (Stevens & Carlson, 2019). Collectively, these findings support metacognitive monitoring judgments' sensitivity to map displays and suggest that variations in the way complexity is operationalized can guide observed rating outcomes.

Despite differences in JOL ratings between map types in Experiments 1 and 2, memory accuracy did not vary. This observation raises questions about the utility of map-based cues for metacognitive judgment formation. Based on transfer appropriate monitoring (TAM) hypotheses, judgments should be generally accurate if the cues referred to at the time of the judgment are relevant for memory later at test (Dunlosky et al., 2005; Koriat, 1997). Results from Experiments 1 and 2 suggested that learners' judgments were sensitive to perceptual display cues, namely map complexity and salience, but these cues were not necessarily predictive of memory. It is possible that learners perceived less complex and salient maps as easier to learn during study (Dobson, 1980; Klippel et al., 2006; Phillips & Noyes, 1982), thereby assigning higher ratings to these maps, but ease of learning was not meaningful for memory outcomes (Miller & Geraci, 2016; Rhodes & Castel, 2008). Future research should consider whether observed JOL rating magnitude differences also guide meaningful metacognitive control decisions during map learning. For instance, determining whether learners who rate themselves less likely to remember

high complexity maps will also choose to restudy high complexity maps can help elucidate the practical utility of these judgments for spatial visual displays. Similarly, learners who rate themselves more likely to remember salient versus standard maps may prefer to use salient maps during environment learning. Future research should explicitly assess whether honoring such preferences results in more effective learning over time, and whether such preferences correspond to other individual differences, such as self-reported sense of direction ability.

In Experiment 2, an important interaction was observed, such that higher JOLs were associated with better location memory for salient maps, specifically. This finding highlights the nuances of metacognitive monitoring during map learning. Though learners' judgments distinguished between map types and memory accuracy did not, the relationship between JOL and ratings was stronger for salient maps. This point raises an important issue: How metacognitive accuracy is quantified and interpreted has implications for uncovering the dynamic nature of metacognitive monitoring in spatial contexts. With complex visuospatial stimuli such as maps, it is challenging to determine what cues learners are referring to when predicting their likelihood to remember map display components in the future. These cues could pertain to components of the map itself, differences between map types, and elements independent of the map display, such as their self-perceived spatial abilities broadly. In Experiment 2, when goal-relevant map display components were made salient, the relationship between JOLs and later memory was stronger. Taken together, presenting map displays with salient goal-relevant landmark information may prompt learners to engage in more effective metacognitive monitoring processes during learning, thereby resulting in more accurate memory predictions.

Across Experiments 1 and 2, learners better remembered target landmark names, followed by context, and then cue landmarks. This pattern may have resulted from prompting learners to make target-absent JOLs, where they rated their likelihood to remember the relative locations of the studied landmarks and were only presented with the cue landmark at the time of judgment. This protocol could have urged learners to actively maintain the target landmark name in mind while forming the subsequent JOL, and such active rehearsal may have supported superior memory for the target landmark names (Ariel et al., 2021; Karpicke & Roediger, 2008; Smith et al., 2016; Thomas & McDaniel, 2007). Context landmarks appeared on all maps, thereby justifying the rate at which learners produced context landmark names on the free recall test. The relatively low memory for cue landmark names again raises questions around the utility of map-based cues during learning. Namely, learners may have paid little attention to “cue” landmarks during JOL formation. This possibility offers further support for the principle that map displays are complex, riddled with cues, and require holistic consideration when investigating the information learners integrate while monitoring their learning (Mason et al., 2022, 2024).

In Experiment 3, learners who studied landmark names in their native language (English), and therefore had access to the semantic meaning of landmark names, had better memory for landmark locations as compared to those who studied landmarks in a language foreign to them. This discovery offers a novel contribution to the literature on map learning, providing evidence for a link between semantic and spatial landmark components for memory. In addition to theoretical insight, this discovery has applied implications. For instance, first-responders in foreign territory may benefit from studying maps with landmark names in English and the foreign language translation to support their memory for the spatial environment.

As observed in prior work (Mason et al., 2024), learners across all three experiments rated themselves less likely to remember landmark locations as the learning span progressed, suggesting that cognitive load resulting from information accumulation influenced judgment formation. Furthermore, individual difference findings suggested that learners' self-reported SOD ability was predictive of JOL rating magnitudes (Experiments 1 and 3) and memory accuracy (Experiment 2). These findings are unique, given that Santa Barbara Sense of Direction (SBSOD) scale scores tend to correlate more highly with knowledge acquired from direct environment experience as compared to that acquired with maps (Hegarty et al., 2002). The present results suggest that the SBSOD measure may also be informative in the context of map-based metacognition tasks.

Collectively, the present study suggests that JOL ratings are sensitive to variations in map features, such as landmark complexity and salience, in addition to subjective components, such as information accumulation across learning and self-rated SOD ability. These conclusions offer support to previous work with simpler materials (images of real-world objects) that found JOLs stemmed from stimulus-intrinsic components, such as ease of encoding, and individual-specific encoding variables, such as flexible cognitive states and stimulus processing (Saito et al., 2022). In other words, learners integrate various streams of information when forming JOLs. Future research could implement a think-aloud paradigm during spatial learning to gather rich qualitative data about learners' rationale during judgment formation. Such information may further elucidate the multifaceted cognitive processes contributing to metacognitive ratings during spatial knowledge acquisition.

Metacognitive monitoring is important, because it can influence one's decision to engage in effective metacognitive control processes and in turn prevent people from becoming spatially

disoriented or lost. Therefore, future research should also consider whether map display variations not only result in unique metacognitive monitoring signatures, but also benefit metacognitive control processes throughout learning. Taken together, evidence from the present work offers further motivation for the development of an integrative cue-utility model to demonstrate the dynamic information learners consider when forming metacognitive monitoring judgments during complex spatial visual display learning (Mason et al., 2022).

Limitations

The present work provides early evidence on the role of map display features for metacognitive judgment formation. This said, several limitations should be acknowledged. In Experiment 1, there were many ways complexity could be operationalized. Prior work has utilized grid lines (Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014; Keil et al., 2020), verbal and graphical detail (Brunyé et al., 2007), turns in routes (Mason et al., 2024), and JPEG file size (Bestgen et al., 2017; Donderi & McFadden, 2005; Edler, Bestgen, et al., 2014; Edler, Dickmann, et al., 2014) to operationalize map complexity, to name a few approaches. In the present study, I adopted the number of landmarks on the map as my definition, and I subsequently confirmed this distinction with JPEG file sizes. Future work should aim to replicate differences in JOL rating magnitude with varying map complexity measures to determine whether the present results are dependent upon the number of landmarks alone, or other perceptual complexity components.

Interestingly, the present study found that learners rated themselves more likely to remember less complex maps. This finding falls contrary to spatial literature, which suggests that additional reference points on maps by serve to support spatial mental model development. Testing whether the presence or absence of grid lines on a map produces systematic differences

in JOL rating magnitudes may be a fruitful path forward in elucidating visual complexity's dynamic role in metacognitive judgment formation during map learning, broadly. Furthermore, recent research (Cheng et al., 2022) investigated how the number of landmarks on a map (i.e., 3 vs. 5 vs. 7 landmarks) impacted spatial learning, cognitive load, and encoding. Findings suggested that learning improved when the number of landmarks increased from three to five, but no further benefits were observed for seven landmarks. Similar research should be conducted with JOLs as the outcome variable to gauge the threshold for which map complexity drives shifts in metacognitive judgments, and whether this threshold varies across individuals.

On half of the map stimuli in Experiment 2, critical landmark pairs were made visually salient, such that the remaining context landmarks were reduced in transparency by 70%. This manipulation produced differences in JOL rating magnitude, such that learners deemed themselves more likely to remember landmarks studied on salient maps. Differences in memory, however, were not observed. It remains unknown whether this salience manipulation did in fact produce variations in attentional resource allocation, or whether it simply produced differences in subjective ease of processing, which were not meaningful for encoding or memory outcomes. Future work may repeat the present task in combination with eye tracking to determine whether unique viewing patterns are observed for salient versus standard maps, and whether these patterns align with metacognitive monitoring signatures.

Findings from Experiment 3 offer a novel perspective to the map learning literature, such that knowledge of landmark name meaning supported landmark location memory. The language used in the present study was Estonian. This language was elected given that it used the same alphabet as English, the names could fit on the map, and it contained a limited number of cognates with English. Future research should aim to replicate the memory finding with other

languages to determine whether the present result has broader implications. Furthermore, in contrast to Experiments 1 and 2, differences in JOL rating magnitude were not observed, potentially because learners could not compare their encoding experience across stimulus types. Therefore, future research should also establish a within-subjects study design that would allow learners to compare across stimulus type. Such comparison could reveal whether semantic accessibility, or knowledge of landmark name meaning, influences the metacognitive judgment formation process in addition to subsequent memory.

Conclusion

Modern technological advancements make finding our destination easier than ever before, with navigation tools guiding our every turn. Yet, navigational aids can fail. Further, findings suggest that referring to a map before navigating an environment benefits learning accuracy compared to only navigating (Meneghetti & Pazzaglia, 2021). In sum, maps maintain practical value even in a technologically-guided era. Despite their observed relationships with memory, however, how learners predict what they will remember from maps remains an understudied question. This question is important, because accurate spatial memory predictions may prevent people from becoming lost. Therefore, the present study explored the role of three map features, namely landmark complexity, salience, and semantic accessibility for metacognitive judgment formation, memory, and metacognitive accuracy.

Findings from this work suggest that learners' metacognitive judgments are sensitive to landmark complexity and salience, and salient map displays result in more accurate predictions of memory. Finally, knowledge of landmark name meaning impacts memory for landmark locations. Collectively, these results suggest that classic JOL findings from verbal learning literature may extend to more complex visuospatial stimuli, such that subjective ease of

processing can impact judgment formation, and these variations in JOL magnitude do not always translate to meaningful differences in memory. Further, knowledge of landmark name meaning aids landmark location memory, offering interesting theoretical insights to the map learning literature, broadly. In addition to map-based components, monitoring judgments also appeared sensitive to characteristics of the learners' experience, such as information accumulation across learning span, and self-reported sense of direction ability. In sum, the present results highlight that further work is warranted to establish a more nuanced understanding of metacognition's dynamic nature during spatial learning, in addition to how semantic information contributes to location memory.

References

- Allen, G. L. (1981). A developmental perspective on the effects of “subdividing” macrospatial experience. *Journal of Experimental Psychology: Human Learning and Memory*, 7(2), 120–132.
- Ariel, R., Karpicke, J. D., Witherby, A. E., & Tauber, S. K. (2021). Do judgments of learning directly enhance learning of educational materials? *Educational Psychology Review*, 33(2), 693–712. <https://doi.org/10.1007/s10648-020-09556-8>
- Ariel, R., Lembeck, N. A., Moffat, S., & Hertzog, C. (2018). Are there sex differences in confidence and metacognitive monitoring accuracy for everyday, academic, and psychometrically measured spatial ability? *Intelligence*, 70, 42–51. <https://doi.org/10.1016/j.intell.2018.08.001>
- Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, 28(5), 610–632. [https://doi.org/10.1016/0749-596X\(89\)90016-8](https://doi.org/10.1016/0749-596X(89)90016-8)
- Bestgen, A.-K., Edler, D., Müller, C., Schulze, P., Dickmann, F., & Kuchinke, L. (2017). Where is it (in the map)? Recall and recognition of spatial information. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 52(1), 80–97. <https://doi.org/10.3138/cart.52.1.3636>
- Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In *In J. Metcalfe & A. P. Shimamura (Eds.) Metacognition: Knowing about knowing* (pp. 185–205). MIT Press.
- Britton, B. K., Meyer, B. J. F., Simpson, R., Holdredge, T. S., & Curry, C. (1979). Effects of the organization of text on memory: Tests of two implications of a selective attention

- hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, 5(5), 496–506.
- Brunyé, T. T., Gardony, A. L., Holmes, A., & Taylor, H. A. (2018). Spatial decision dynamics during wayfinding: Intersections prompt the decision-making process. *Cognitive Research: Principles and Implications*, 3(1), 13. <https://doi.org/10.1186/s41235-018-0098-3>
- Brunyé, T. T., Smith, A. M., Hendel, D., Gardony, A. L., Martis, S. B., & Taylor, H. A. (2020). Retrieval practice enhances near but not far transfer of spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(1), 24–45. <https://doi.org/10.1037/xlm0000710>
- Brunyé, T. T., Taylor, H. A., & Worboys, M. (2007). Levels of detail in descriptions and depictions of geographic space. *Spatial Cognition & Computation*, 7(3), 227–266. <https://doi.org/10.1080/13875860701515472>
- Bunch, R. L., & Lloyd, R. E. (2006). The cognitive load of geographic information. *The Professional Geographer*, 58(2), 209–220. <https://doi.org/10.1111/j.1467-9272.2006.00527.x>
- Caduff, D., & Timpf, S. (2008). On the assessment of landmark salience for human navigation. *Cognitive Processing*, 9(4), 249–267. <https://doi.org/10.1007/s10339-007-0199-2>
- Cartwright, W., Peterson, M. P., & Gartner, G. F. (Eds.). (1999). *Multimedia cartography*. Springer.
- Chan, J. C. K., Manley, K. D., Davis, S. D., & Szpunar, K. K. (2018). Testing potentiates new learning across a retention interval and a lag: A strategy change perspective. *Journal of Memory and Language*, 102, 83–96. <https://doi.org/10.1016/j.jml.2018.05.007>

- Cheng, B., Wunderlich, A., Gramann, K., Lin, E., & Fabrikant, S. I. (2022). The effect of landmark visualization in mobile maps on brain activity during navigation: A virtual reality study. *Frontiers in Virtual Reality*, 3, 981625.
<https://doi.org/10.3389/frvir.2022.981625>
- DeCaro, R., & Thomas, A. K. (2020). Prompting retrieval during monitoring and self-regulated learning in older and younger adults. *Metacognition and Learning*, 15(3), 367–390.
<https://doi.org/10.1007/s11409-020-09230-y>
- Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14(2), 238–257. <https://doi.org/10.1177/1088868309352251>
- Dobson, M. W. (1980). The Influence of the amount of graphic information on visual matching. *The Cartographic Journal*, 17(1), 26–32. <https://doi.org/10.1179/caj.1980.17.1.26>
- Donderi, D. C., & McFadden, S. (2005). Compressed file length predicts search time and errors on visual displays. *Displays*, 26(2), 71–78. <https://doi.org/10.1016/j.displa.2005.02.002>
- Dunlosky, J., Rawson, K. A., & Middleton, E. L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language*, 52(4), 551–565.
<https://doi.org/10.1016/j.jml.2005.01.011>
- Edler, D., Bestgen, A.-K., Kuchinke, L., & Dickmann, F. (2014). Grids in topographic maps reduce distortions in the recall of learned object locations. *PLoS ONE*, 9(5), e98148.
<https://doi.org/10.1371/journal.pone.0098148>
- Edler, D., Dickmann, F., Bestgen, A.-K., & Kuchinke, L. (2014). The effects of grid line separation in topographic maps for object location memory. *Cartographica: The*

- International Journal for Geographic Information and Geovisualization*, 49(4), 207–217.
<https://doi.org/10.3138/carto.49.4.2674>
- Garden, S., Cornoldi, C., & Logie, R. H. (2002). Visuo-spatial working memory in navigation. *Applied Cognitive Psychology*, 16(1), 35–50. <https://doi.org/10.1002/acp.746>
- Gardony, A. L., Hendel, D. D., & Brunyé, T. T. (2022). Identifying optimal graphical level of detail to support orienting with 3D geo-visualizations. *Spatial Cognition & Computation*, 22(1–2), 135–160. <https://doi.org/10.1080/13875868.2021.1892696>
- Hegarty, M., Canham, M. S., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(1), 37–53.
<https://doi.org/10.1037/a0017683>
- Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, L. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5), 425–447. [https://doi.org/10.1016/S0160-2896\(02\)00116-2](https://doi.org/10.1016/S0160-2896(02)00116-2)
- Hegarty, M., Smallman, H. S., & Stull, A. T. (2012). Choosing and using geospatial displays: Effects of design on performance and metacognition. *Journal of Experimental Psychology: Applied*, 18(1), 1–17. <https://doi.org/10.1037/a0026625>
- Hirtle, S. C., & Mascolo, M. F. (1986). Effect of semantic clustering on the memory of spatial locations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(2), 182–189.
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30(6), 685–701. [https://doi.org/10.1016/0749-596X\(91\)90032-F](https://doi.org/10.1016/0749-596X(91)90032-F)

- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194–203. <https://doi.org/10.1038/35058500>
- Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966–968. <https://doi.org/10.1126/science.1152408>
- Keil, J., Edler, D., Kuchinke, L., & Dickmann, F. (2020). Effects of visual map complexity on the attentional processing of landmarks. *PLOS ONE*, 15(3), e0229575. <https://doi.org/10.1371/journal.pone.0229575>
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1
- Klippel, A. (2003). Wayfinding choremes. In W. Kuhn, M. F. Worboys, & S. Timpf (Eds.), *Spatial Information Theory. Foundations of Geographic Information Science* (Vol. 2825, pp. 301–315). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-39923-0_20
- Klippel, A., Freksa, C., & Winter, S. (2006). You-are-here maps in emergencies –the danger of getting lost. *Journal of Spatial Science*, 51(1), 117–131. <https://doi.org/10.1080/14498596.2006.9635068>
- Knight, M. J., & Tlauka, M. (2017). Interactivity in map learning: The effect of cognitive load. *Spatial Cognition & Computation*, 17(3), 185–198. <https://doi.org/10.1080/13875868.2016.1211661>
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126(4), 349–370. <https://doi.org/10.1037/0096-3445.126.4.349>

- Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 1133–1145. <https://doi.org/10.1037/0278-7393.32.5.1133>
- Kozlowski, L. T., & Bryant, K. J. (1977). Sense of direction, spatial orientation, and cognitive maps. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 590–598. <https://doi.org/10.1037/0096-1523.3.4.590>
- LaBerge, D. (1995). *Attentional Processing: The brain's art of mindfulness*. Harvard University Press.
- Mason, L. A., Thomas, A. K., & Taylor, H. A. (2022). On the proposed role of metacognition in environment learning: Recommendations for research. *Cognitive Research: Principles and Implications*, 7(1), 104. <https://doi.org/10.1186/s41235-022-00454-x>
- Mason, L. A., Thomas, A. K., & Taylor, H. A. (2024). Metacognitive judgment formation during map learning: Evidence for global monitoring. *Cognition*, 246, 105743. <https://doi.org/10.1016/j.cognition.2024.105743>
- Mayer, R. E. (2001). *Multi-media Learning*. Cambridge University Press.
- Meneghetti, C., & Pazzaglia, F. (2021). Navigating in virtual environments: Does a map or a map-based description presented beforehand help? *Brain Sciences*, 11(6), 773. <https://doi.org/10.3390/brainsci11060773>
- Merrill, A. A., & Baird, J. C. (1987). Semantic and spatial factors in environmental memory. *Memory & Cognition*, 15(2), 101–108. <https://doi.org/10.3758/BF03197021>
- Michon, P. E., & Denis, M. (2001). *Spatial information theory: Foundations of geographic information science: international conference, COSIT 2001, Morro Bay, CA, USA*,

- September 19-23, 2001: proceedings* (D. R. Montello, Ed.). COSIT 2001, Berlin ; New York. Springer.
- Miller, T. M., & Geraci, L. (2016). The influence of retrieval practice on metacognition: The contribution of analytic and non-analytic processes. *Consciousness and Cognition*, 42, 41–50. <https://doi.org/10.1016/j.concog.2016.03.010>
- Montello, D. R. (2002). Cognitive map-design research in the twentieth century: Theoretical and empirical approaches. *Cartography and Geographic Information Science*, 29(3), 283–304. <https://doi.org/10.1559/152304002782008503>
- Murayama, K., Sakaki, M., Yan, V. X., & Smith, G. M. (2014). Type I error inflation in the traditional by-participant analysis to metamemory accuracy: A generalized mixed-effects model perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1287–1306. <https://doi.org/10.1037/a0036914>
- Nelson, & Narens. (1994). *Why investigate metacognition*.
<https://doi.org/10.7551/mitpress/4561.003.0003>
- Nelson, T. O. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 26, pp. 125–173). Academic Press.
[https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- Pezdek, K., & Evans, G. W. (1979). Visual and verbal memory for objects and their spatial locations. *Journal of Experimental Psychology: Human Learning and Memory*, 5(4), 360–373.
- Phillips, R. J., & Noyes, L. (1977). Searching for names in two city street maps. *Applied Ergonomics*, 8(2), 73–77. [https://doi.org/10.1016/0003-6870\(77\)90056-4](https://doi.org/10.1016/0003-6870(77)90056-4)

- Phillips, R. J., & Noyes, L. (1982). An investigation of visual clutter in the topographic base of a geological map. *The Cartographic Journal*, 19(2), 122–132.
<https://doi.org/10.1179/caj.1982.19.2.122>
- Pichert, J. W., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology*, 69(4), 309–315.
- Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, 137(4), 615–625. <https://doi.org/10.1037/a0013684>
- Saito, J. M., Kolisnyk, M., & Fukuda, K. (2022). Judgments of learning reveal conscious access to stimulus memorability. *Psychonomic Bulletin & Review*.
<https://doi.org/10.3758/s13423-022-02166-1>
- Schwarz, N. (2018). *Offluency, beauty, and truth* (Vol. 1). Oxford University Press.
<https://doi.org/10.1093/oso/9780198789710.003.0002>
- Scott, B. M., & Schwartz, N. H. (2007). Navigational spatial displays: The role of metacognition as cognitive load. *Learning and Instruction*, 17(1), 89–105.
<https://doi.org/10.1016/j.learninstruc.2006.11.008>
- Shamon, H., & Berning, C. C. (2020). Attention check items and instructions in online surveys: Boon or bane for data quality? *Survey Research Methods*, 55-77 Pages.
<https://doi.org/10.18148/SRM/2020.V14I1.7374>
- Siegel & White. (1975). The development of spatial representations of large-scale environments. *Advances in Child Development and Behavior*, 10, 9–55. [https://doi.org/doi:10.1016/s0065-2407\(08\)60007-5](https://doi.org/doi:10.1016/s0065-2407(08)60007-5)

- Smith, A. M., Floerke, V. A., & Thomas, A. K. (2016). Retrieval practice protects memory against acute stress. *Science*, 354(6315), 1046–1048.
<https://doi.org/10.1126/science.aah5067>
- Son, L. K., & Metcalfe, J. (2005). Judgments of learning: Evidence for a two-stage process. *Memory & Cognition*, 33(6), 1116–1129. <https://doi.org/10.3758/BF03193217>
- Stevens, C. A., & Carlson, R. A. (2019). Is simpler always better? Effects of perceptual detail and viewpoint on spatial cognition and metacognition. *The American Journal of Psychology*, 132(3), 293–302.
- Taylor, H. A., Naylor, S. J., & Chechile, N. A. (1999). Goal-specific influences on the representation of spatial perspective. *Memory & Cognition*, 27(2), 309–319.
<https://doi.org/10.3758/BF03211414>
- Thomas, A. K., Bonura, B. M., & Taylor, H. A. (2012). The influence of semantic relationships on older adult map memory. *Psychology and Aging*, 27(3), 657–665.
<https://doi.org/10.1037/a0028504>
- Thomas, A. K., & McDaniel, M. A. (2007). The negative cascade of incongruent generative study-test processing in memory and metacomprehension. *Memory & Cognition*, 35(4), 668–678. <https://doi.org/10.3758/BF03193305>
- Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. *Spatial Information Theory: A Theoretical Basis for GIS*, 14–24.
- Ungerleider, L. (1994). “What” and “where” in the human brain. *Current Opinion in Neurobiology*, 4(2), 157–165. [https://doi.org/10.1016/0959-4388\(94\)90066-3](https://doi.org/10.1016/0959-4388(94)90066-3)
- Valcke, M. (2002). Cognitive load: Updating the theory? *Learning and Instruction*, 8.

van Merriënboer, J. J. G., Schuurman, J. G., de Croock, M. B. M., & Paas, F. G. W. C. (2002).

Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction*, 12(1), 11–37.

[https://doi.org/10.1016/S0959-4752\(01\)00020-2](https://doi.org/10.1016/S0959-4752(01)00020-2)

Weisberg, S. M., & Newcombe, N. S. (2016). How do (some) people make a cognitive map?

Routes, places, and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(5), 768–785. <https://doi.org/10.1037/xlm0000200>

Wessel, G., Ziemkiewicz, C., Chang, R., & Sauda, E. (2010). GPS and road map navigation: The

case for a spatial framework for semantic information. *Proceedings of the International Conference on Advanced Visual Interfaces*, 207–214.

<https://doi.org/10.1145/1842993.1843030>

Yeh, M., & Wickens, C. D. (2001). Attentional filtering in the design of electronic map displays:

A comparison of color coding, intensity coding, and decluttering techniques. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(4), 543–562.

<https://doi.org/10.1518/001872001775870359>

Appendix A.**Landmark name pairs (Experiment 1 & 2)**

List A Pairs	
Movie Theatre	Zoo
Bank	Auditorium
Pharmacy	Art Gallery
Car Wash	Post Office
Train Station	Realtor
Gym	Hotel
Nail Salon	School
Restaurant	Laundromat
Golf Course	Apartment
Church	Flower Shop

List B Pairs	
Arcade	Bakery
Grocery Store	Tennis Court
ATM	Synagogue
Auto Shop	Pool
Bus Stop	Garden
Library	Pub
Barber	Hospital
Courthouse	Factory
Travel Agent	Museum
Motor Home	Aquarium

Appendix B.

Multilevel model formulas and associated AIC values (Experiment 1)

Model	Hypothesized Models	AIC
E1J0	JOL~ 1 + (1 Participant)	4528.2
E1J1	JOL~ Map Complexity + (1 Participant)	4521.5
E1J2	JOL~ Map Complexity + (1 + Map Complexity Participant)	4500.3
E1M0	Distance Score~ 1 + (1 Participant)	13107.6
E1M1	Distance Score~ Complexity + (1 Participant)	13109.5
E1M2	Distance Score~ Complexity*JOL + (1 Participant)	13112.6
Model	Exploratory Models	AIC
E1J3	JOL~ Learning Trial*Map Complexity + (1 Participant)	4497.0
E1J4	JOL~ Learning Trial*Map Complexity + (1+Learning Trial Participant)	4436.2
E1RT0	TestRT~ 1 + (1 Participant)	15954.4
E1RT1	TestRT~ Map Complexity*JOL + (1 Participant)	15956.5
E1FR0	Free Recall Score~ 1 + (1 Participant) + (1 Landmark)	2635.9
E1FR1	Free Recall Score~ Category + (1 Participant) + (1 Landmark)	2623.1
E1J5	JOL~ SBSOD + (1 Participant)	4525.4
E1J6	JOL~ SBSOD*Map Complexity + (1 Participant)	4513.7

Appendix C.

Multilevel model formulas and associated AIC values (Experiment 2)

Model	Hypothesized Models	AIC
E2J0	$JOL \sim 1 + (1 Participant)$	4352.7
E2J1	$JOL \sim \text{Map Saliency} + (1 Participant)$	4259.4
E2M0	$\text{Distance Score} \sim 1 + (1 Participant)$	13089.9
E2M1	$\text{Distance Score} \sim \text{Saliency} + (1 Participant)$	13091.9
E2M2	$\text{Distance Score} \sim \text{Saliency} * JOL + (1 Participant)$	13089.6
Model	Exploratory Models	AIC
E2J2	$JOL \sim \text{Learning Trial} * \text{Map Saliency} + (1 Participant)$	4475.4
E2J3	$JOL \sim \text{Learning Trial} * \text{Map Saliency} + (1 + \text{Learning Trial} Participant)$	4431.1
E2RT0	$\text{TestRT} \sim 1 + (1 Participant)$	15988.5
E2RT1	$\text{TestRT} \sim \text{Map Complexity} * JOL + (1 Participant)$	15993.8
E2FR0	$\text{Free Recall Score} \sim 1 + (1 Landmark)$	2529.6
E2FR1	$\text{Free Recall Score} \sim \text{Category} + (1 Landmark)$	2497.5
E2M3	$\text{Distance Score} \sim \text{SBSOD} + (1 Participant)$	13143.9
E2M4	$\text{Distance Score} \sim \text{SBSOD} * \text{Map Saliency} + (1 Participant)$	13147.4

Appendix D.

Landmark name pairs in English and Estonian (Experiment 3). Landmark names changed from Experiments 1 and 2 to prevent the presence of cognates are italicized.

List A Pairs	
Movie Theatre	Zoo
Bank	<i>Concert Hall</i>
Pharmacy	Art Gallery
Car Wash	Post Office
Train Station	Realtor
Gym	<i>Track</i>
Nail Salon	School
Restaurant	Laundromat
Golf Course	Apartment
Church	Flower Shop

List A Pairs	
Kino	Loomaaed
Pank	<i>Kontserdisaal</i>
Apteek	Kunstigalerii
Autopesula	Postkontor
Rongijaam	Kinnisvaramaakler
Jõusaal	<i>Rada</i>
Küünesalong	Kool
Kohvik	Pesumaja
Golfirada	Korter
Kirik	Lillepood

List B Pairs	
<i>Amusement Park</i>	Bakery
Grocery Store	Tennis Court
ATM	Synagogue
<i>Acupuncturist</i>	Pool
Bus Stop	Garden
Library	<i>Tavern</i>
Barber	Hospital
Courthouse	Factory
Travel Agent	<i>Circus</i>
Motor Home	Aquarium

List B Pairs	
<i>Lõbustuspark</i>	Pagariäri
Toidupood	Tenniseväljak
Sularahaaautomaat	Sünagoog
<i>Nõelraviarst</i>	Bassein
Bussipeatus	Aed
Raamatukogu	<i>Kõrts</i>
Juuksur	Haigla
Kohtumaja	Tehas
Reisibüroo	<i>Tsirkus</i>
Haagissuvilla	Akvaarium

Appendix E.

Multilevel model formulas and associated AIC values (Experiment 3)

Model	Hypothesized Models	AIC
E3J0	$JOL \sim 1 + (1 Participant)$	7008.9
E3J1	$JOL \sim \text{Semantic Accessibility} + (1 Participant)$	7010.6
E3M0	$\text{Distance Score} \sim 1 + (1 Participant)$	21371.6
E3M1	$\text{Distance Score} \sim \text{Semantic Accessibility} + (1 Participant)$	21369.3
E3M2	$\text{Distance Score} \sim \text{Semantic Accessibility} * JOL + (1 Participant)$	21370.9
Model	Exploratory Models	AIC
E3JE0	$JOL \sim 1 + (1 Participant)$	9554.0
E3JE1	$JOL \sim \text{Semantic Accessibility}^3 + (1 Participant)$	9556.6
E3ME0	$\text{Distance Score} \sim 1 + (1 Participant)$	28881.7
E3ME1	$\text{Distance Score} \sim \text{Semantic Accessibility}^3 + (1 Participant)$	28881.0
E3ME2	$\text{Distance Score} \sim \text{Semantic Accessibility}^3 * JOL + (1 Participant)$	28882.6
E3JE2	$JOL \sim \text{Learning Trial} * \text{Semantic Accessibility}^3 + (1 Participant)$	9434.1
E3JE3	$JOL \sim \text{Learning Trial} * \text{Semantic Accessibility}^3 + (1 + \text{Learning Trial} Participant)$	9315.3
E3RT0	$\text{TestRT} \sim 1 + (1 Participant)$	41166.4
E3RT1	$\text{TestRT} \sim \text{Condition} * JOL + (1 Participant)$	41162.9
E3JE4	$JOL \sim \text{SBSOD} + (1 Participant)$	9548.0
E3JE5	$JOL \sim \text{SBSOD} * \text{Semantic Accessibility}^3 + (1 Participant)$	9552.1