Understanding Metacognition's Role During Map Learning: Evidence for Global Monitoring

A master's thesis submitted by

Lauren A. Mason

in partial fulfillment of the requirements for the degree of

Master of Science

in

Psychology

Tufts University

May 2023

Advisors:

Dr. Holly A. Taylor

Dr. Ayanna K. Thomas

Abstract

Maps are complex: They contain landmark locations, routes between landmarks, and even landmark names. Complex information requires cognitive effort to process, and people do not remember everything from complex visuospatial stimuli. Verbal research shows that metacognitively monitoring learning can benefit memory. Since the relationship between metacognitive monitoring and map learning is understudied, it remains unknown whether monitoring benefits memory for map-based materials. In Experiment 1, we explored the reactive role of monitoring judgments (JOLs) for landmark name memory. Findings suggest that JOL-reactivity occurs for map-based stimuli and learners could accurately predict their memory for maps, even when learning and test contexts were incongruent. With support that monitoring prompts could bolster learning and predictive accuracy for map components, we turned our attention to investigate the conditions necessary for effective monitoring prompts during spatial learning. We explored the relationship between learning goals and JOL prompts (Experiment 2), and JOL prompts and tests (Experiment 3). We expected better memory accuracy and metacognitive accuracy when goals aligned with JOL prompts, and when learners were tested on map elements that they monitored. Findings suggest that congruency between learning goals and JOL prompts contributes to metacognitive accuracy, but it depends on the order learners attend to differing map components. When learning goals and JOLs were congruent, learners could accurately predict their memory, irrespective of the map components on which they were tested. Taken together, our results suggest that learners can accurately predict their memory accuracy, and cues such as map complexity and time influence rating magnitude. Our findings provide novel insights on the nature of metacognitive monitoring judgments during map-based learning.

Table of Contents

List of Figures & Tables

List of Figures & Tables Cont.

Introduction

Imagine being in a new city for the first time. You prepare to travel from the conference hotel to a restaurant. Looking at a map on your phone, you find your present location (the hotel), and your destination (the restaurant). How likely are you to remember the *location* of the landmark you are traveling to if asked again in the future? What about the *route* between where you are now and where you are going? Or even the *name* of the landmark destination?

As garnered from this scenario, we gain a lot of environment information from looking at maps, including landmark locations, routes between landmarks, and even landmark names that allow us to reference environment components (Westerbeek & Maes, 2013). In other words, map displays are complex, information-rich, and require cognitive effort to process (Bunch & Lloyd, 2006). We typically use a map when we are unfamiliar with an environment (Aretz, 1988; Coluccia, 2008) and processing unfamiliar information also increases cognitive load (Tuovinen & Sweller, 1991).

Cognitive load can be understood as the amount of effort required to acquire and process information in working memory (Bunch & Lloyd, 2006; Knight & Tlauka, 2017). The burden of extraneous cognitive load for learning with complex materials is well-established (Garden et al., 2002; Kirschner et al., 2006; Scott & Schwartz, 2007; Valcke, 2002; van Merriënboer et al., 2002). As an example, research finds when participants engage in a concurrent task such as spatial tapping while learning a map, absolute positions and route knowledge is hindered, potentially due to working memory interference (Coluccia, 2008). Similarly, when participants under high cognitive load are required to actively learn maps, encoding is detrimentally affected (Knight & Tlauka, 2017). These discoveries suggest active spatial learning is cognitively demanding and impeded when mental resources are consumed. Furthermore, people do not remember everything from

complex visuospatial stimuli such as maps. Yet, reduction of concurrent cognitive demands while learning from maps is not always ecologically plausible. Because successful memory for map components is critical to prevent people from becoming spatially disoriented or lost, the present study considers strategies to support learning. Specifically, we consider how metacognition and learning goal instantiation can scaffold map learning. We begin by discussing the role of metacognition as a learning scaffold.

Map Learning Scaffolds: Metacognition

Various factors can impact what information someone remembers from learned materials. One way to potentially impact what people remember from learned materials is through metacognitive monitoring (Janes et al., 2018). Metacognition is a complex process that consists of monitoring one's own knowledge and controlling learning through certain strategies (Flavell, 1979; T. Nelson & Narens, 1994; T. O. Nelson, 1990; Osman & Hannafin, 1992; Schraw, 2001). The critical role of metacognition for successful learning with complex stimuli is well-cited throughout the multimedia literature (Cuevas et al., 2002; Cuevas et al., 2004; Schwartz et al., 2004; Scott & Schwartz, 2007). For instance, some researchers find that participants higher in metacognitive skills more accurately judge their learning comprehension (Cuevas et al., 2004) and show better knowledge acquisition and performance following complex task training (Ford et al., 1998).

However, these studies typically assess metacognition generally, rather than focusing on one metacognitive component, such as monitoring. These studies also tend to measure metacognitive skills separate from the learning task, via questionnaire measures, rather than engendering metacognitive strategies as part of the experimental design. These approaches provide less insight on the specific metacognitive processes involved during complex learning tasks. Therefore, the

present study expands on findings from the multimedia literature that suggest metacognition is essential for successful learning and memory of complex stimuli. In the current experiments, metacognitive monitoring is engendered and assessed on a trial-by-trial basis during learning.

Findings show that the act of making JOLs can positively impact memory for learned materials. One possible explanation for this result is that JOLs impact people's learning goals, causing them to modify learning strategies and maximize memory for easier, related materials (Janes et al., 2018). Other possibilities are that JOLs enhance attention (Mitchum et al., 2016), promote deeper encoding (Double et al., 2018), and thereby influence memory for studied items (Dougherty et al., 2005). This effect is referred to as JOL-reactivity. However, to this point, little is known about the reactive role of metacognitive monitoring judgments during complex spatial learning. In the present study, we assess whether JOL-reactivity is observed during map-based learning.

There is reason to believe that metacognitive monitoring might positively impact retention of complex spatial stimuli. Findings from multimedia literature suggest that metacognitive evaluation and monitoring are essential for effective interaction with complex navigational maps, but metacognitive manipulations must be relevant to instructional goals for them to successfully support learning (Cuevas et al., 2002; Mayer, 2001; Scott & Schwartz, 2007). However, to our knowledge this work does not explicitly engender metacognitive monitoring. We begin by testing whether JOL-reactivity effects are observed for map-based stimuli in Experiment 1. In real-world settings, it is common to see landmark names on a map but be tasked with remembering them without the context of the map. Therefore, in Experiment 1, we also test the role of study-test congruency for landmark name memory. Finally, we assess whether the act of making JOLs aids memory for incongruent study-test scenarios.

With knowledge on the role of monitoring prompts during map-based learning, we turn our attention to investigate the conditions necessary for effective monitoring prompts during spatial learning. Namely, we explore the relationship between learning goals and JOL prompts in Experiment 2, and the relationship between JOL prompts and tests in Experiment 3. We address the motivation and predictions for this portion of the study in more detail following Experiment 1.

Experiment 1

In Experiment 1, we investigated JOL-reactivity and congruency between study-test contexts. Participants studied landmark name pairs on maps or in word-pair format, without the context of the map. Half the participants made JOLs about their likelihood to remember the second landmark name of the pair if prompted with the first landmark name on a future memory test. At test, participants saw the first landmark of a pair and had to complete the pair. All landmark pairs were tested in word-pair format.

The aims of this experiment were three-fold. First, we were interested in whether the act of making JOLs could impact landmark name pair retention (JOL-reactivity). Second, we were interested in whether better memory and metacognitive accuracy were observed for congruent study-test contexts, such as when participants studied landmark name pairs without the context of a map and were tested without maps, compared to when pairs were studied with a map and tested without a map. Finally, we were particularly interested in whether the act of making JOLs supports memory when there was an incongruency between study and test contexts. We refer to this latter point as the JOL-Bridge Hypothesis.

JOL-Reactivity

A primary goal of Experiment 1 was to determine whether the act of making JOLs impacts memory retention for landmark name pairs. According to verbal learning research, monitoring

one's likelihood to remember something in the future may support memory for studied materials (Janes et al., 2018). This is commonly referred to as JOL-reactivity. Much of the work on the JOL-reactivity effect is conducted using word pairs, however. Therefore, the current experiment extends these findings to more complex stimuli, namely by displaying landmark name pairs on maps. We predicted that the act of making JOLs would support memory for landmark name pairs as compared to not making such judgments.

The current experiment measures JOL-reactivity, specifically. Half of the participants made a JOL after each learning trial and half generated a random number (RNG). In metacognition research, JOLs are often implemented without the RNG group comparison. By including a no-JOL group, it is possible to compare recall across group type and determine whether there is a difference in performance among participants asked to judge their learning (Janes et al., 2018).

Study-Test Congruency

Another goal of Experiment 1 was to investigate the role of congruency for memory and metacognitive accuracy. Specifically, we assessed whether studying landmark name pairs *without* maps and testing *without* maps results in better memory and metacognitive accuracy than studying pairs on maps and testing *without* the map context. Prior work finds that a match between judgment and test contexts may be important for monitoring accuracy (Dunlosky et al., 2005), while a match between study and test contexts may contribute to memory accuracy at test (Morris et al., 1977). In the present study, we extend these classic findings to more complex stimuli.

Based on Kulhavy's conjoint retention theory (Kulhavy et al., 1985) and Paivio's dual coding theory (Paivio, 1986), geographic maps provide visual information that can be used to acquire and store information from the map. Moreover, when a map and text are learned together, it is possible to make cross-code connections between the two. These connections allow for

information from one code to function as a retrieval cue for information from the other code. This theory assumes that the spatial information from the map and the related text information are held simultaneously in working memory, which result in increased recall of both. In other words, it is possible that learning landmark names on maps may offer additional cues to recall associate names at test. However, due to the complex nature of map displays, it is also possible that studying landmark names on maps might produce disadvantageous working memory demands and study tactics which do not aid recall when landmark names are presented without maps at test.

Therefore, we predicted better memory and metacognitive accuracy for landmark name pairs that were studied *without* the map and tested *without* the map (congruent study-test context), as compared to pairs that were studied *with* the map and tested *without* the map (incongruent study-test context). Furthermore, we predicted congruency between study and test contexts would aid participants' ability to accurately predict their future memory performance. We hypothesized greater metacognitive accuracy for the JOL group on items studied (and subsequent metacognitive judgments provided) *without* maps and tested *without* maps (congruent study-test context), as compared to those studied (and subsequent metacognitive judgments provided) *with* maps and tested *without* maps (incongruent study-test context).

JOL-Bridge Hypothesis

The final interest of Experiment 1 was to explore whether the act of making JOLs benefits memory when there is an incongruency between study and test contexts. In other words, we considered whether making JOLs serves as a bridge to support memory in complex learning scenarios. Based on the verbal learning literature, the act of making JOLs can aid retention potentially by influencing the way learners attend to (Mitchum et al., 2016) or encode information (Double et al., 2018). Returning to findings from the multimedia literature, metacognition may be

critical to successfully learn from complex stimuli (Cuevas et al., 2002, 2004; Schwartz et al., 2004; Scott & Schwartz, 2007). Therefore, there is reason to believe that making JOLs about one's likelihood to remember specific information from complex stimuli may support memory for that information at test.

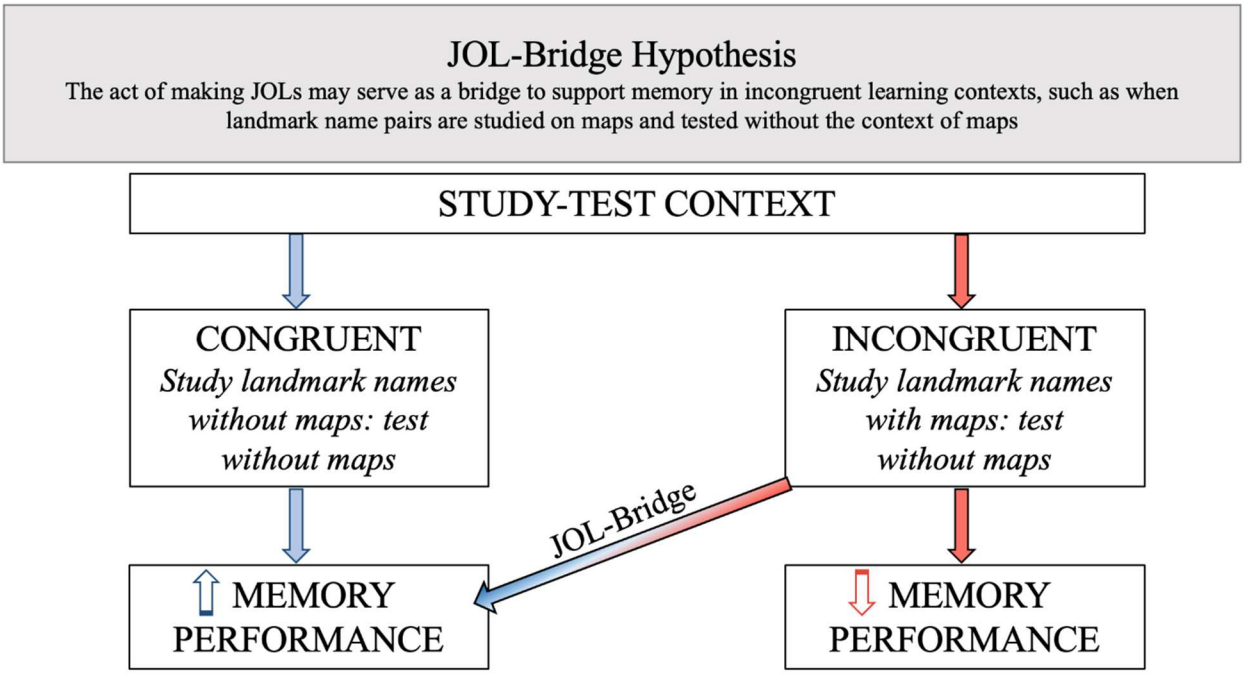
In the present experiment, participants are told that their memory for landmark names will be tested. Half the participants rated their likelihood to remember the second landmark if prompted with the first landmark of the pair on a future memory test. These specific monitoring judgments may scaffold learning by directing participants to attend to goal-relevant information, namely the associated landmark name. This scaffold may be particularly important in more complex learning scenarios, such as when the landmark names appear on a map. In addition to displaying the landmark names spatially, the map stimuli in the current experiment also included routes. In other words, there are a variety of features that participants can attend to when studying landmark pairs on maps. Yet, attending to these additional features may not aid memory for the associated landmark name at test.

With these components in mind, the act of making JOLs about one's likelihood to remember the second landmark if prompted with the first on a future memory test may direct attention, scaffold learning, and support memory in incongruent learning contexts, such as when landmark name pairs are studied on maps and tested without the map context. We predicted that making JOLs, compared to picking a random number (also referred to as random number generation or RNG), would result in better memory on incongruent trials. That is, we predicted metacognitive monitoring would serve as a bridge to support memory performance when faced with study-test incongruency. We hypothesized greater memory for landmark name pairs that were studied with maps and tested without the context of maps (incongruent study-test context) among

the JOL group as compared to the RNG group. A visual presentation of the JOL-bridge hypothesis is provided below (Figure 1).

Figure 1

Visual representation of JOL-bridge hypothesis. The act of making JOLs should support learning in complex learning scenarios, resulting in better performance when study-test contexts are incongruent.



Metacognitive Accuracy-Congruency Relationship

Lastly, we anticipated that congruency between study and test contexts would aid participants' ability to accurately predict their future memory performance. Based on the transfer appropriate monitoring (TAM) hypothesis, judgment accuracy depends on the match between judgment and test contexts (Dunlosky et al., 2005). In the present experiment, participants made JOLs during the study phase. Therefore, we predicted greater metacognitive accuracy for the JOL

group on items studied without maps and tested without maps (congruent study-test context), as compared to those studied with maps and tested without maps (incongruent study-test context).

Method

Participants

Participants were recruited from Tufts University through SONA and compensated with class credits. They used a laptop/desktop using Chrome or Firefox and had screen resolution greater than 1000x600 to complete the experiment.

This experiment implemented novel spatial stimuli in a metacognitive paradigm loosely modeled after those found in the word-pair literature. Therefore, due to the exploratory nature of this work, we were unable to conduct a representative power analysis. The desired sample size for the present experiment was determined by holistic consideration of previous work from the 1) verbal 2) spatial and 3) multimedia learning literatures.

Within the transfer appropriate processing literature, Blaxton (1989) sought to separate the effects of the memory system tapped by a test from effects produced by the type of processing required by that test. In 2 of 3 experiments, the author implemented a 2 x 2 mixed design, and stimuli included between 126-128 words. There were 60 participants in Experiment 1 and 68 in Experiment 2.

Turning now to the work on spatial learning, a study by Bonura (2011) investigated the relationship between location and identity, memory, aging, and metacognitive monitoring using a visuospatial-working memory paradigm. The author implemented a 2 x 3 x 4 mixed factorial design and had participants view objects on grids. There were 20 objects and 400 grids, with 50 participants in each of the four experiments.

Moreover, prior work (Thomas et al., 2012) implemented a 2 x 2 mixed factorial design to examine map learning in older adults using 8 maps depicting 4 environments with 12 landmarks for each environment included 47 participants. A second experiment from the same study implemented a 2 x 3 mixed factorial design and included 99 participants. Further work (Dai, 2016) assessed how spatial organization, semantic association, strategic processing, and aging affect encoding and binding between location and object processing in visuospatial working memory. The author implemented a 2 x 2 x 3 x 2 design, and stimuli included 144 grids containing 5 objects over four experiments. The number of participants ranged from 51-72 across experiments.

In consideration of multimedia research, (Schwartz et al., 2004) assessed how metacognitive skill and prior experience impact memory retention of a hypermedia learning environment. Participants (n= 28) were students ranging in age from nine to seventeen years. To measure locational memory, learners found 10 Web pages in a rainforest-themed site based on verbal cues they were given about salient information contained on the page. They were given 90 seconds to find each page. To measure recall, learners recorded everything they remembered from the module they searched.

Furthermore, (Cuevas et al., 2002) examined how instructional strategies can support knowledge acquisition and metacomprehension of complex, computer-based training environments. Specifically, they were interested in the role of diagrams to facilitate learning outcomes. The authors implemented a 2 x 2 mixed between/within design. Participants (n= 78) learned about principles of flight from a tutorial that was divided into three modules.

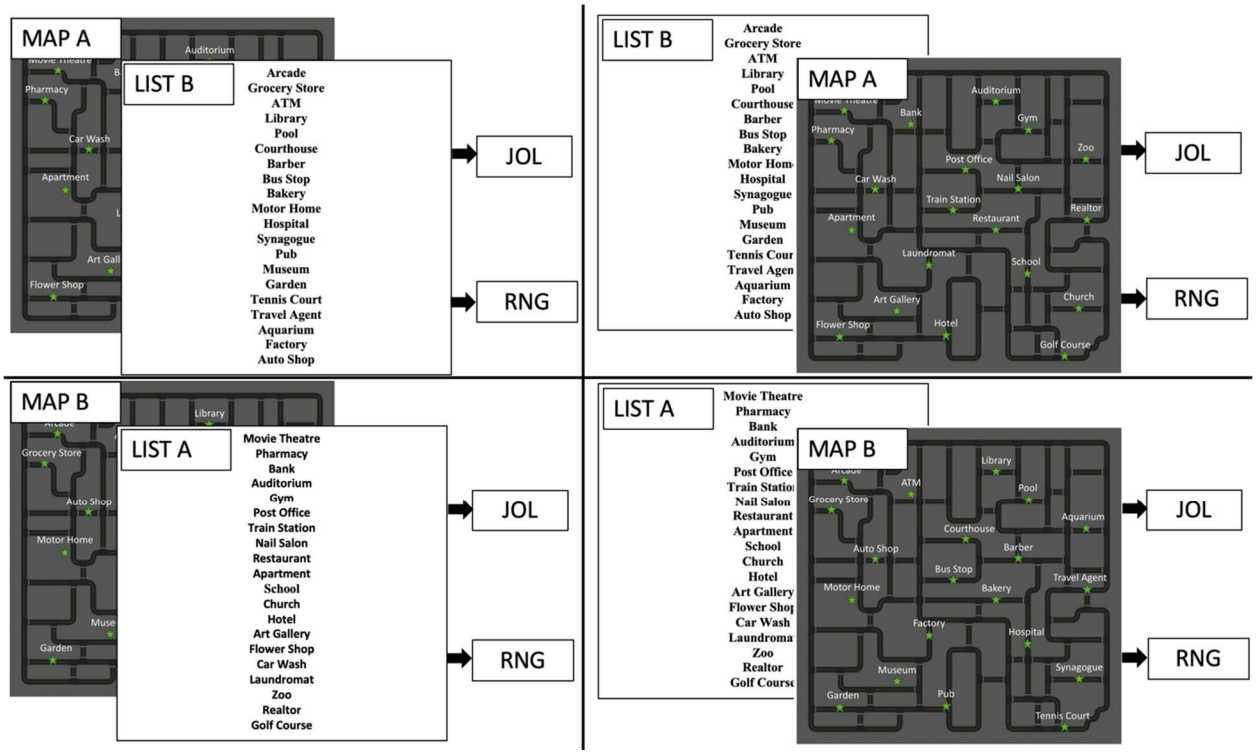
From the experiments described, the total number of participants ranges from 28 at the low end to 99 at the high end. For the present study, we analyzed data from 66 (Experiment 1), 102 (Experiment 2) and 94 (Experiment 3) participants.

Design

A 2 (Study Context: Map, Word-Pair) x 2 (Judgment Group: JOL, RNG) mixed design was employed. Study context was a within-participant variable; participants studied half of the landmark pairs on maps and half of the landmark pairs in word-pair format. Judgment condition was a between-participant variable; half the participants were randomly assigned to the JOL condition and half to the RNG condition. See Figure 2 for a visual schematic of this experiment. All participants were tested on all landmark pairs in word-pair format.

Figure 2

Schematic of conditions in Experiment 1: Content from learning and judgment phases.



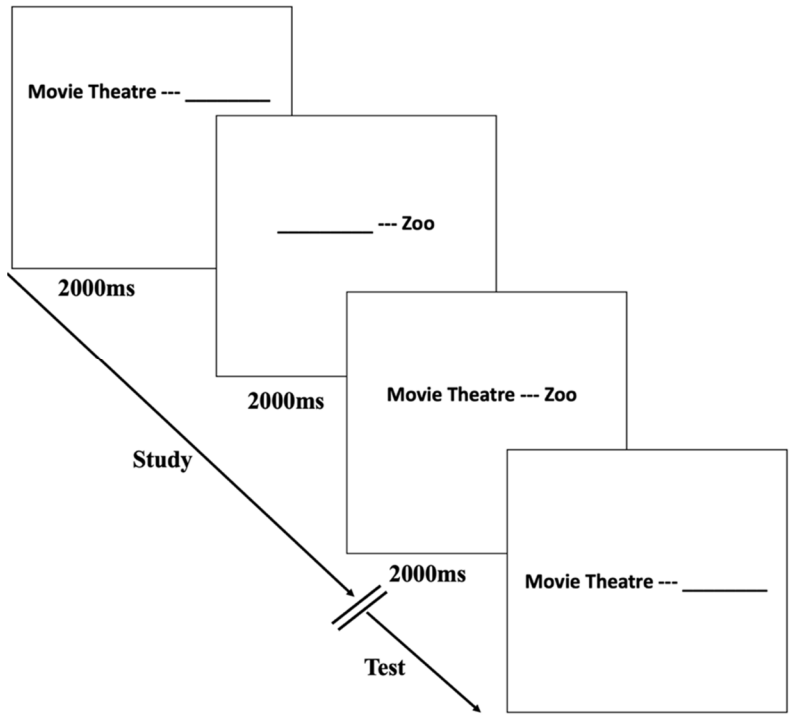
Materials

Landmark name pairs. There were two lists of 10 landmark pairs (20 total landmarks, Appendix A). Landmark names were inspired by those typically found in cities. The landmark names were general and based on function, such as "Movie Theatre" and "Grocery Store".

Participants studied a block of 10 landmark pairs on maps and a block of 10 landmark pairs in word-pair format. Order of landmark pair presentation was randomized within each block, and the blocks were counterbalanced. Half of the participants studied landmark pairs from list (A) on a map and those from List B as word-pairs, while the other half studied landmark pairs from list (A) in word-pair format and those from List B on a map. After each learning trial, half of the participants made a JOL, while the other half picked a number between 1 and 4. At test, participants studied and were tested on all landmark pairs in word-pair format. Landmark pairs were randomized at test. See Figure 3 (word pairs) and 4 (maps) for sample stimuli.

Figure 3

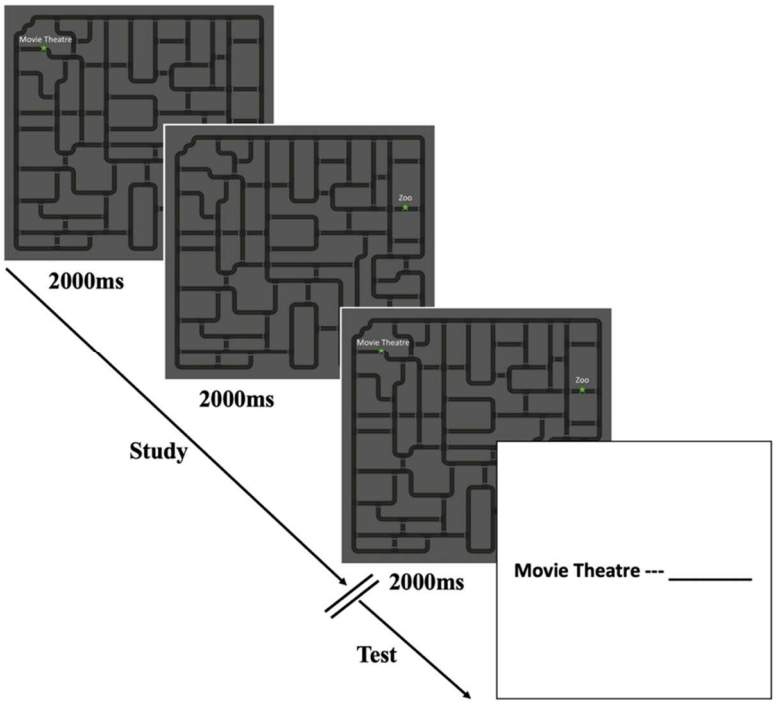
Sample learning stimuli with landmark pairs in word pair format. During test, all cues were displayed in word format.



.

Figure 4

Sample learning stimuli with landmark pairs on maps. During test, all cues were displayed in word format.



Note. Font size presented on maps is not representative of actual font size during the experiment.

Maps. The first author developed the map stimuli using a VR environment map as a base. This map was then modified to incorporate additional map features. The individual map stimuli containing landmark names were then created using Microsoft PowerPoint. The maps were approximately 34 x 33 cm in size. See Appendix B for complete maps for both landmark lists.

Questionnaires. Participants completed 4 questionnaires: a demographics questionnaire, an encoding-based strategies questionnaire, and two sense-of-direction questionnaires. They also answered three questions at the end of the study that ask whether they had difficulty loading the displays. Demographics questions asked for participant gender, ethnicity, highest degree

completed, and age. The encoding-based strategies questionnaire assessed how much participants paid attention to and subjectively remembered map components, including landmark locations, routes between landmarks, and landmark names (Appendix C). Participants responded using a Likert scale and had the option describe their responses in an answer box. Finally, participants rated the overall difficulty of remembering the information needed at test on a scale from 1 (not difficult) to 7 (very difficult). The sense-of-direction questionnaires included the Santa Barbara Sense of Direction scale (Hegarty, 2002) and a single question about their sense of direction ("when traveling somewhere in a new city, please rate your overall sense of direction"). Participants responded using a sense of direction slider ranging from 0 or "terrible" to 10 or "excellent".

Attention checks. Participants saw two attention checks within 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).

Retention interval. Participants completed a six-minute retention interval after the study phase and before the test. The interval consisted of a hidden picture image shown for 2 minutes. Participants were told that they would complete a visual search task. They were instructed that they would be shown a photo with hidden items in them. The list of hidden items was provided at the bottom of each of the photos. They were given 2 minutes to study the photos and find the

hidden items. Then, they reported how many of the items they found. The hidden picture printable for adults was pulled from the internet (https://www.printablee.com).

Procedures

Santa Barbara Sense of Direction Scale. At the start of the semester, all participants completed the Santa Barbara Sense of Direction scale (Hegarty, 2002) as part of the SONA systems prescreening process.

Experiment. The experiment was carried out over the internet using Qualtrics Survey Software. Informed consent was obtained from participants at the start of the experiment. The experiment was split into three phases: a practice session, learning with ratings, and a final test.

Practice. The task began with two practice trials. In one trial, they were prompted to study landmark name pairs on a map. In another trial, they studied landmark names in word-pair format. Specifically, they studied the landmark name pairs, "leaf --- rock" and "tree --- flower." Half of the participants saw "leaf --- rock" on a map, while the other half saw it in word-pair format, similarly to the way it is shown here. The same is true for "tree --- flower."

JOLs. After studying each pair, the pairs were removed from the screen and half of the participants made a JOL. They were prompted with the question, "How likely are you to remember the second landmark when presented with the first if tested at a later time?" with the response options as follows: "1 not likely to remember/ 2 somewhat likely to remember/ 3 likely to remember/ 4 highly likely to remember". The other half of participants were prompted to "Pick a number between one and four," with the response options as: "1/ 2/ 3/ 4". This RNG group was designed to parallel baseline processes inherent to the JOL group, without assessing future retrievability.

Tests. Next, participants took a practice landmark cued-recall test. They were instructed to, "Type the name of the associated landmark that completes the pair. If you cannot remember the name of the second landmark, try to guess. If you cannot guess, type an 'X' in the answer box. Press the arrow to proceed." Then, they were cued with one of the landmark names, such as "leaf," and they were expected to type the target landmark name that completed the pair (in this case, "flower"). After they completed the practice phase, the learning phase began.

During the learning phase, participants studied 20 landmark name pairs. Prior to beginning study, they were informed that once they had finished studying all of the pairs, their memory for the landmark pairs would be tested. They were told, "during testing, you will be presented with a landmark and you will have to try to retrieve the landmark with which it was paired. Please type the target landmark when prompted with the cue." Then, the landmark pair presentation began. Participants briefly studied twenty landmark names at the start of the learning phase: half of the participants were presented with a list of twenty landmark names, while half of them saw twenty landmark names on a map. They are allotted 30 seconds to study all of the landmark names. Next, the individual pair study began. Participants were first presented with one landmark name, such as "Gym", then the next landmark flashed, such as "Hotel", and finally both were shown together. These either appeared in word pair format (i.e., "Gym --- Hotel") or on a map. Each landmark name presentation persisted for 2000 ms before the next landmark name was presented. Each pair of landmark names was then presented for 2000 ms. Then, the participant was prompted with a judgment (JOL or RNG) and the next trial began.

The first attention check was placed at the end of study block 1 and prompted participants to enter a color name in the box provided.

After they studied ten landmark name pairs (20 total names) in the first study block (either in word pair format, or on the map), the next study block began. In the next block, they studied ten more landmark name pairs in the opposite format to the one shown previously. So, if they first studied landmark name pairs on a map, they moved on to study ten new landmark name pairs in word-pair format. As was done during practice, participants were prompted to either make a JOL or RNG following each successive learning trial. After they finished studying the next set of 10 landmark name pairs, the test phase began.

Retention interval. Participants completed the six-minute retention interval after the study phase. Then, the test phase began.

During the cue-recall test, the first landmark of a pair served as the cue and the second landmark was the to-be-retrieved target. Participants were provided with the instructions, "Type the name of the associated landmark that completes the pair. If you cannot remember the name of the second landmark, try to guess. If you cannot guess, type an 'X' in the answer box. Press the arrow to proceed". All twenty studied landmark name pairs were tested in word-pair format without the context of the map. Participants were cued with a landmark name, such as "Gym", and they were expected to type in the target landmark name that completed the pair (in this case, "Hotel"). The test was self-paced. Participants proceeded to the next test trial after providing their response.

The second attention check was placed after the test phase and had participants select the extent to which the statement "I was born before 1920" applied to them.

Questionnaires. After the test phase of the landmark cued-recall task, they completed the demographics, encoding-based strategies, single-item sense of direction, Santa Barbara Sense of Direction scale (Hegarty, 2002), and technical difficulty questionnaires.

Experiment 1 Results

In Experiment 1, we collected data from 71 participants. We excluded 5 participants due to technical difficulties. We analyzed data from 66 participants ages 18-23 years. Most participants identified as white (47%) or Asian (24%) females (50%). All participants were college students.

Memory and Metacognitive Accuracy Results

We estimated linear mixed models (LMMs) to assess our primary hypotheses (see Table 1). In all models, we included participant as a random effect, because some participants may be more skilled at map-based learning. One common way to assess model fit is to begin by running an analysis with only the intercept terms (i.e., the null model) and comparing the AIC of the null model to the hypothesized model (Starkweather, 2010). We utilize this approach in the analyses (Table 1). First, we tested whether better memory accuracy was observed for word pairs compared to maps, for the JOL condition compared to the RNG condition. This model was a better fit for the data as compared to the null model, (i.e., lower AIC).

Table 1

Model formulas and associated AIC values predicting memory accuracy

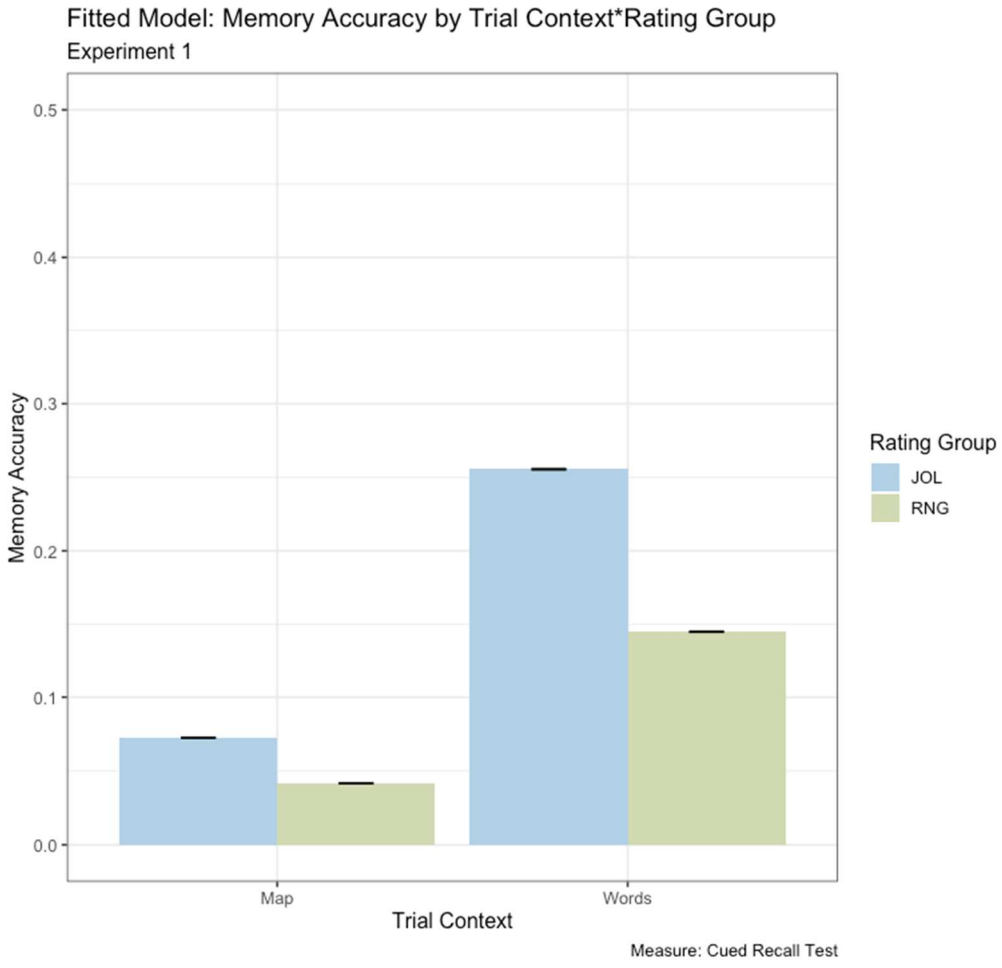
Model	Model formula	AIC
E1MN	Accuracy~ 1 + (1|PID)	1117.5
E1M1	Accuracy~ Trial Context*Rating Group + (1|PID)	1048.2
E1M2	Accuracy~ Trial Context*Rating Group*Rating + (1|PID)	1007.1

We observed main effects of trial context, rating group, and a trial context*rating group interaction ($p < .001$). Both groups more accurately remembered landmark name pairs that

appeared in word pair format as compared to those that appeared on maps. The JOL group more accurately remembered landmark names than the RNG group (Fig 5).

Figure 5

E1M1. Mean memory accuracy as a function of trial context and rating group. Error bars show SE.

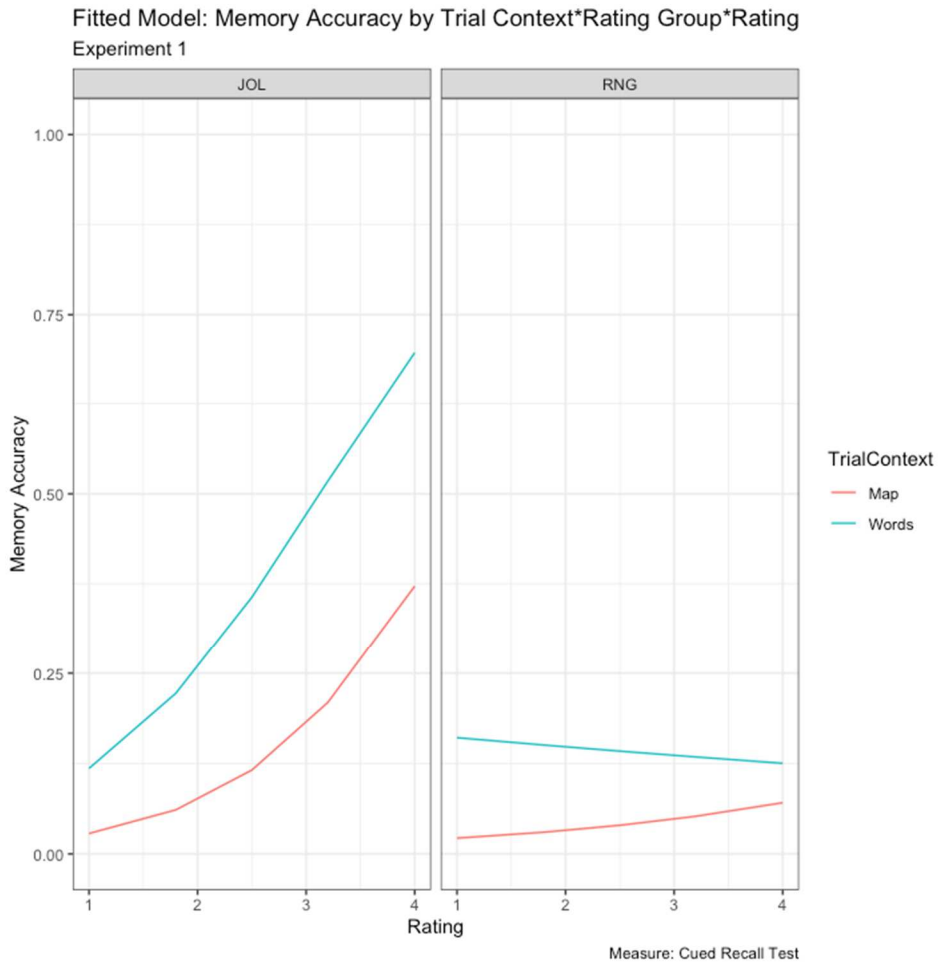


Next, we tested whether JOL ratings were a significant predictor of memory accuracy and whether metacognitive accuracy differed across trial contexts. We entered rating into the model (Table 1, E1M2) to see whether rating significantly predicted memory accuracy and whether there

was a rating group*trial context interaction. We saw a significant improvement in the overall model fit by adding rating to the model. We observed a significant interaction between rating group and rating ($p < .001$). (Fig 6). Rating was a significant predictor of memory accuracy only among the JOL group ($p = .02$). This model was an even better fit for the data as compared to E1M1.

Figure 6

E1M2. Mean memory accuracy as a function of rating across rating groups and trial contexts.



Exploratory JOL Magnitude Results

Next, we estimated exploratory models to test whether trial context and rating group predicted JOL magnitude (Table 2, E1M3). While rating group significantly predicted rating ($p < .001$), trial context did not (Fig 7). The JOL group rated themselves significantly less likely to remember the associated landmark name in the future, as compared to the RNG group. This likely reflects that JOLs meaningfully relate to learning, while RNGs do not. That said, there were no differences in rating magnitudes between learning contexts for the JOL group, suggesting that the context of the learned materials did not impact their judgment magnitude.

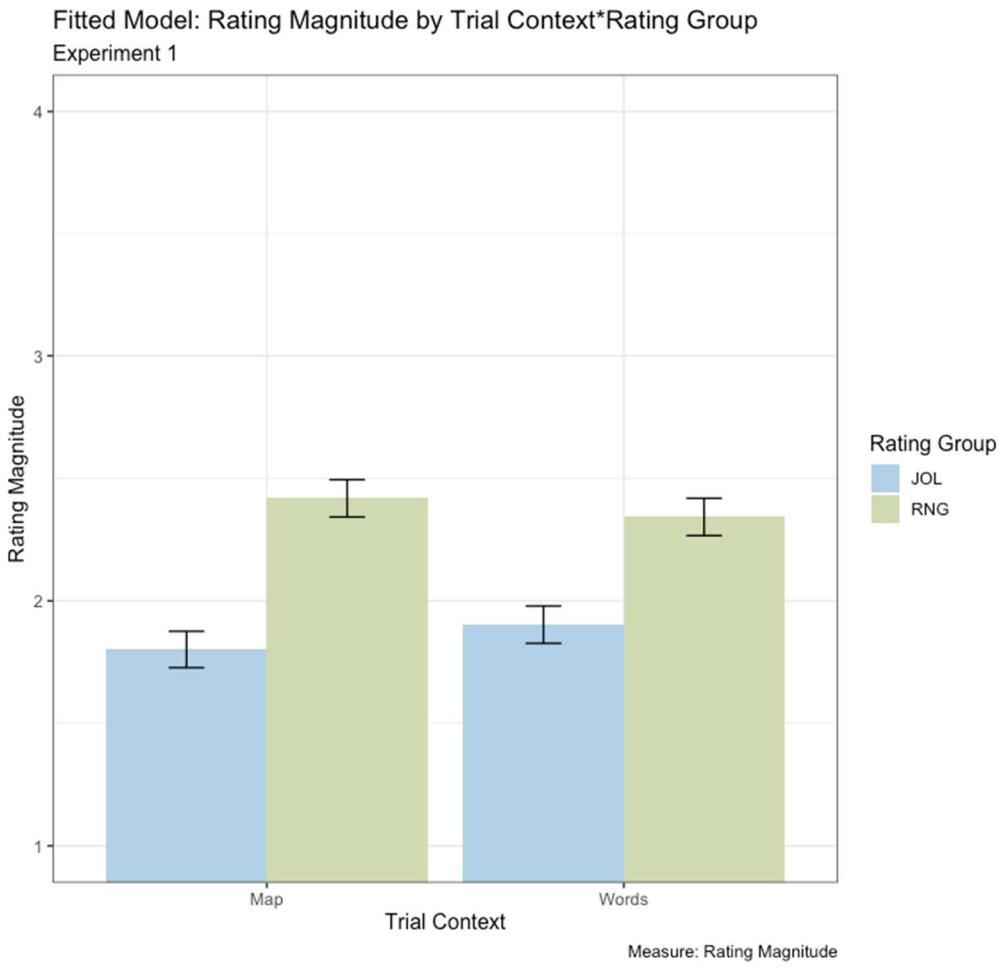
Table 2

Exploratory model formulas and associated AIC values predicting JOL magnitude

Model	Model formula	AIC
E1MNJ	Rating~ 1 + (1|PID)	3589.012
E1M3	Rating~ Trial Context*Rating Group + (1|PID)	3581.602

Figure 7

E1M3. Mean rating magnitude as a function of trial context and rating group. Error bars show SE.



Encoding Based Strategies

No between-group differences were observed on the encoding-based strategies questionnaire ratings, including the single item sense of direction (SOD) measure, so none are reported here.

Experiment 1 Discussion

In Experiment 1, we tested whether JOL-reactivity was found with map stimuli. Additionally, we assessed whether classic transfer appropriate processing and monitoring effects were observed with complex map materials, and whether the act of making JOLs would aid memory performance in incongruent study-test contexts, which we referred to as the JOL-bridge hypothesis.

Results reveal that JOL-reactivity is observed with complex map materials, such that the JOL group was significantly more accurate at remembering landmark names in both map and word-pair contexts as compared to the RNG condition. This provides support for the JOL-bridge hypothesis. That said, both groups had more accurate memory for landmark names that appeared in word-pair contexts during learning and at test, meaning a transfer appropriate processing (TAP) finding emerged.

JOL ratings were a significant predictor of memory accuracy, generally. No interaction with trial context emerged. These results are contrary to classic transfer appropriate monitoring (TAM) findings, which would predict more accurate JOLs for name pairs that were studied in word-pair format and tested in word-pair format, as compared to those that were studied on maps. Furthermore, rating magnitude did not differ between learning contexts among the JOL group, suggesting that visual complexity of learned materials did not contribute to the judgment formation. It is possible that the act of monitoring learning for complex map materials yields benefits that transfer to less complex, related materials at test.

In sum, findings from Experiment 1 suggest that there is utility in monitoring learning for complex map materials, and that people are relatively accurate at doing so. In Experiments 2 and 3, we dove deeper into the relationship between monitoring judgments and learning goals

(Experiment 2) and monitoring judgments and tests (Experiment 3), to uncover the factors that contribute to accurate memory and metacognitive monitoring during map learning.

Experiments 2 & 3

Map learning is complex as it requires learners to encode information about relative landmark locations, routes between landmarks, and even landmark names. In real-world settings, learners may attend to one map component based on a particular learning goal, but later be tasked with remembering a different map component. For instance, when someone looks at a map, their goal may be to gain information about the *relative location* of their destination but later be tasked with remembering the *route* between two landmarks. The opposite could also be true. Learners can seek out maps to gain information about the *route* between two landmarks, but later be tasked with remembering their *relative landmark locations*. This suggests that incongruency between learning goals and tests may play a role in the way that people learn and remember information from maps. We address this point in the following two experiments.

In Experiment 1, we saw that prompting learners to engage in metacognitive monitoring during map-based learning benefitted memory for learned components. There also is reason to believe that metacognitive monitoring plays a natural role in map learning. For instance, as learners view a static map, they seek out information that aligns with their learning goals and assess whether they are likely to remember that information before navigating. The act of predicting one's likelihood to remember information in the future is a metacognitive monitoring judgment known as a judgment of learning, or JOL. If a learner predicts that they will remember a map component, such as a route between their current location and their destination, but later cannot remember it, they risk becoming disoriented or lost. Therefore, predictive accuracy is important in real-world scenarios.

In Experiment 1, we found that the act of monitoring learning benefitted memory outcomes and learners were relatively accurate at predicting their likelihood to remember landmark names from maps, even when they were tested without the context of the map. However, little is known about what conditions are necessary for effective metacognitive monitoring prompts during complex spatial learning. First, it remains unknown how the relationship between learning goals and metacognitive monitoring judgments guides learning processes and subsequent memory for map materials. Prior work reveals that learning goals influence knowledge gained about an environment by increasing attention to goal-relevant information (Taylor et al., 1999). Therefore, goals can be leveraged to scaffold learning for specific information on maps. Experiments 2 and 3 further explore the relationship between learning goals and monitoring judgments.

Understanding the Role of Learning Goals

One way to support complex spatial learning is by guiding learners to relevant information through goal instantiation. Goals guide attention during learning, help learners select relevant stimuli, and instantiate frameworks to interpret new information (Britton et al., 1979; LaBerge, 1995; Pichert & Anderson, 1977; Taylor et al., 1999). Based on multimedia literature, it is especially important to explicitly guide people when learning novel information because they lack pre-existing frameworks into which new information can be integrated (Kintsch, 1998; Kirschner et al., 2006; Snow & Lohman, 1984). In other words, learning goals serve as necessary scaffolds to support learning for complex, novel information.

Guiding learners to goal-relevant information also addresses the cognitive load burden. Cognitive load theory suggests free exploration of complex materials is detrimental to learning due to excessive cognitive load (Baddeley, 1986; Mayer, 2001; Paas et al., 2003, 2004; Scott & Schwartz, 2007). Specifically, having to search for relevant information places working memory

demands. Therefore, the working memory system cannot contribute to knowledge accumulation in long-term memory because it is being used to search for relevant information rather than being used to learn (Kirschner et al., 2006). Moreover, literature on cognitive load with map learning suggests that ambiguous task instructions that require searching for relevant information result in extraneous cognitive load and impair performance (Bunch & Lloyd, 2006). As offered by these authors, one solution to this problem is directing the learner's attention to critical map information. With these considerations in mind, we believe assigning people explicit map learning goals is a strategic way to guide attention to goal-relevant information, support the cognitive system despite demands, and promote integration of materials into memory. Therefore, the present study assigns people explicit map learning goals.

Findings from Taylor et al. (1999) further emphasize the notion that learning goals increase memory for specific aspects of the environment. Participants were assigned to learn the layout of an environment (survey goal) or to learn the fastest way to a location (route goal). Spatial memory was assessed through survey-based and route-based tasks. Results show that people who learn with a survey-focused goal are more accurate at survey-based tasks, while those with a route-focused goal are more accurate at route-based tasks. Moreover, these findings reveal that memory for certain aspects of a spatial environment can benefit from learning goals that are specific in nature. The present study expands on this finding by assigning people explicit learning goals for particular map components. Participants were either instructed to learn routes between landmarks or to learn relative landmark locations. In addition to being provided with specific learning goals, learners were also prompted to metacognitively monitor their likelihood to remember specific map components, otherwise referred to as JOL prompts.

Understanding the Role of JOL Prompts

To date, most studies measuring the role of metacognitive monitoring during learning largely utilize word list or word-pair stimuli. Word pairs are less complex than maps because they lack a spatial component. Therefore, pre-existing metacognitive approaches might require reconsideration in this spatial learning context. For instance, JOL prompts for word-pair stimuli are general in nature and require participants to rate their likelihood to remember a word if tested in the future. This is a much simpler judgment task compared to the type of monitoring learners are required to engage in during real-world map learning.

Some metacomprehension researchers acknowledge the limitation of general JOL prompts for more complex materials. In fact, Dunlosky and colleagues (2005) suggest that the standard procedure for obtaining judgments of learning for lengthy text may be to blame for poor metacognitive accuracy. As an example, participants are typically asked to read a 200–400 word text and then are asked, "How well will you be able to answer questions about this text?" This broad JOL prompt can produce a variety of cognitive processes that likely differ from processes elicited at test, resulting in a poor correlation between learners' predicted and actual performance. This general JOL strategy may be even less effective during map learning, since maps contain a variety of components that participants could attend to during learning. With this in mind, the current study prompted participants with specific JOLs that required them to rate their likelihood to remember routes or relative landmark locations on a map.

In summary, because metacognition is generally a valuable contributor to performance outcomes during complex multimedia learning, and metacognitive monitoring prompts aid memory for verbal learning materials, there is reason to believe that JOLs will promote successful learning outcomes during map learning. Because maps contain many components, it's important

to gain a better understanding of the conditions necessary for effective metacognitive monitoring prompts. We suggest that successful metacognitive monitoring bolsters learning outcomes and predictive accuracy. In other words, when metacognitive monitoring prompts are effective, this should be reflected by 1) participants' future memory and 2) the accurate prediction of their future memory. Therefore, the present study extends findings from the verbal learning literature to the spatial learning context by explicitly engendering specific JOLs during map learning.

To account for the cognitive demands of learning complex spatial stimuli, we consider whether metacognitive strategies are best implemented through a *congruent specificity framework.* Based on cognitive load theory, learners should be explicitly guided to cognitively manipulate material in a way that is consistent with learning goals (Kirschner et al., 2006). Therefore, to promote effective metacognitive monitoring, we test whether learning goals and metacognitive monitoring prompts should be both specific and congruent with one another. Referring to the earlier example, we measure whether learners who are directed to learn *rout*es and monitor their memory for *routes* have greater memory and metacognitive accuracy than those who monitor their memory for *relative locations.*

Furthermore, existing literature finds that spatial tasks demanding recall from the learned perspective are easier than those requiring recall from a perspective different than the one learned (Muffato et al., 2019). Therefore, we also assess whether metacognitive monitoring prompts and tests should be congruent. We measure whether learners who are prompted to monitor their memory for *routes* and are tested on *routes* have greater memory and metacognitive accuracy than those who are tested on *relative locations.*

Strengthening the Scaffold: Congruent Specificity Framework

We propose that learning goals and metacognitive monitoring can serve as scaffolds to support learning and memory for complex spatial stimuli such as maps. Now, we consider ways to strengthen this scaffold. We consider successful metacognitive monitoring to consist of two parts. First, when metacognitive monitoring is effectively engendered, memory at test should be accurate. Second, future memory predictions should be accurate. In other words, a positive relationship between predicted memory performance and actual memory performance should be observed. Yet, research frequently shows that JOLs do not always align with objective performance. Despite many attempts to identify which factors play a role in poor accuracy, conclusions remain mixed and are primarily tailored toward less complex learning materials, as suggested by Burkett and Azevedo (2012). Therefore, we introduce a congruent specificity framework to strengthen the scaffold and promote memory and metacognitive accuracy as observable outcomes. In the present study, we explore two congruent specificity framework components, namely congruency between learning goals and JOL prompts, and congruency between JOL prompts and information tested.

Learning Goal-JOL Congruency

One of the congruent specificity framework components we explored is learning goal-JOL congruency. Said differently, we tested whether a match between learning goals and JOL prompts resulted in better memory accuracy and metacognitive accuracy than when learning goals and JOL prompts were incongruent. As previously described, providing learners with specific learning goals should direct their attention to relevant map components, reduce cognitive load, and promote accurate memory. Furthermore, based on the multimedia literature, for metacognition to be effective, the metacognitive processes engaged must be relevant to learning goals (Cuevas et al.,

2002; Mayer, 2001; Scott & Schwartz, 2007). This proposal is further supported by work on guidance during learning, which states, "the goal is to give learners specific guidance about how to cognitively manipulate information in ways that are consistent with a learning goal and store the result in long-term memory" (Kirschner et al., 2006, p. 77). Therefore, in Experiment 2, participants are prompted with specific JOLs. Critically, the congruency between learning goals and JOL prompts is manipulated.

To our knowledge, we are the first to explicitly assess how congruency between JOL prompts and learning goals impacts memory and metacognitive accuracy. We predicted that prompting learners with JOLs that were congruent with learning goals would lessen extraneous working memory load, where all conscious cognitive processing occurs (Paas et al., 2003), and allow sufficient resources to yield JOLs that predict later performance. This prediction is supported by the notion that performance reflects load (Bunch & Lloyd, 2006). To this end, congruent specificity between learning goals and JOL prompts should promote accurate metacognitive predictions.

In summary, the present study provided participants with specific learning goals and specific JOL prompts. Participants were either instructed to learn routes between landmarks or to learn relative landmark locations, and they rated their likelihood to remember routes or relative landmark locations. To assess the relationship between JOL prompts and learning goals, congruency between learning goals and JOL prompts was manipulated in Experiment 2. We predicted that the congruent learning goal-JOL prompt conditions would result in better memory and metacognitive accuracy than incongruent learning goal-JOL prompt conditions.

JOL-Test Congruency

The second congruent specificity framework component we explored is JOL prompt-test congruency. Said differently, we tested whether a match between what participants were prompted to monitor (JOL) and what information was tested resulted in better memory and metacognitive accuracy than when JOL prompt and tested information was mismatched. Based on the transfer appropriate processing (TAP) hypothesis, the value of acquisition activities should not only be defined relative to particular learning goals, but also to tests that are congruent with these goals (Morris et al., 1977). Morris et al. (1977) found that when they directed participants to attend to rhymes during learning, participants performed better on rhyming tests than when they directed participants to process semantic meanings during learning. The present study extends these classic findings to the spatial learning context by manipulating congruency between JOL prompts and tests of map memory in Experiment 3.

Based on the transfer appropriate monitoring (TAM) hypothesis, judgment accuracy is a direct function of the match between properties of the judgment and test contexts (Dunlosky et al., 2005), however, this hypothesis is largely tested within the verbal learning literature. Our prediction on the important role of congruent specificity for monitoring accuracy between JOLs and test aligns with concluding remarks from Weaver and Kelemen (2003). Specifically, these researchers acknowledge that more complex materials allow for a wider range of encoding strategies and processing during judgments and test; therefore, it might be important to match processing for such materials. However, the more complex stimulus materials referred to by these researchers are passages of text, which are arguably less complex than spatial map displays. Therefore, in Experiment 3 we extend findings from the verbal learning literature to more complex map learning materials. Participants either rated their likelihood to remember routes or relative

landmark locations, and they were tested on their memory for routes or relative landmark locations. We predicted congruent JOL prompt-test conditions would result in better memory accuracy and metacognitive accuracy than incongruent JOL prompt-test conditions.

Exploring the Role of Cues

A final point addressed in the present study is the role of cues for memory accuracy and judgment formation. Little is known about what cues people refer to during map learning or how these cues impact memory and metacognitive judgment formation about learned components. In the present study, we embraced an exploratory approach to gain insight on the role of two possible cues relevant to map learning, namely map complexity and time across the learning span.

Map Complexity

In real-world settings, some map information is more complex than others. For instance, the route between the hotel and the nearby restaurant is shorter than the route between the hotel and the airport we fly into. Usually when looking at a shorter route on map, we see fewer landmarks than we do when looking at a longer route. In other words, there is often a direct relationship between route complexity and landmark complexity. In the present study, we explored the role of map complexity.

Variation in map complexity is important because it may play a role in memory accuracy, metacognitive judgment formation, and even the relationship between predicted and actual memory for map components, otherwise described as metacognitive accuracy. Complex learning materials contain a variety of information that learners can attend to and encode or fail to encode (Weaver & Kelemen, 2003). Based on TAP and TAM hypotheses, if there is alignment between the cognitive processes triggered by the judgment and test prompts, then judgment accuracy should be observed (Dunlosky et al., 2005).

It is also possible that learners could base their judgments on a mnemonic sense that they will or will not remember the map information. If people perceive more complex maps as harder to remember, this would reflect in lower JOL magnitudes. It is also possible that highly complex map information is in fact more difficult to remember. Whether or not someone actually remembers the information at test dictates whether a positive relationship between JOLs and memory is observed. For example, if they rate themselves as unlikely to remember information and are unable to remember it at test, we would observe a positive relationship between the JOL and the memory score.

Alternatively, it could be the case that highly complex map information is salient, or memorable to learners. Furthermore, attending to highly complex map information may also result in a comprehensive or detailed representation of map components. With these two considerations in mind, learners could rate themselves as more likely to remember highly complex map information, as indicated by a high JOL magnitude. Whether or not the cues utilized to form the judgment are predictive of memory at a later test should dictate the observed relationship between JOLs and memory accuracy. If the cues learners refer to when forming their JOL are predictive of actual memory at test, then a positive relationship between the JOL and memory score should again be observed. In the present study, we explored memory accuracy and JOL magnitude as a function of learned complexity.

Time

The next cue that we explore in the present study is time. For the purpose of this work, we conceptualize time as indicated by trials across the learning span. To this point, it remains unknown whether JOL magnitudes change across the spatial learning span. Due to the cognitive demands of learning complex spatial stimuli, learners could progressively rate themselves as less likely to

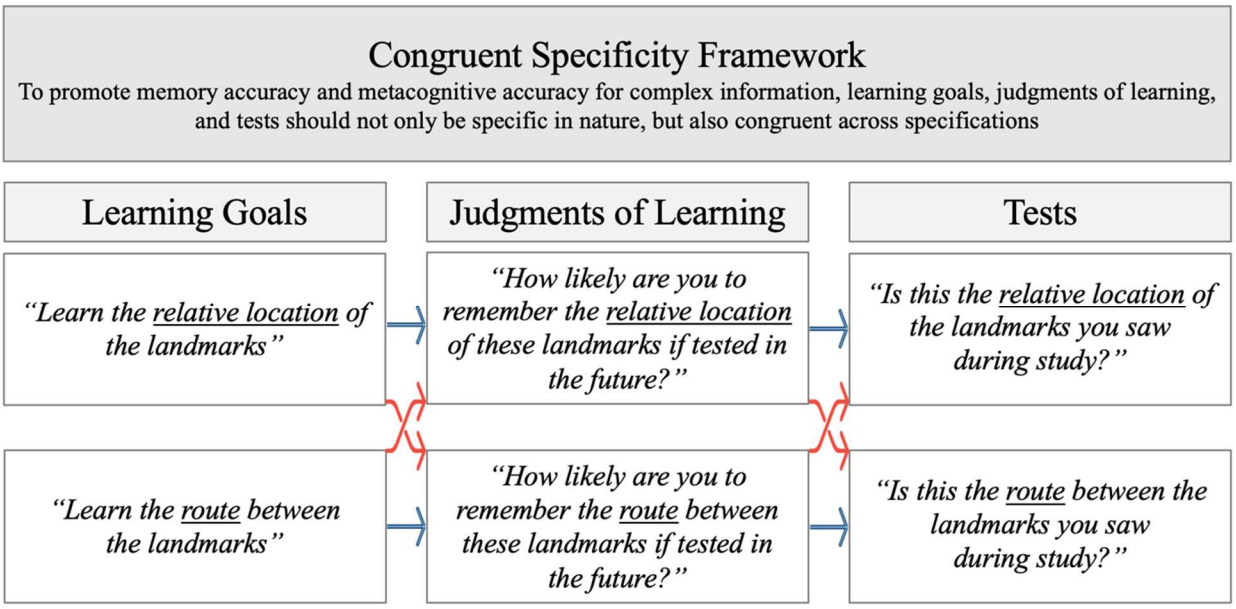
remember map components as information accumulates and memory demands increase. This would result in higher JOL magnitudes early on in learning, and lower ones later in learning. It is also possible that learners' strategies could improve across the learning span. Improved learning strategies over time could cause learners to feel more likely to remember information in the future. This would reflect as lower JOL magnitudes early in learning, and higher JOL magnitudes later in learning. In the present study, we explored the distribution of JOL magnitudes across time and as a function of learned complexity.

The Present Study

The overarching purpose of Experiments 2 and 3 was to consider the conditions necessary for effective metacognitive monitoring prompts. We suggest that effective JOL prompts should bolster memory and metacognitive accuracy. In this sense, effective JOL prompts should work to scaffold map learning. We introduced a congruent specificity framework to strengthen this scaffold and promote memory and metacognitive accuracy as observable outcomes. With this framework, we proposed that specificity alone is not enough to aid memory during complex learning. To strengthen the scaffold provided by assigned learning goals and JOL prompts, specifications must also be congruent with one another (Figure 8).

Figure 8

Congruent Specificity Framework



Note. The blue arrows represent congruent learning goal/test-JOL prompt, and learning goal/JOL prompt-test scenarios. The red arrows represent incongruent learning goal/test-JOL prompt and learning goal/JOL prompt-test scenarios.

In Experiments 2 and 3, participants were assigned a specific map learning goal, studied complex map displays, made JOLs about specific map components, and were tested on their memory for specific map components. With this design, congruence could exist between learning goals and JOL prompts and between JOL prompts and tested information. First, we considered the role of congruent learning goals and JOL prompts in Experiment 2. To assess the relationship between JOL prompts and learning goals, congruency between learning goals and JOL prompts was manipulated. Next, Experiment 3 considered the role of congruent JOL prompts and tests. To assess the relationship between JOL prompts and tests, congruency between the two was manipulated.

Experiment 2

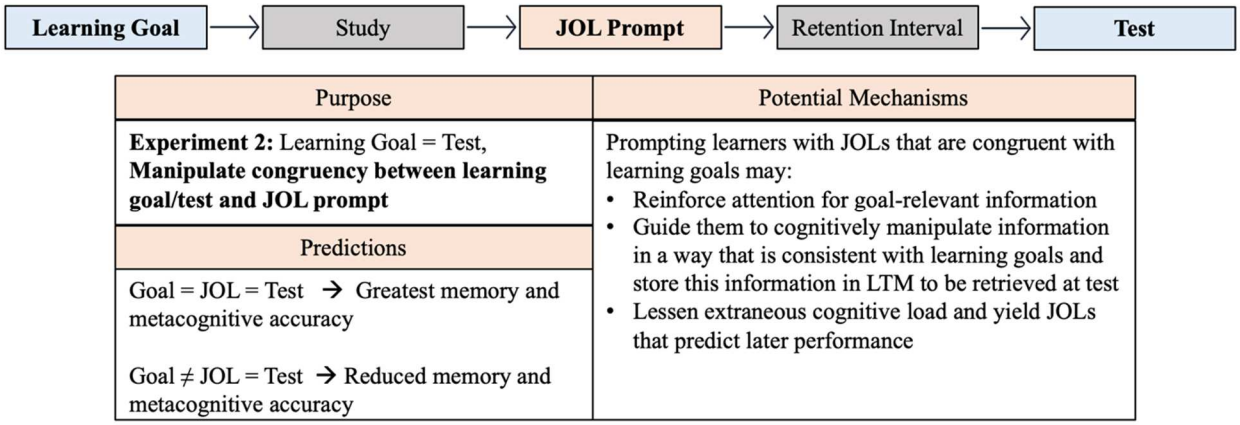
In Experiment 2, we considered the role of congruent learning goals and JOLs prompts. Participants were assigned explicit learning goals for particular map components, and they were prompted with specific JOLs. To assess the relationship between JOL prompts and learning goals, congruency between learning goals and JOL prompts was manipulated. Participants were either instructed to learn routes between landmarks (route goal) or to learn relative landmark locations (survey goal), and they rated their likelihood to remember routes or relative landmark locations. We predicted congruent learning goal-JOL prompt conditions would result in better memory accuracy and metacognitive accuracy than incongruent learning goal-JOL prompt conditions. To test this, we used a novel map learning paradigm.

Participants were assigned to a survey or a route learning goal. They were given instructions to learn the relative location of the landmarks or to learn the routes between landmarks. They studied maps containing a starting landmark, an ending landmark, and a route between the two. There were also contextual landmarks scattered across the map that were held constant throughout the experiment. After studying each map, they either rated their likelihood to remember the relative location of the landmarks (survey), or the route between the landmarks (route). After studying all maps, participants received a survey or route memory test. Participants completed a 2-AFC test to assess their memory for map components. Participants taking a survey test saw two maps that each contained landmarks. On one of the maps, the target (destination) landmark was moved to a different position than what they saw during learning (lure). Participants taking a route test saw maps that contained drawn routes between landmarks. On one of the maps, the route was different than what they saw during study (lure). Participants were asked to select the correct map that they saw during study.

We anticipated better memory accuracy when assigned learning goals and JOL prompts were congruent as opposed to when they were incongruent. We also predicted that congruent learning goals and JOL prompts would aid participants' ability to accurately predict their future memory on a test that matched the learning goal. We predicted better metacognitive accuracy when assigned learning goals and JOL prompts were congruent. For further conceptualization of the purpose, predictions, and potential mechanisms at play in Experiment 2, see Figure 9.

Figure 9

Visual conceptualization of the purpose, predictions, and potential mechanisms at play in Experiment 2



Note. In this experiment, learning goals and tests are always matched, while congruency between learning goals and JOL prompts is manipulated.

Method

Participants

Participants were recruited through Prolific, an online participant recruitment platform. They were compensated with the standard Prolific rate. All participants were 18-35 years old and within the United States. They were required to use a desktop to complete the experiment.

This experiment implemented novel spatial stimuli in an entirely novel paradigm. Therefore, due to the exploratory nature of this work, we were unable to conduct a representative power analysis. As in Experiment 1, the desired sample sizes for the present experiments in this study were determined by consideration of previous work from the 1) verbal 2) spatial and 3) multimedia learning literatures.

Design

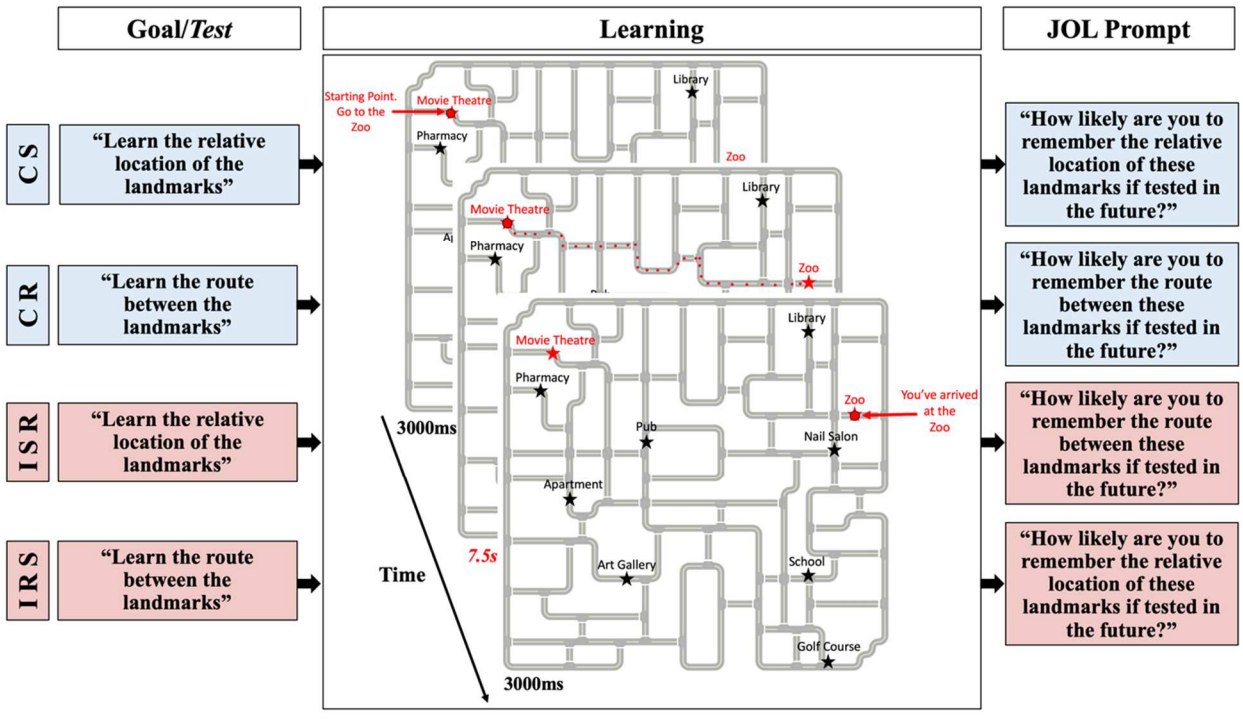
A 2 (Assigned Goal/Test: Survey, Route) x 2 (JOL Prompt: Survey, Route) between-subjects design was employed. Participants were assigned either a survey or route goal and test, and they made JOLs about survey or route information after each study trial. The assigned learning goal perspective always matched the test perspective. In other words, if a participant was assigned a survey learning goal, they were always tested on their survey memory. The same was true for route learning goals and tests.

All participants were assigned a learning goal, made JOLs after each learning trial, and were tested on their memory. They were assigned to one of four conditions. Two of the conditions contained congruent learning goals and JOL prompts. In the congruent Survey Goal/JOL/Test condition, they were assigned a survey learning goal and made JOLs about survey information. They were also tested on survey information. In the congruent Route Goal/JOL/Test condition, they were assigned a route learning goal and made JOLs about route information. They were also tested on route information.

The remaining two conditions contained incongruent learning goals and JOLs prompts. In the incongruent Survey Goal/Test, Route JOL condition, they were assigned a survey learning goal and tested on survey information, but they made JOLs about route information. In the incongruent Route Goal/Test, Survey JOL condition, they were assigned a route learning goal and tested on route information, but they made JOLs about survey information. For a visual schematic of the separate conditions comprising the experiment, as indicated by assigned learning goal-JOL prompt congruency, see Figure 10.

Figure 10

Schematic of conditions, as indicated by assigned learning goal-JOL prompt congruency, in Experiment 2



Note. CS= Congruent Survey Goal/JOL/Test; CR= Congruent Route Goal/JOL/Test; ISR= Incongruent Survey Goal/Test, Route JOL; IRS= Incongruent Route Goal/Test, Survey JOL

Materials

Landmark name pairs. There were two lists of 8 landmark pairs (16 total landmarks; Appendix D). Landmark names were pulled from Experiment 1.

Context landmarks. In addition to the 8 landmark pairs, there are 8 additional context landmarks. These landmarks were present on the maps throughout the study and test but were not part of the test. For a map containing only the 8 context landmarks, see Appendix E.

Maps. The first author developed the map stimuli using the VR environment map from Experiment 1 as a base. This map was then modified to incorporate additional map features. Each of the individual map stimuli contain a starting landmark destination, an ending landmark destination, and an animated route between the two landmarks. In addition to the starting and ending landmarks, there are also the 8 context landmarks on each of the individual stimuli. The individual map stimuli were created using Microsoft PowerPoint. The maps were approximately 30 x 30 cm in size. For a full map containing all landmarks, see Appendix F. Note that participants never saw a complete map with all landmarks provided.

For the purpose of the test, lure maps were also generated. Survey lures contained a target, or "destination" landmark that had been moved so that it was no longer in the same location shown during study. To maintain consistency across survey lures, the author set a specific circumference range within which the new placement could fall. Route lures showed a different route between the cue, or "starting point" landmark and the target, or "destination" landmark than shown during study. To maintain consistency across route lures, the first author strategically designed the average number of turns for lure routes to mirror the average number of turns for routes seen during study. The average number of turns for the studied routes was 6.8, while the average number of turns for the lure routes was 6. Participants were shown 32 maps at test. Half (16) of them were

the same as those seen during study and half (16) of them were lures. The true and lure maps were shown in pairs, for a total of 16 test trials.

Map Learning Task. The experiment was carried out over the internet using the Gorilla experiment building system. The map task was split into three phases: practice, learning, and test.

Practice Phase. Prior to beginning the experimental task, participants completed two practice trials. The practice trials were representative of their assigned condition. In other words, if a participant is assigned to the congruent Survey Goal/JOL/Test condition, then during practice they were instructed to learn the relative location of the landmarks, rated their likelihood to remember the relative location of the landmarks in the future, and were tested on their memory for relative landmark locations. The stimuli presentation order was randomized during learning and test.

Learning Goals. Participants were provided with an assigned learning goal. If they were in the congruent Survey Goal/JOL/Test condition, or the incongruent Survey Goal/Test, Route JOL condition, they were instructed to learn the relative location of the landmarks. In contrast, if participants were in the congruent Route Goal/JOL/Test condition, or the incongruent Route Goal/Test, Survey JOL condition, they were instructed to learn the route between landmarks.

Next, they saw a map that contained two critical landmarks in red font. One was labeled "Starting Point" and told them where to go next. As an example, a practice trial had an arrow pointing to the landmark "flower" and stated, "Starting Point. Go to the Rock". There were also four context landmarks in black font scattered across the map. Then, participants saw an animated route between the starting point and the destination landmark. After the route animation was complete, the ending point landmark was labeled. An arrow pointed to the landmark and stated, "You've arrived at the Rock".

JOLs. After each trial, they made a JOL. Participants in the congruent Survey Goal/JOL/Test condition or the incongruent Route Goal/Test, Survey JOL condition were prompted with the question, "How likely are you to remember the relative location of these landmarks if tested in the future?" In contrast, participants in the congruent Route Goal/JOL/Test condition or the incongruent Survey Goal/Test, Route JOL condition were prompted with the question, "How likely are you to remember the route between these landmarks if tested in the future?" The response scales ranged from 0 (not likely to remember) to 100 (highly likely to remember).

Tests. After the practice learning phase, the 2-alternative forced choice (2-AFC) practice test began. Participants were shown two maps that each contained two critical landmarks in red and a route between the two landmarks. There were also four context landmarks scattered across the maps. One of the maps shown was the same as the map seen during test, and one of the maps contained a lure. In the congruent Survey Goal/JOL/Test condition and the incongruent Survey Goal/Test, Route JOL condition, participants were shown a map that contained a target or "destination" landmark that had been moved from the location shown during study. The cue, or "starting point" landmarks were always in the correct place as in study. Participants were told to "Select the map that contains the relative location of the landmarks that you learned during study."

In contrast, in the congruent Route Goal/JOL/Test condition and the incongruent Route Goal/Test, Survey JOL condition, participants were shown a map that contained a route between the "starting point" and "destination" landmark that was different from the route shown during study. Participants were told to "select the map that contains the route between the landmarks that you learned during study." Participants clicked on the map to select it.

Participants received feedback on their response following the first practice test trial only. If participants taking a survey test selected the correct map, they were shown, "Correct! That is the map that contains the relative location of the landmarks that you learned during study." Participants taking a route test were shown, "Correct! That is the map that contains the route between the landmarks that you learned during study."

If participants selected the incorrect (lure) map, they were shown, "Oops, that is not the correct answer. The correct map that contains the relative location of the landmarks (survey test conditions) or the route between the landmarks (route test conditions) that you learned during study is shown below. Participants were then shown the correct map that they should have selected.

Following the second practice test trial, participants were prompted to provide a confidence rating about their map response. After selecting the map they believed to be correct, they were asked, "How confident are you that the map you selected is the correct map?" Participants rated their confidence on a 5-point scale ranging from 1 (Complete guess) to 5 (Absolutely sure). After the practice was complete, the learning phase of the experimental task began.

Learning Phase. During the experimental learning phase, participants studied a series of 16 map stimuli. At the start of each trial, prior to viewing the map stimuli, they were reminded of their assigned learning goal. As an example, a participant who was randomly assigned to a survey learning goal condition was instructed, "Learn the relative location of the landmarks". A participant who was randomly assigned to a route learning goal condition was instructed, "Learn the route between the landmarks". Then the map stimulus was presented. Each of the maps contained two critical landmarks in red font. Just as in practice, one of the landmarks was labeled "Starting Point" and told them where to go next. As an example, a trial had an arrow pointing to the landmark "Movie Theatre" and stated, "Starting Point. Go to the Zoo". Next, participants saw

an animated route between the starting point and the destination landmark. After the route animation was complete, the ending point landmark was labeled, such as, "You've arrived at the Zoo". Therefore, each of the map stimuli contained three events (the starting point, the animated route, and the ending point). The starting point and ending point event remained on the screen for 3000 ms. The route animation ranged from 3-16.5 seconds with the average route lasting around 9.5 seconds. The timing of routes was dependent on path complexity. There were also 8 context landmarks in black font scattered across the map. These context landmarks remained constant throughout the experiment.

Participants made a JOL after viewing each of the stimuli. Participants in the congruent Survey Goal/JOL/Test condition received a survey learning goal and made JOLs about survey information. They were told to, "Learn the relative location of the landmarks." Then, when making a JOL, they were asked, "How likely are you to remember the relative location of these landmarks if tested in the future?" Participants in the congruent Route Goal/JOL/Test condition received a route learning goal and make JOLs about route information. They were told to "Learn the route between the landmarks." Then, when making a JOL, they were asked, "How likely are you to remember the route between these landmarks if tested in the future?"

In contrast, participants in the incongruent Survey Goal/Test, Route JOL condition were assigned a survey learning goal, but they made a JOL about route information. They were told to, "Learn the relative location of the landmarks," but their JOL prompt asked, "How likely are you to remember the route between these landmarks if tested in the future?" Lastly, participants in the incongruent Route Goal/Test, Survey JOL condition were assigned a route learning goal and made JOLs about survey information. They were told to "Learn the route between the landmarks," bur

their JOL prompt asked, "How likely are you to remember the relative location of these landmarks if tested in the future?"

All participants rated their likelihood to remember on a scale from 0 (not likely to remember) to 100 (highly likely to remember). After making the JOL, the next trial began. After studying the 16 map stimuli, the test phase began. The 0-100 rating scale was used to promote JOL-sensitivity.

The first attention check was placed at the end of the study phase and prompted participants to enter a color name in the box provided. Then, they completed the two-minute retention interval before beginning the test phase.

Test Phase. The test phase largely mirrored the practice test implemented during the practice phase at the start of the experiment. Participants completed a 2-AFC test to assess their memory for map components. At the start of the test phase, they were told they would be shown two maps and asked to select the correct map that they saw during study. Then, they rated how confident they were that the map they selected was the correct map that they learned during study.

The context of the test (survey or route) matched the context of the learning goal they participants are assigned at the start of the experiment. Participants randomly assigned a survey learning goal (Survey Goal/JOL/Test and Survey Goal/Test, Route JOL conditions) took a survey test. Those randomly assigned a route learning goal (Route Goal/JOL/Test
and Route Goal/Test, Survey JOL conditions) took a route test.

Participants were shown 32 maps, half of which were the same as the ones seen during learning, and half of which were lures. Participants taking a survey test saw two maps that each contained two critical landmarks in red font and 8 context landmarks in black font. These were the same context landmarks shown during learning. On one of the maps, the critical landmark was in

the same position as in study. In a second map, the target (destination) landmark was moved (lure). Participants had to select the correct map that contained the relative landmark locations they saw during study.

Participants taking a route test also saw two maps that each contained two critical landmarks in red font and the 8 context landmarks in black font. Drawn routes were shown between the two critical landmarks. On one of the maps, the drawn route was the same route they saw during study. In a second map, it was a different route (lure). Participants had to select the correct map that contained the route between the landmarks that that they saw during learning.

After selecting a map, participants were prompted to rate their confidence about their map response. After selecting the map they believed to be correct, they were asked, "How confident are you that the map you selected is the correct map?" Participants rated their confidence on a 5-point scale ranging from 1 (Complete guess) to 5 (Absolutely sure). Then the next test trial began.

The second attention check was placed after the test phase and had participants select the extent to which the statement "I was born before 1920" applied to them.

Questionnaires. After the test phase of the experiment, participants completed the same questionnaires as in Experiment 1. The encoding-based strategies questionnaire was modified for the purpose of the experiment (Appendix G).

Experiment 2 Results

In Experiment 2, we collected data from 120 participants. We excluded data from 17 participants due to technical difficulties. Additionally, 1 participant was identified as an extreme outlier and removed from analyses. We analyzed data from 102 participants ages 21-35 years. Most participants identified as white (75%) males (65%). Participants came from a wide array of educational backgrounds, with 61% holding a bachelor's degree or greater.

Memory and Metacognitive Accuracy Results

We estimated linear mixed models (LMMs) to assess our primary hypotheses (see Table 3). Before analyzing the memory accuracy data, we gave participants a score of zero for any test trial where they reported guessing. Therefore, the following memory accuracy analyses take guessing into account. Though we our hypotheses predicted a difference between groups based on congruency, it was also possible that goals/test types could drive differences in memory accuracy. We looked at the conditions in experiment 2 and 3 models separately due to this reason.

First, we tested whether memory accuracy varied between conditions. The results of this model (Table 3, E2M1) were non-significant, and did not fit the data better than the null model (Table 3, E2MN). Next, we tested whether JOL ratings were a significant predictor of memory accuracy and whether better metacognitive accuracy was observed in congruent as compared to incongruent conditions (Table 3, E1M2).

Table 3

Model formulas and associated AIC values predicting memory accuracy

Model	Model formula	AIC
E2MN	Accuracy~ 1 + (1|PID)	2190.5
E2M1	Accuracy~ Condition + (1|PID)	2195.1
E2M2	Accuracy~ Condition*JOL + (1|PID)	2171.1

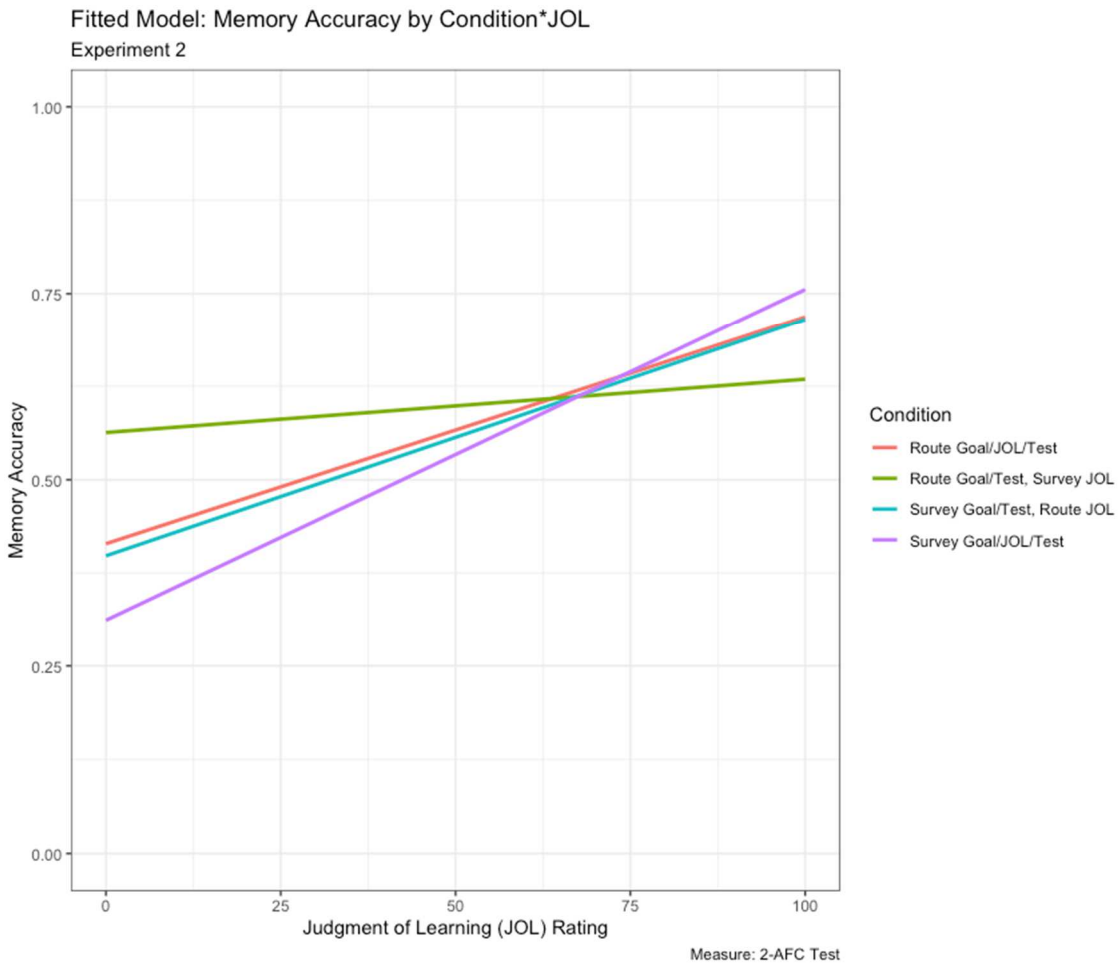
We saw improvement in the overall model fit by adding JOL to the model.

We observed a main effect of JOL ($p < .001$), meaning that JOL ratings did predict test accuracy. (Fig 11). The incongruent Route Goal/Test, Survey JOL condition was less accurate at predicting

their memory performance only when compared to the congruent Survey Goal/JOL/Test condition, respectively.

Figure 11

E2M2. Mean memory accuracy as a function of JOL and condition



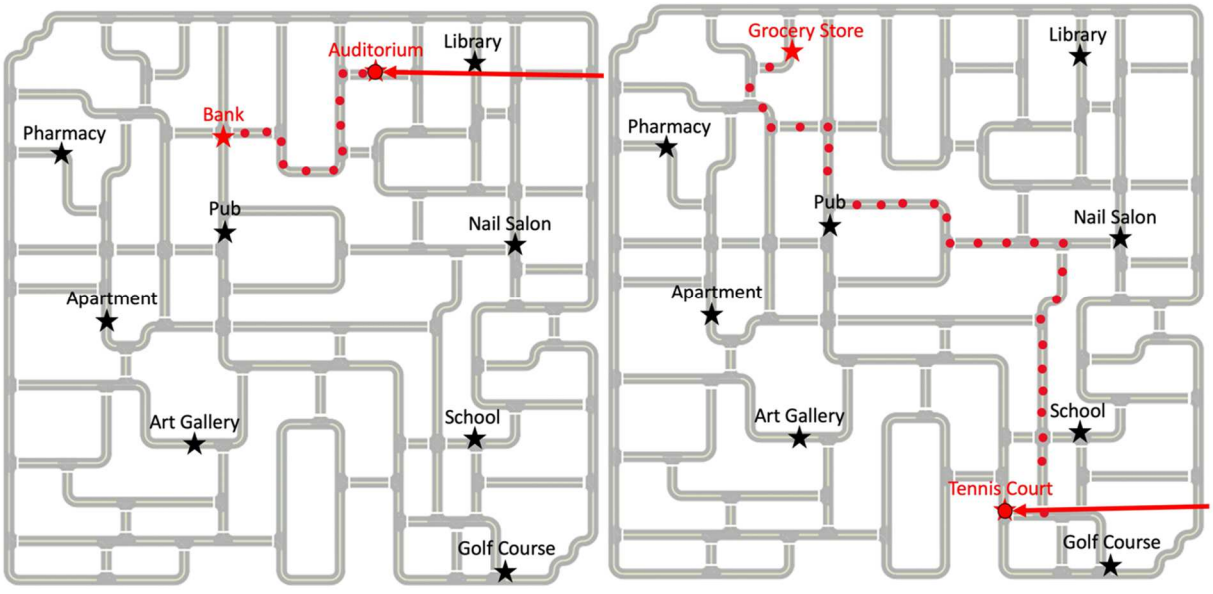
Note: This model was calculated using a centered JOL variable. For interpretability, it is plotted here using the 0-100 JOL scale.

Exploratory Memory and Metacognitive Accuracy Results

Next, we conducted exploratory analyses to assess whether map complexity impacted memory accuracy and metacognitive accuracy respectively, and whether these relationships varied across conditions. To determine learned complexity, independent coders calculated the number of turns per route on each of the 16 studied maps. From this metric, it is inherent that when routes have more turns, more landmarks exist in route proximity on average. Therefore, variation in route complexity can also translate to variation in landmark complexity on studied maps. For analysis purposes, maps containing routes with 6 or fewer turns were categorized as "low complexity", while those with 7 or more turns were categorized as "high complexity." Examples of a low and high complexity map are shown below (Fig 12).

Figure 12

A low complexity (left) and high complexity map (right)



Note: Low complexity maps contain routes with fewer turns and therefore fewer landmarks exist in route proximity on average, as compared to higher complexity maps

First, we tested whether learned complexity played a role in memory accuracy and if this role differed between groups. To address this question, we entered condition*learned complexity into the model (Table 4, E2M3).

Table 4

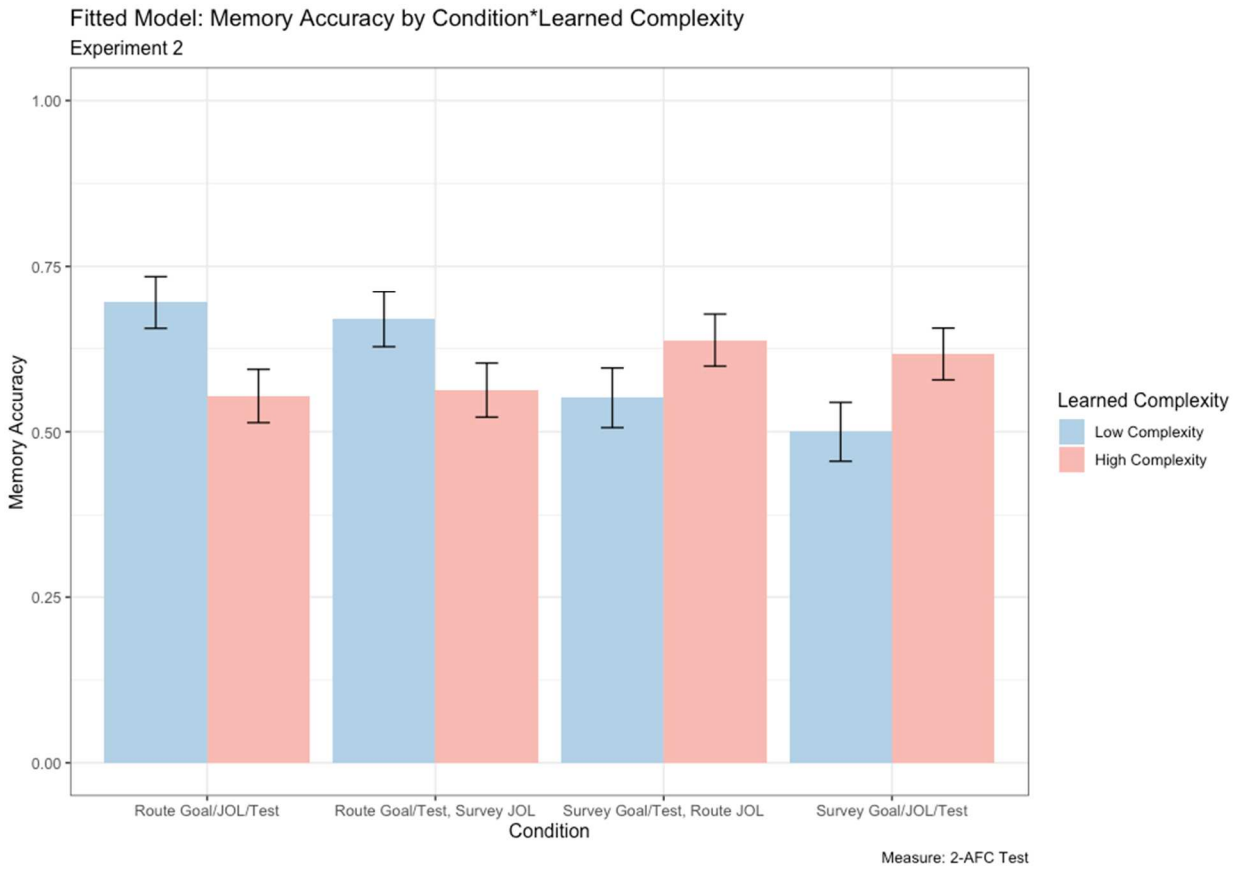
Exploratory model formulas and associated AIC values predicting memory accuracy

Model	Model formula	AIC
E2M3	Accuracy ~ Condition*Learned Complexity + (1|PID)	2181.8
E2M4	Accuracy ~ Condition*JOL*Learned Complexity + (1|PID)	2160.3

We saw a main effect of learned complexity ($p < .05$) and a complexity*condition interaction (Fig 13). While memory accuracy for participants with survey learning goals/tests was lower for low, compared to high, complexity maps, the opposite trend was observed for participants with route goals/tests. Namely, memory accuracy for those with route learning goals/tests was higher on low, compared to high, complexity maps ($p < .005$). The fit of the model with complexity improved compared to the model with condition alone.

Figure 13

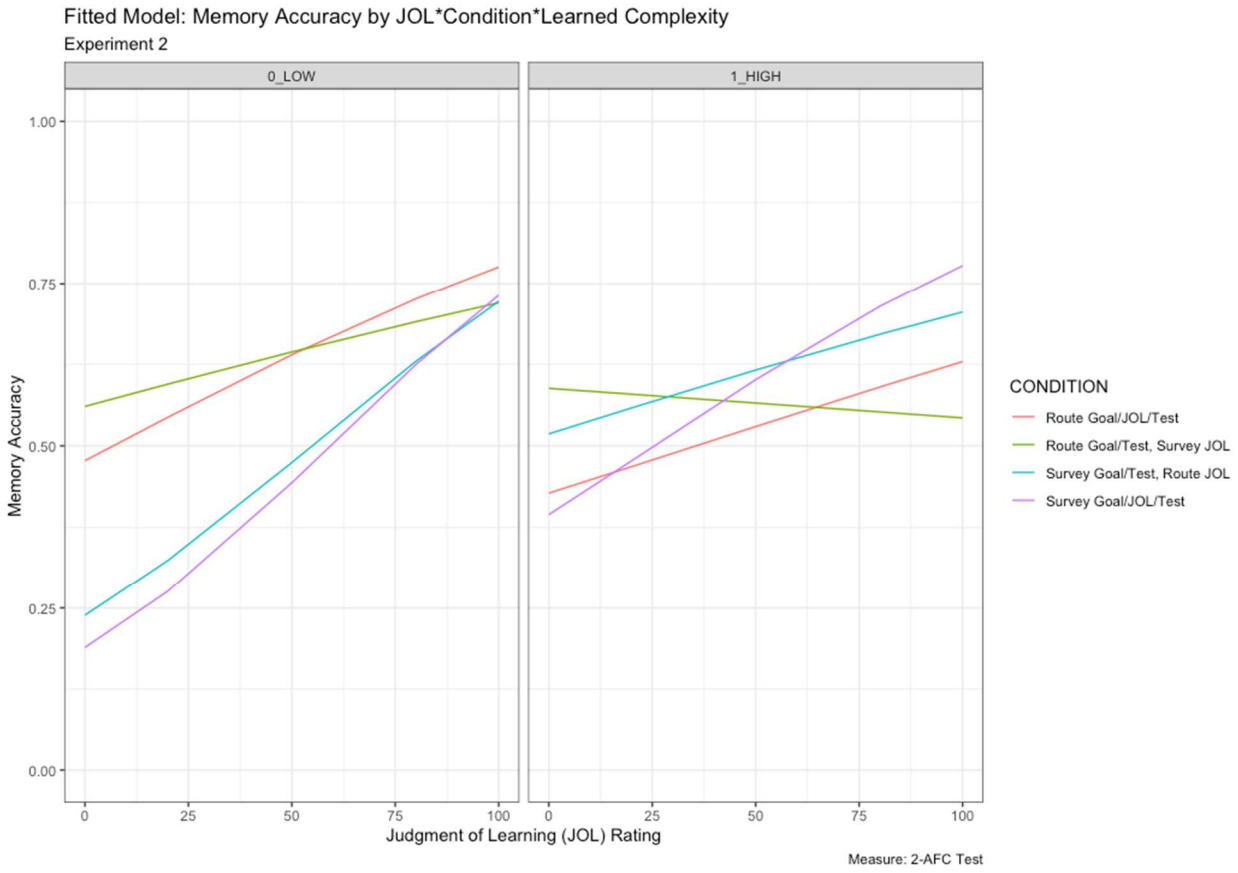
E2M3. Mean memory accuracy as a function of condition and learned complexity. Error bars show SE.



Next, we tested whether learned complexity played a role in metacognitive accuracy and if this role differed between groups (Table 4, E2M4). We entered condition*JOL*learned complexity into the model. While a generally positive relationship was observed between JOL ratings and memory accuracy across conditions, the same could not be said for the incongruent Survey Goal/Test, Route JOL condition on high complexity maps based on visual interpretation (Fig 14). Though we saw an improvement in the model fit by adding JOL, none of the three-way interactions were significant.

Figure 14

E2M4. Mean memory accuracy as a function of JOL across levels of learned complexity and condition



Exploratory JOL Magnitude Results

Next, we explored factors that might predict JOL magnitude. We began by testing whether condition alone predicts JOL magnitude (Table 5, E2M5).

Table 5

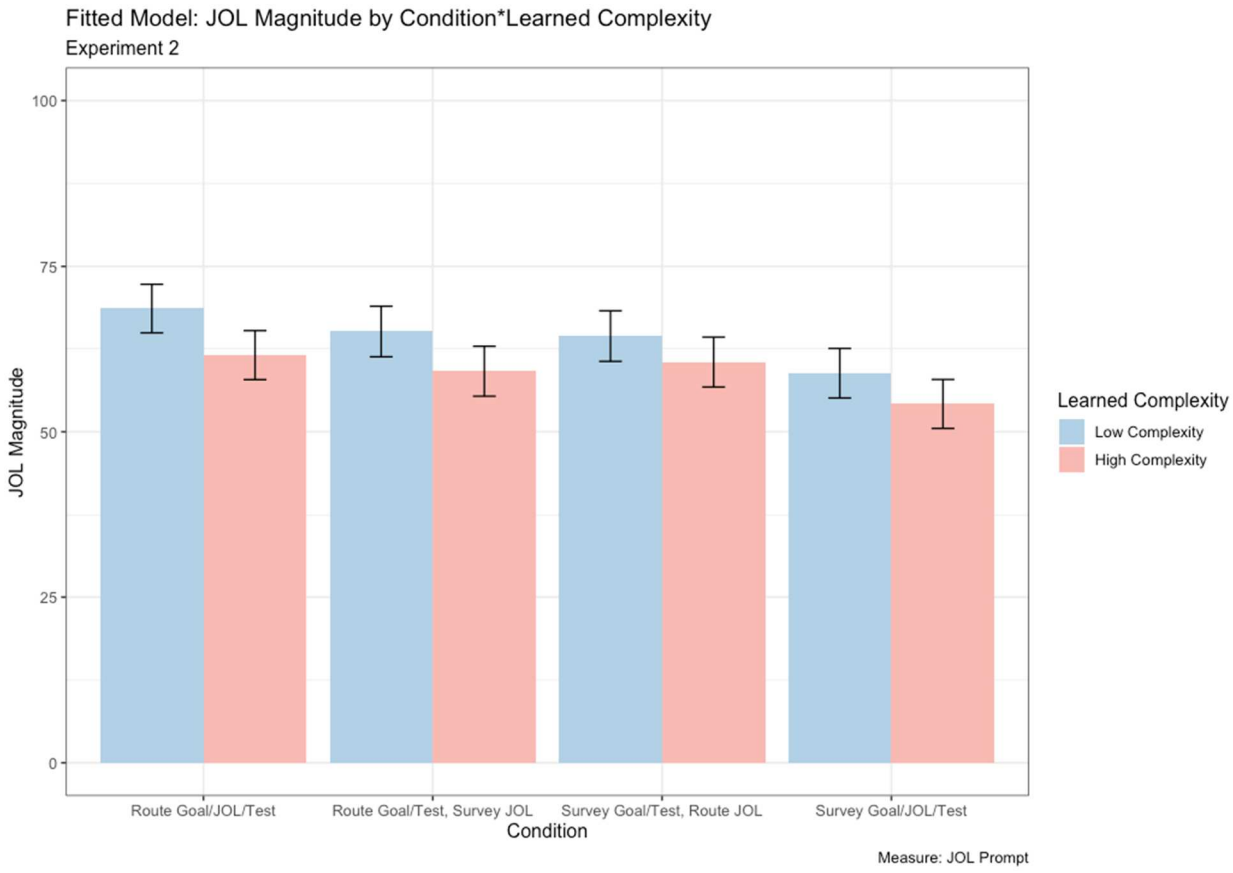
Exploratory model formulas and associated AIC values predicting JOL magnitude

Model	Model formula	AIC
E2MNJ	JOL ~ 1 + (1|PID)	14155.2
E2M5	JOL ~ Condition + (1|PID)	14143.64
E2M6	JOL ~ Condition*Learned Complexity + (1|PID)	14101.9
E2M7	JOL ~ Condition*Learning Trial + (1|PID)	13995.81

Though this model was a better fit than the null model, it was non-significant, meaning the groups did not differ in their average rating magnitude. Next, we entered learned complexity into the model (Table 5, E2M6). We observed a main effect of learned complexity ($p = .005$), such that higher JOLs were observed for less complex maps across all groups (Fig 15). The model fit significantly improved with the addition of learned complexity as a predictor.

Figure 15

E2M6. Mean JOL magnitude as a function of condition and learned complexity. Error bars show SE.

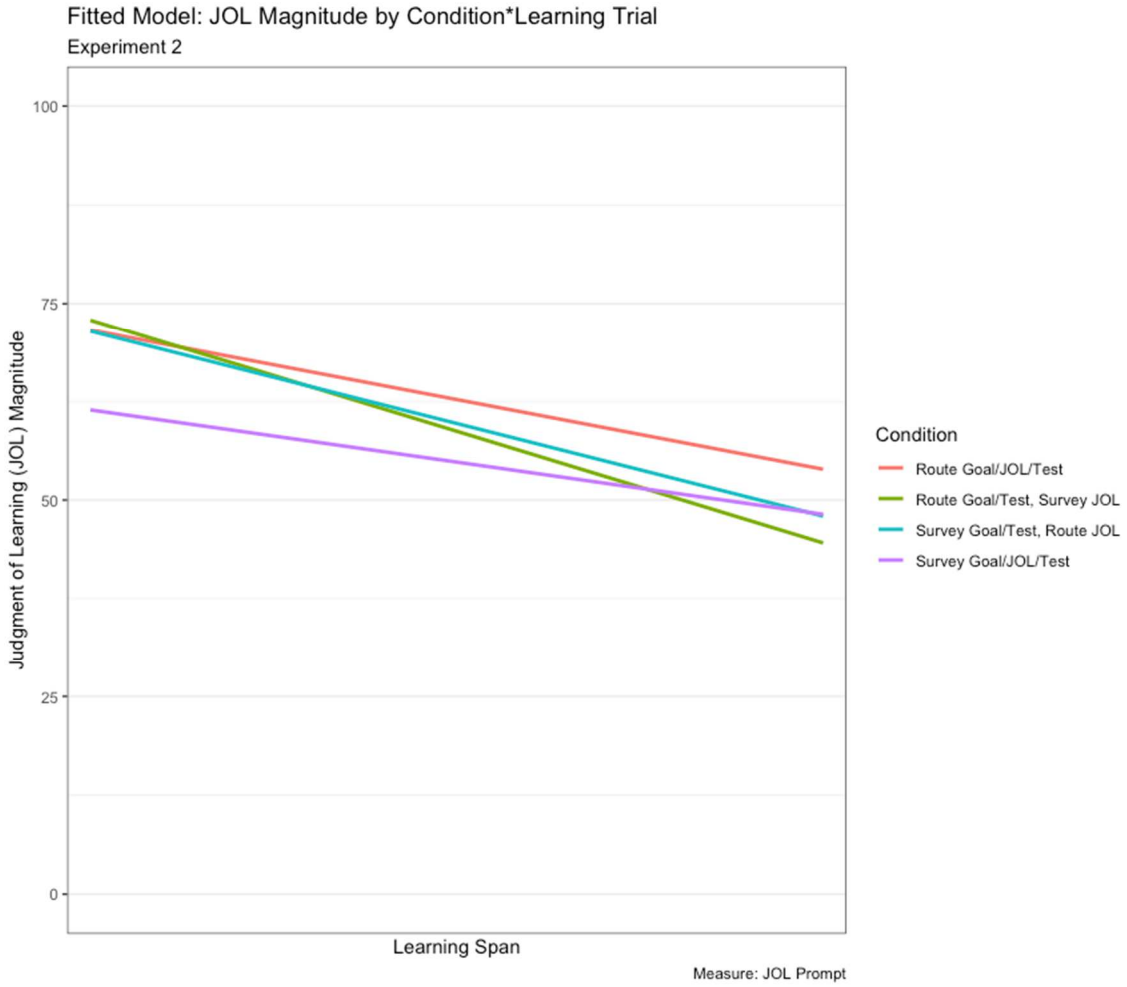


Next, we tested whether JOLs differed as a function of time (across learning trials), and whether this relationship differed across groups. We entered condition*learning trial into the model (Table 5, E2M7). We observed a main effect of learning trial ($p < .001$), such that JOLs decreased across time for all groups (Fig 16). JOL magnitudes decreased at a steeper rate for the incongruent Survey Goal/Test, Route JOL group as compared to the congruent Survey Goal/JOL/Test group ($p < .05$). JOL magnitudes similarly decreased at a steeper rate for the incongruent Route Goal/Test, Survey JOL group as compared to the congruent Route

Goal/JOL/Test ($p < .05$) and congruent Survey Goal/JOL/Test groups respectively ($p < .005$). The fit of this model was significantly better than the previously estimated models predicting JOL magnitude.

Figure 16

E2M7. Mean JOL magnitude as a function of the learning span across conditions



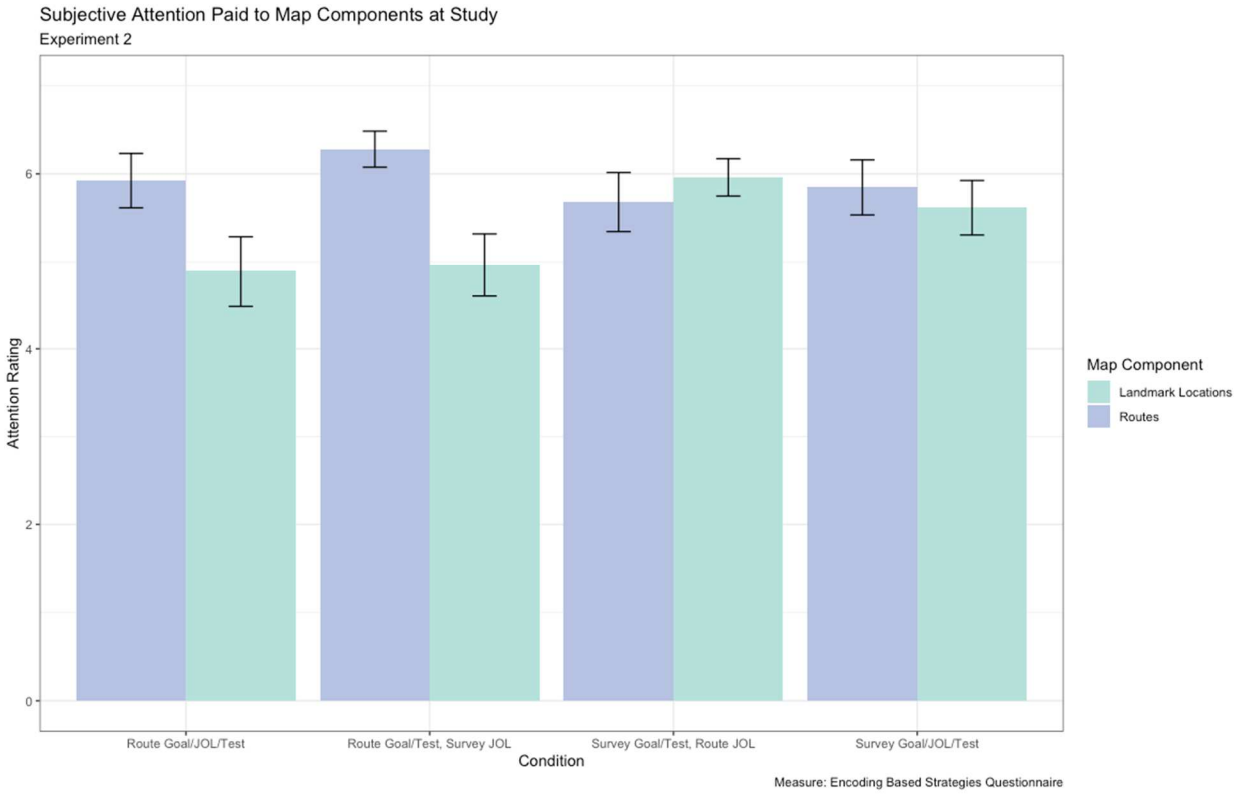
Encoding Based Strategies

Next, we assessed participant responses on the subjective encoding-based strategies (EBS) questionnaire. Specifically, we conducted a mixed factor ANOVA to determine whether the amount of attention participants reportedly paid to different map components (relative landmark

locations and routes between landmarks) differed. We observed a main effect of component ($F(1) = 10.881$, $p = .001$)) and a condition*component interaction ($F(3) = 4.356$, $p = .006$)) (Fig 11). The Route Goal/JOL/Test and Route Goal/Test, Survey JOL conditions reported attending to routes more than relative landmark locations ($F(1) = 9.11$ $p = .02$, $F(1) = 16$, $p = .002$, respectively) (Fig 17).

Figure 17

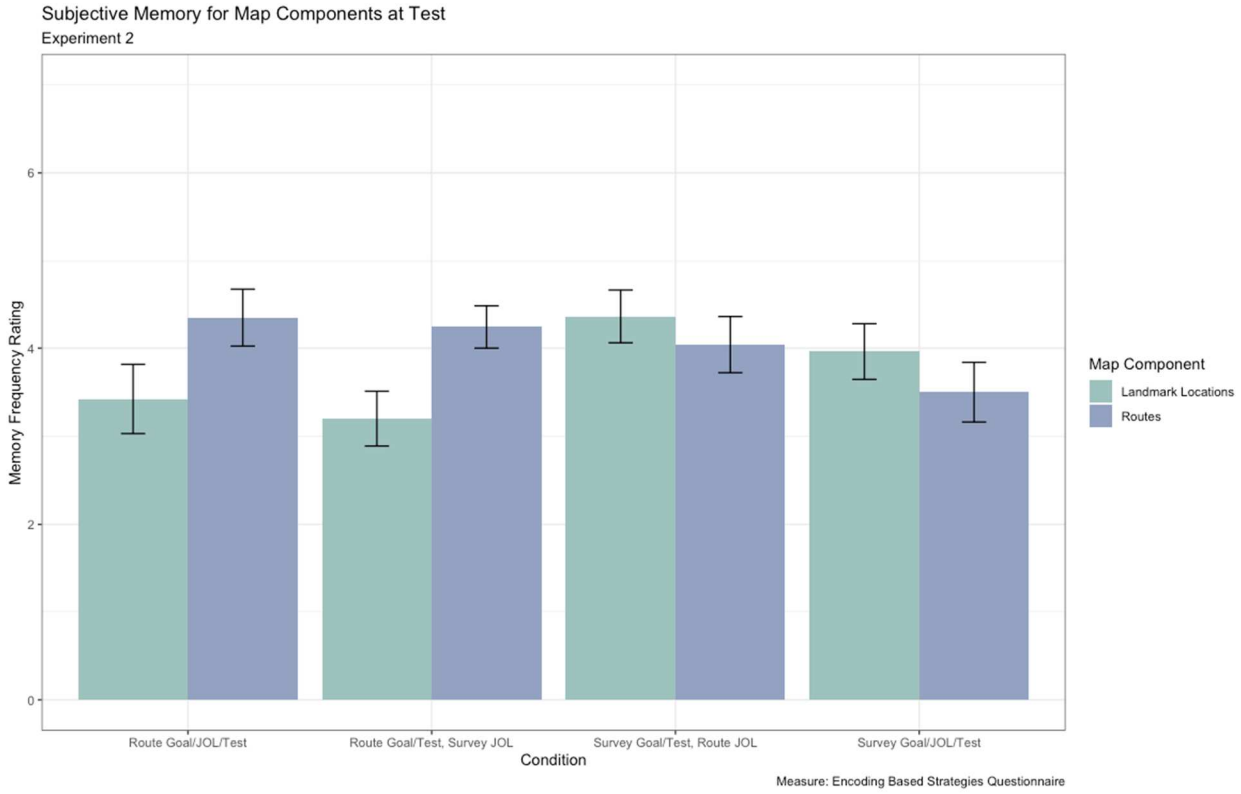
Mean attention rating as a function of condition and map component. Error bars show SE.



The Route Goal/Test, Survey JOL condition also reported remembering route information more than relative landmark locations ($F(1) = 8.02$, $p < .05$)) (Fig 18). This effect was only marginally significant for the congruent Route Goal/JOL/Test group ($F(1) = 5.78$, $p = .09$)) (Fig 12).

Figure 18

Mean memory rating as a function of condition and map component. Error bars show SE.



There were no significant differences between the groups on the single-item sense of direction (SOD) measure.

Experiment 2 Discussion

In Experiment 2, we investigated the relationship between learning goals and JOL prompts. While incongruency between learning goals and JOL prompts did not play a role in memory accuracy, it did contribute to the relationship between JOLs and memory accuracy, otherwise described as metacognitive accuracy. While a positive relationship between JOL ratings and memory accuracy was observed across conditions, this relationship was less pronounced for individuals in the incongruent Route Goal/Test, Survey JOL condition who were assigned route

learning goals, made JOLs about their likelihood to remember relative landmark locations, and were tested on their route knowledge. Yet, the same was not true for those in the incongruent Survey Goal/Test, Route JOL condition. This finding suggests that in complex spatial contexts, congruency may be relevant to metacognitive accuracy, but how information is initially learned matters. It could be the case that if learners' attention is directed to multiple map information streams during learning, it is more advantageous to begin by directing them to relative landmark location information and then to route information. Directing attention to route information first, followed by survey information, seems to impair metacognitive accuracy.

When asked about their attention and memory for particular map components, participants in the Route Goal/Test, Survey JOL condition reported attending to and remembering routes more than relative landmark locations. This suggests that they directed their attention primarily to the goal-oriented information, as opposed to the JOL-oriented information, and that they subjectively remembered the goal-oriented information more often. Their attention and memory patterns closely mirror those in the congruent Route Goal/JOL/Test condition. In sum, these exploratory findings suggest that the role of goals in this context may override the role of JOLs.

Exploratory Discoveries

We estimated exploratory models to gain insight on the role of complexity for memory and metacognitive accuracy, and whether this relationship varies based on learning goal-JOL congruency. Exploratory findings suggest that complexity predicts memory accuracy. Critically, this relationship varied depending on the learning goal. Participants with survey learning goals had higher memory accuracy for highly complex materials, while those with route learning goals had higher for low complexity materials. These findings suggest that map complexity matters, but the

relationship between complexity and memory accuracy varies as a function of the learning goal (and test), irrespective of the JOL.

When we added JOL to the model to assess whether metacognitive accuracy varied based on complexity and congruency condition, a generally positive relationship between JOLs and memory accuracy appeared across conditions. This exploratory evidence suggests that learners can predict their future map memory, even when materials are highly complex and they are tasked with monitoring information different from their learning goal or what they're tested on. Additional research is needed to further understand the role of transfer appropriate monitoring in spatially complex learning scenarios, since learners in this study were relatively accurate at predicting their memory.

Finally, we estimated exploratory models to examine influences on JOL magnitude. Little is known about what cues people use to form metacognitive judgments during complex spatial learning. We considered the role of map complexity and time, specifically. Exploratory results suggest that both map complexity and time, as indicated by learning trial, are significant predictors of JOL magnitude. Higher JOLs are observed for less complex maps, suggesting that people believe themselves to be more likely to remember less complex materials. Furthermore, JOLs decrease across the learning span, suggesting that people refer to the accumulation of information over time as a cue that they will not remember the map information in the future. In Experiment 3, we shifted our focus to assess the relationship between JOLs and tests.

Experiment 3

In Experiment 3 we considered the role of congruent JOL prompts and tests. As in Experiment 2, participants were assigned explicit learning goals for particular map components and were prompted with specific JOLs. To assess the relationship between JOL prompts and tests,

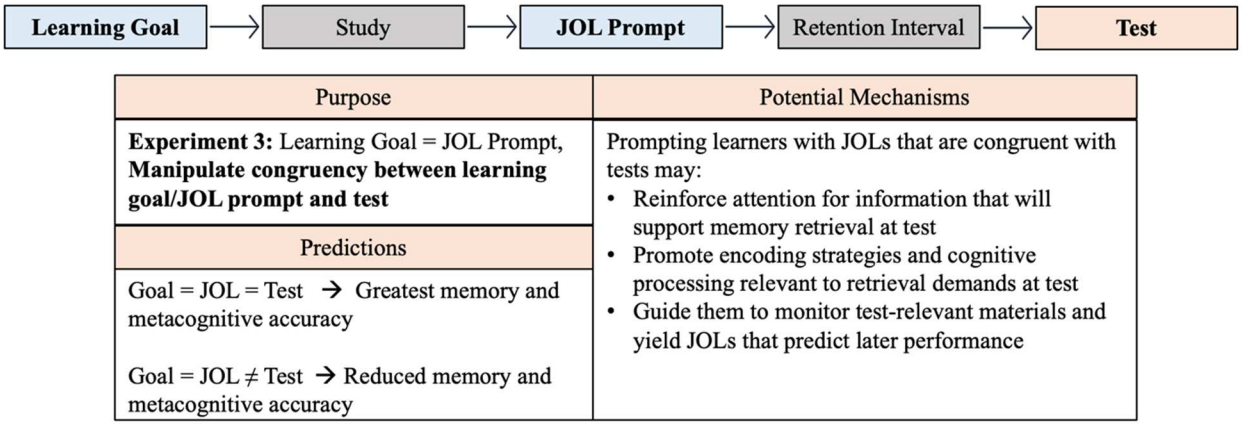
congruency between JOL prompts and tests was manipulated. Participants rated their likelihood to remember routes or relative landmark locations, and they were tested on their memory for routes or relative landmark locations. We predicted congruent JOL prompt-test conditions would result in better memory accuracy and metacognitive accuracy than incongruent JOL prompt-test conditions.

As in Experiment 2, participants were assigned a survey or a route learning goal. They were given instructions to learn the relative location of the landmarks or to learn the routes between landmarks. Next, they studied maps containing a starting landmark, an ending landmark, and a route between the two. There were also contextual landmarks scattered across the map that were held constant throughout the experiment. After studying each map, participants made a JOL about survey or route information. After studying all maps, participants took a survey or route memory test.

In this experiment, goals and JOL prompts always matched, but half of the participants received tests that were incongruent with the goal/JOL prompt. Therefore, we expected better memory and metacognitive accuracy when tests necessarily called upon the information that participants were instructed to learn (goal) and monitored (JOL). For further conceptualization of the purpose, predictions, and potential mechanisms at play in Experiment 3, see Figure 19 below.

Figure 19.

Visual conceptualization of the purpose, predictions, and potential mechanisms at play in Experiment 3



Note. In this experiment, learning goals and JOL prompts are always matched, while congruency between Learning Goals/JOL prompts and tests is manipulated.

Method

*Participant*s

Participants were recruited through Prolific and compensated with the standard Prolific rate. They were 18-35 years old, within the United States, and used a desktop to complete the experiment. This experiment implemented the same novel spatial stimuli and entirely novel paradigm as Experiment 2. Therefore, due to the exploratory nature of this work, we were unable to conduct a representative power analysis. As in Experiment 2, the desired sample size was determined by holistic consideration of previous work from the 1) verbal 2) spatial, and 3) multimedia learning literatures.

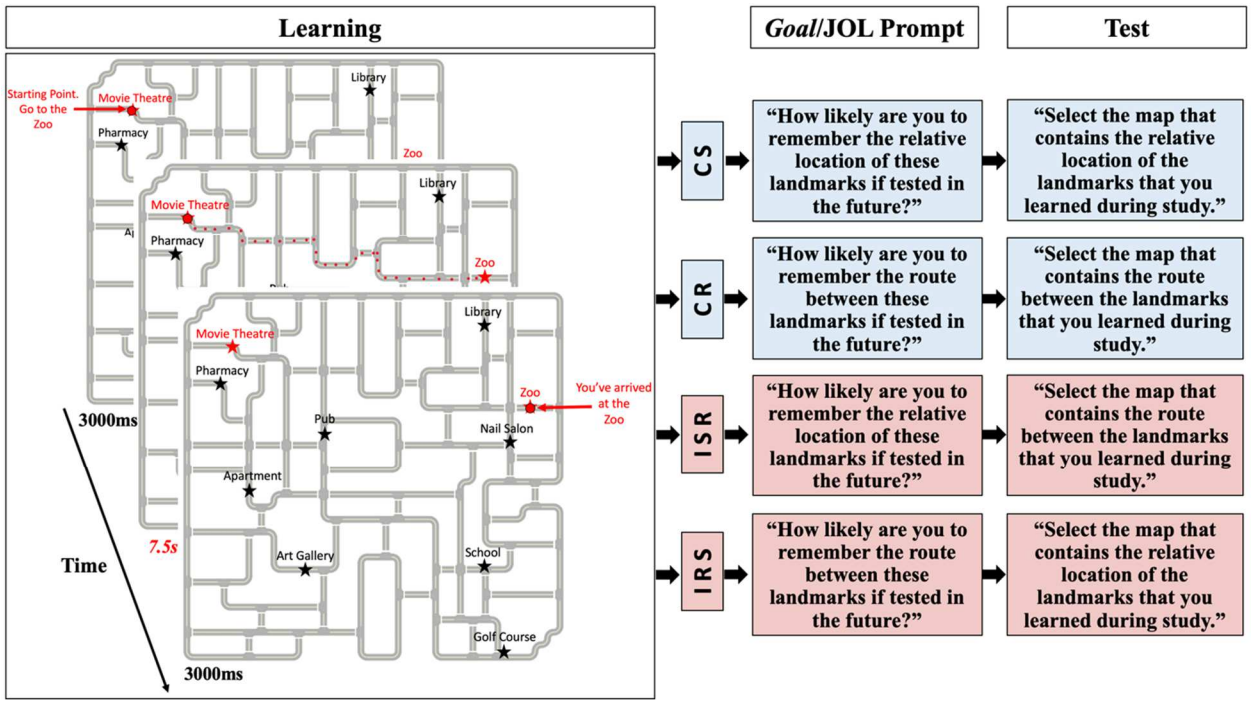
Design

A 2 (Assigned Goal/JOL Prompt: Survey, Route) x 2 (Test: Survey, Route) between-subjects design was employed. The assigned learning goal perspective always matched the JOL prompt perspective. If a participant was assigned a survey learning goal, they also received a survey JOL prompt after each learning trial. The same was true for route learning goals and JOL prompts. Then, they were tested on their survey or route memory. Two conditions contained congruent learning goal/JOL prompts and tests. In the congruent Survey Goal/JOL/Test condition, they had a survey learning goal/made JOLs about survey information and were tested on their survey memory. In the congruent Route Goal/JOL/Test condition, they had a route learning goal/made JOLs about route information and were tested on their route memory.

The remaining two conditions contained incongruent learning goal/JOL prompts and tests. The incongruent Survey Goal/JOL, Route Test condition had a survey learning goal/made JOLs about survey information, but they were tested on their route memory. The incongruent Route Goal/JOL, Survey Test condition had a route goal/made JOLs about route information, but they were tested on their survey memory. For a visual schematic of the separate conditions comprising the experiment, as indicated by learning goal/JOL prompt-test congruency, see Figure 20.

Figure 20

Schematic of conditions, as indicated by JOL prompt-test congruency, in Experiment 3



Note. CS= Congruent Survey Goal/JOL/Test; CR= Congruent Route Goal/JOL/Test; ISR= Incongruent Survey Goal/JOL, Route Test; IRS= Incongruent Route Goal/JOL, Survey Test

Landmark name pairs. The pairs were the same as those used in Experiment 2.

Context landmarks. These context landmarks were the same as those used in Experiment 2.

Maps. The learning and test map stimuli were the same as those used in Experiment 2.

Questionnaires. Participants completed the same demographics, encoding-based strategies, SBSOD, single item sense of direction, and displays questionnaire used in Experiment 2.

Attention checks. The attention checks were the same as in Experiment 2.

Retention interval. The retention interval was the same as in Experiment 2.

Map Learning Task. As in Experiment 2, the experiment was caried out over Gorilla, the map task was split into 3 phases, and participants completed two practice trials representative of their assigned condition before beginning the task. The stimuli presentation order was randomized during learning and test.

Learning Goals. Participants were provided with an assigned survey or route learning goal, just as in experiment 2. They are either instructed to learn the relative location of the landmarks or the route between landmarks. Next, they saw the same practice map used in Experiment 2.

JOLs. After each trial, they made a JOL. Participants in the congruent Survey Goal/JOL/Test condition or the incongruent Survey Goal/JOL, Route Test condition were prompted with the question, "How likely are you to remember the relative location of these landmarks if tested in the future?" Participants in the congruent Route Goal/JOL/Test condition or the incongruent Route Goal/JOL, Survey Test condition were prompted with the question, "How likely are you to remember the route between these landmarks if tested in the future?" The response scales ranged from 0 (not likely to remember) to 100 (highly likely to remember).

Tests. Then, the 2-alternative forced choice (2-AFC) practice test began. The practice test was the same as Experiment 2. In the congruent Survey Goal/JOL/Test condition and the incongruent Route Goal/JOL, Survey Test condition, one of the maps contained a target landmark that was moved from the location shown during study. Participants were told to "Select the map that contains the relative location of the landmarks that you learned during study." In the congruent Route Goal/JOL/Test condition and the incongruent Survey Goal/JOL, Route Test condition, one of the maps contained a route between the "starting point" and "destination" landmark that was different from the route shown during study. Participants were told to "select the map that contains the route between the landmarks that you learned during study." Just as in Experiment 2,

participants were provided with feedback on their selection only following the first practice test trial. After the second trial, they rated their confidence about their selection. Then, the learning phase began.

During the learning phase, participants studied the same map stimuli as in Experiment 2. At the start of each trial, they were reminded of their learning goal. A participant assigned to a survey learning goal/judgment condition saw, "Learn the relative location of the landmarks". A participant assigned to a route learning goal/judgment condition saw, "Learn the route between the landmarks". Then the map stimulus was presented.

Participants made a JOL after viewing each stimulus. The congruent Survey Goal/JOL/Test condition received a survey learning goal, made JOLs about survey information, and were tested on their survey memory. They were told to "Learn the relative location of the landmarks". When making a JOL, they were asked, "How likely are you to remember the relative location of these landmarks if tested in the future?" After study, they took a survey test. The congruent Route Goal/JOL/Test condition received a route learning goal, make JOLs about route information, and were tested on their route memory. They were told to "Learn the route between the landmarks". When making a JOL, they were asked, "How likely are you to remember the route between these landmarks if tested in the future?" After study, they took a route test.

The other half of participants were assigned to a condition with an incongruent learning goal/JOL prompt and test. Participants in the incongruent Survey Goal/JOL, Route Test condition were assigned a survey learning goal and made JOLs about survey information, but they were tested on their route memory. They were told to, "Learn the relative location of the landmarks," and when making a JOL, they were asked, "How likely are you to remember the relative location of these landmarks if tested in the future?" However, they were tested on their

memory for the route between the landmarks. Participants in the incongruent Route Goal/JOL, Survey Test condition were assigned a route learning goal and made JOLs about route information, but they were tested on their survey memory. They were told to "Learn the route between the landmarks," and when making a JOL, they were asked, "How likely are you to remember the route between these landmarks if tested in the future?" However, they were tested on their memory for the relative location of the landmarks.

All participants rated their likelihood to remember on a scale from 0 (not likely to remember) to 100 (highly likely to remember). After making the JOL, the next trial began. After studying the 16 map stimuli, the test phase began. Just as in experiment 2, the color attention check was placed at the end of the study phase. Then, they completed the two-minute retention interval. Next, the test phase began. The 2-AFC test was the same as Experiment 2. The 1920 attention check was placed after the test phase.

Questionnaires. After test, participants completed the same questionnaires as in Experiment 1 and 2.

Experiment 3 Results

In Experiment 3, we collected data from 106 participants. We excluded data from 10 participants due to technical difficulties and 2 participants for failed attention checks. We analyzed data from 94 participants ages 21-35 years. Most participants identified as white (73%) females (55%). Participants came from a wide array of educational backgrounds, with 57% holding a bachelor's degree or greater.

Memory and Metacognitive Accuracy Results

We estimated linear mixed models to assess our primary hypotheses. First, we tested whether better memory accuracy varied between conditions. The results of this model (E3M1,

Table 6) were non-significant, and it was not a better fit for the data as compared to the null model (E3MN, Table 6).

Table 6

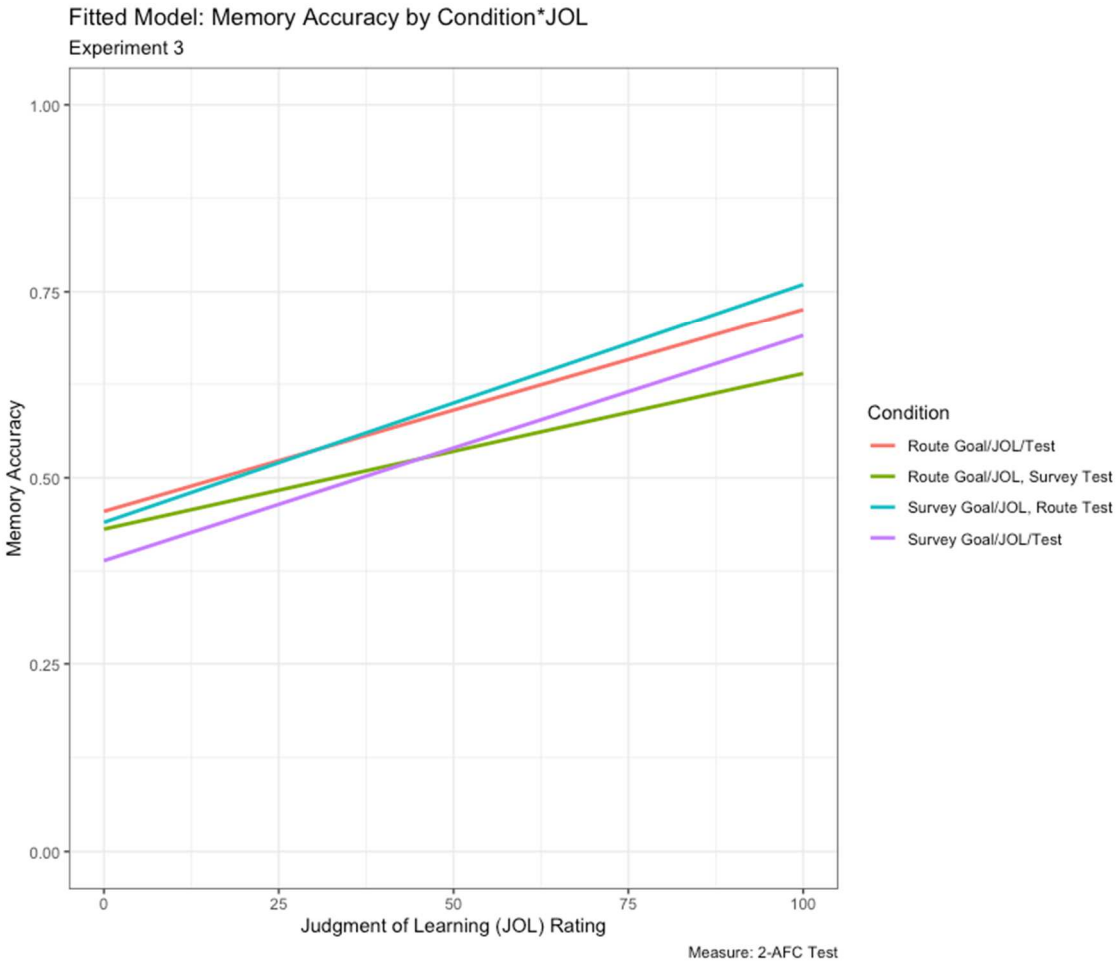
Model formulas and associated AIC values predicting memory accuracy

Model	Model formula	AIC
E3MN	Accuracy~ 1 + (1|PID)	1962.1
E3M1	Accuracy ~ Condition + (1|PID)	1965.8
E3M2	Accuracy ~ Condition*JOL + (1|PID)	1958.8

Next, we entered JOL into the model (E3M2, Table 6) to see whether rating predicted test accuracy, and whether there was a condition*JOL interaction. We saw improvement in the overall fit by adding JOL to the model. We observed a main effect of JOL ($p < .001$), meaning that JOL ratings did predict test accuracy (Fig 21). There were no significant interactions.

Figure 21

E3M2. Mean memory accuracy as a function of JOL and condition



Exploratory Memory and Metacognitive Accuracy Results

Just as in Experiment 2, we conducted exploratory analyses. First, we assessed whether learned complexity played a role in memory accuracy and if this role differed between groups. To address this question, we entered condition*learned complexity into a model (E2M3, Table 7).

Table 7

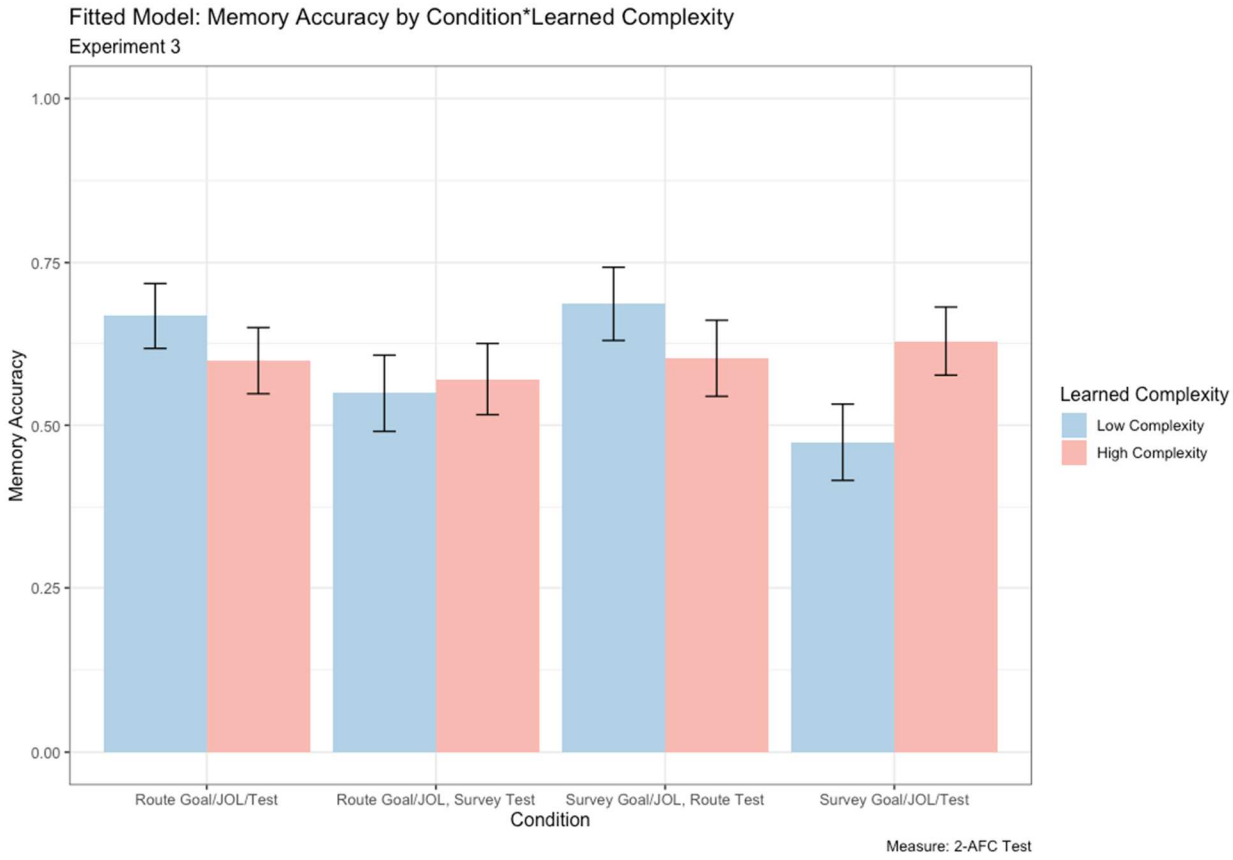
Exploratory model formulas and associated AIC values predicting memory accuracy

Model	Model formula	AIC
E3M3	Accuracy~ Condition*Learned Complexity + (1|PID)	1961.5
E3M4	Accuracy ~ Condition*JOL*Learned Complexity + (1|PID)	1959.8

We observed a main effect of learned complexity ($p = .005$) and significant condition*complexity interactions. Specifically, those taking route tests (Route Goal/JOL/Test and Survey Goal/JOL, Route Test) had higher accuracy on low complexity maps, independent of learning goals and JOLs ($p = .003$). In contrast, the congruent Survey Goal/JOL/Test group had higher accuracy on high complexity maps ($p = .003$). This effect was non-significant for the Route Goal/JOL, Survey Test group (Fig 22).

Figure 22

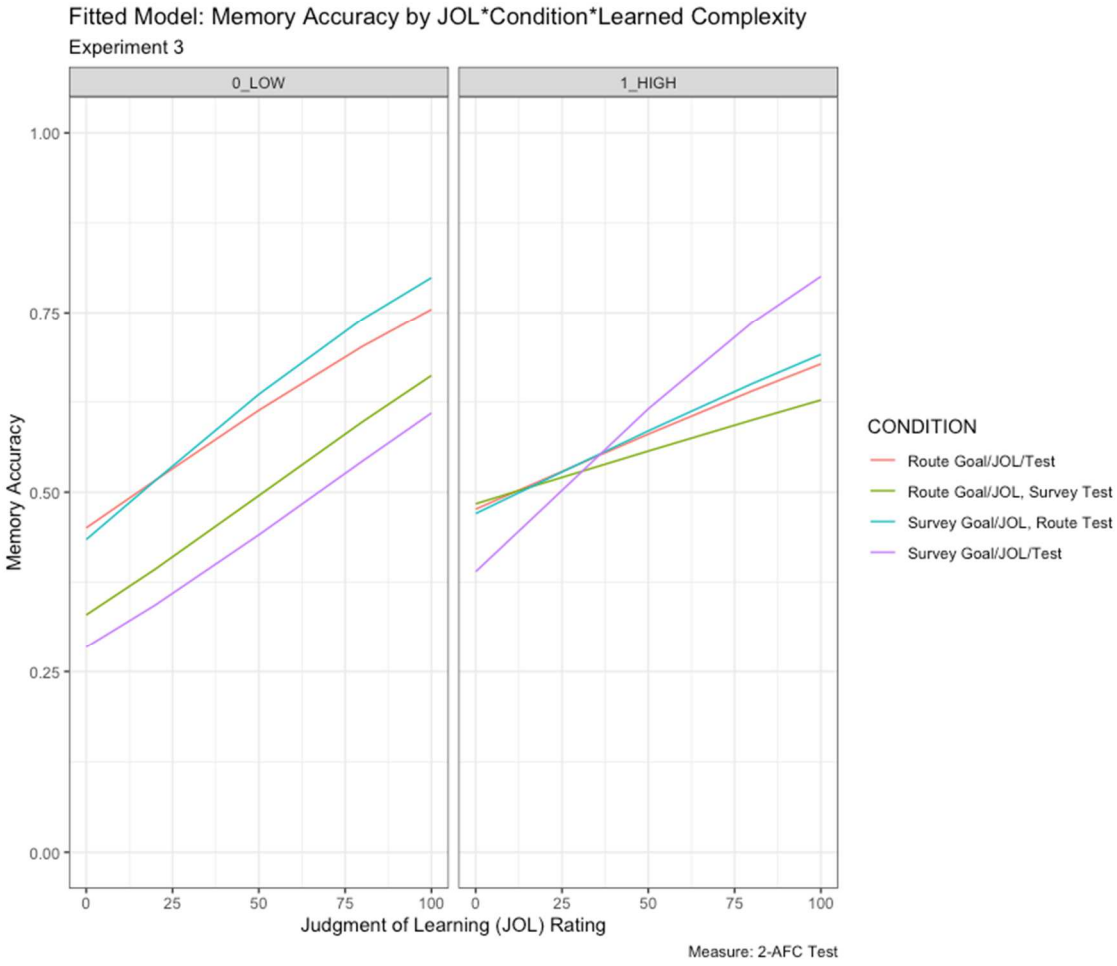
E3M3. Mean memory accuracy as a function of condition and learned complexity. Error bars show SE.



We tested whether learned complexity played a role in metacognitive accuracy and if this role differed between groups. We entered condition*JOL*learned complexity into the model (E3M4, Table 7). A generally positive relationship was observed between JOL ratings and memory accuracy across conditions. Though the model fit improved by adding JOL, the result of the three-way interaction was non-significant (Fig 23).

Figure 23

E3M4. Mean memory accuracy as a function of JOL across levels of learned complexity and conditions.



Exploratory JOL Formation Results

Next, we explored factors that might predict JOL formation. We began by testing whether condition alone predicts the JOL magnitude differentially across groups (E3M5, Table 8).

Table 8

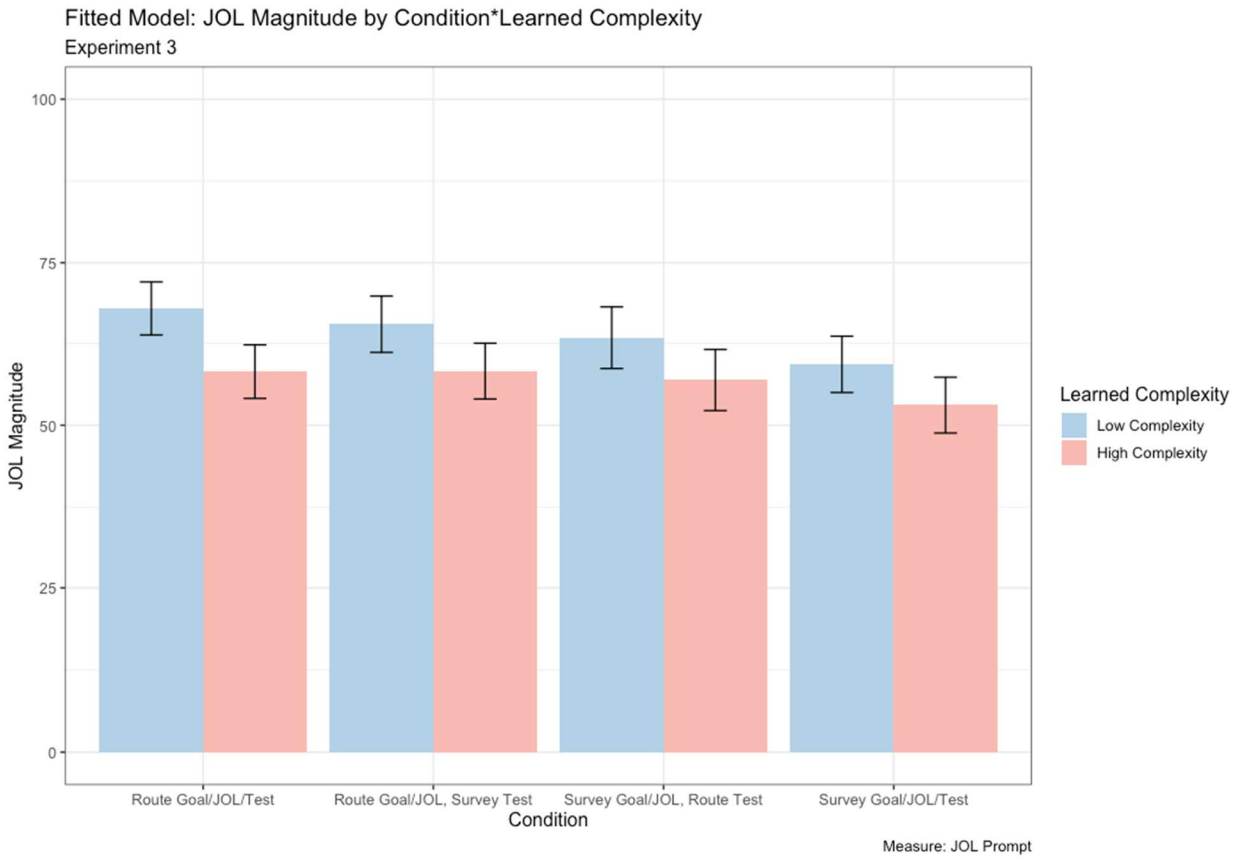
Exploratory model formulas and associated AIC values predicting JOL magnitude

Model	Model formula	AIC
E3MNJ	JOL ~ 1 + (1|PID)	12944.71
E3M5	JOL ~ Condition + (1|PID)	12933.64
E3M6	JOL ~ Condition*Learned Complexity + (1|PID)	12851.01
E3M7	JOL ~ Condition*Learning Trial + (1|PID)	12850.98

Though the fit of this model was significantly better than the null, it was non-significant, meaning the groups did not differ in their average rating magnitudes. Next, we entered learned complexity into the model (E3M6, Table 8). We observed a main effect of learned complexity ($p < .001$), such that JOL values for low complexity maps were significantly higher as compared to high complexity ones (Fig 24). The model fit improved by adding learned complexity.

Fig 24

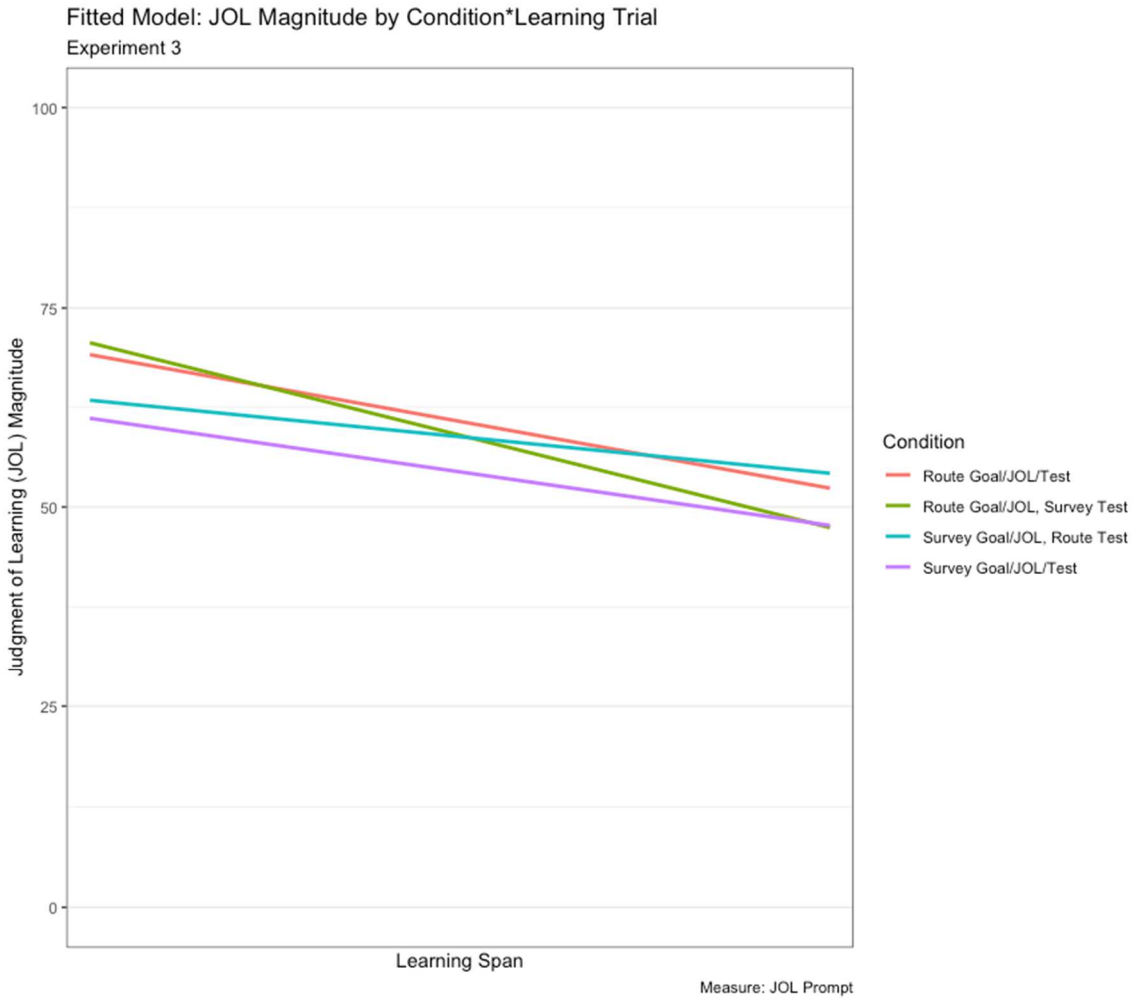
E3M6. Mean JOL magnitude as a function of condition and learned complexity. Error bars show SE.



Next, we tested whether JOLs differed as a function of time. We entered learning trial into the model (E3M7, Table 8). We observed a main effect of learning trial ($p < .001$), such that JOLs decreased across time for all groups, and a significant condition*learning trial interaction. Namely, JOL magnitudes decreased at a steeper rate for the incongruent Route Goal/JOL, Survey Test group as compared to the Survey Goal/JOL/Test group ($p < .05$) and the Survey Goal/JOL, Route Test group ($p < .005$) (Figure 25). The fit of this model was no better than the previous learned complexity model at predicting JOL magnitude, though both models were better than the model with condition alone.

Figure 25

E3M7. Mean JOL magnitude as a function of the learning span across conditions



Encoding Based Strategies

Next, we assessed participant responses on the subjective encoding-based strategies (EBS) questionnaire. Specifically, we conducted a mixed factor ANOVA to determine whether the amount of attention participants reportedly paid to relative landmark locations and routes between landmarks differed. We observed a condition*component interaction ($f(3) = 7.14$, $p = .0002$)) (Figure 26). The Route Goal/JOL/Test condition reported attending to routes significantly more

than relative landmark locations (f(1) = 13.6 p = .004)). The Route Goal/JOL/Test condition also reported remembering route information significantly more than relative landmark locations (f(1) = 8.3, p < .05)) (Fig 27). None of the other conditions were significant.

Figure 26

Mean attention rating as a function of condition and map component. Error bars show SE.

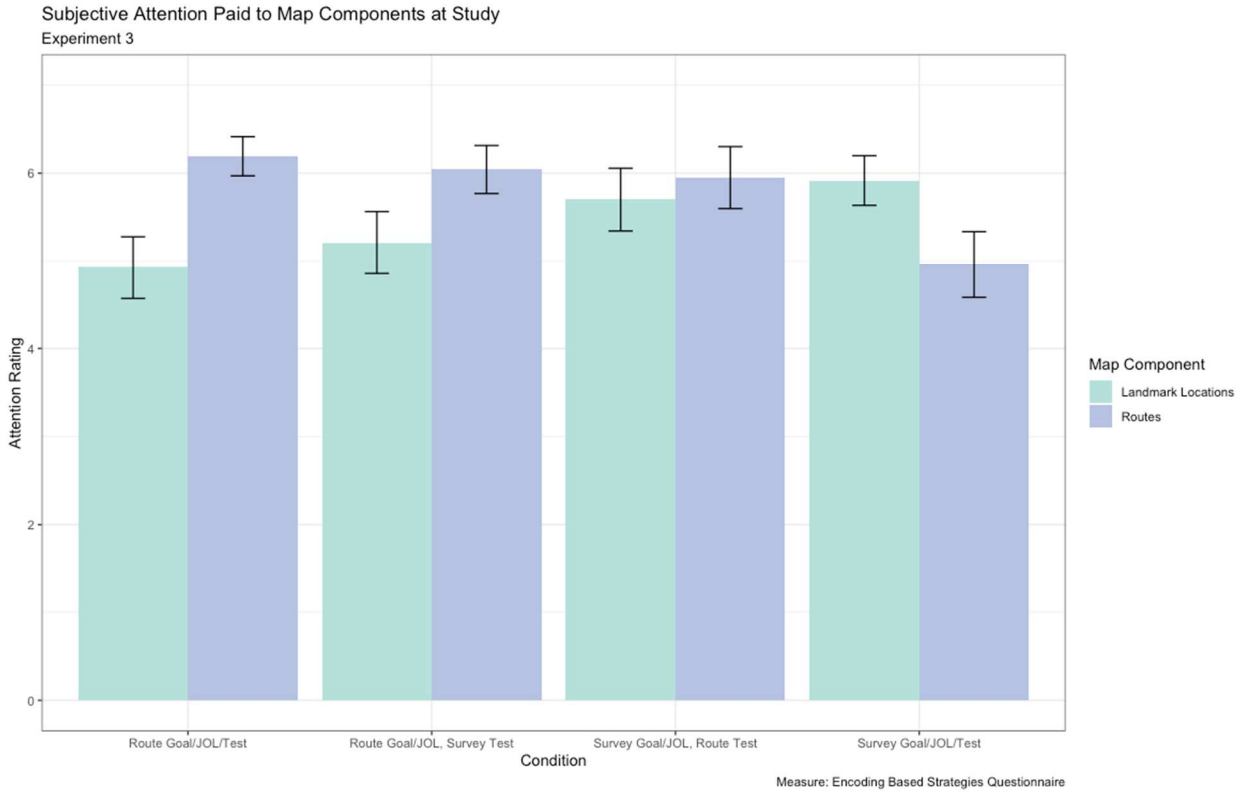
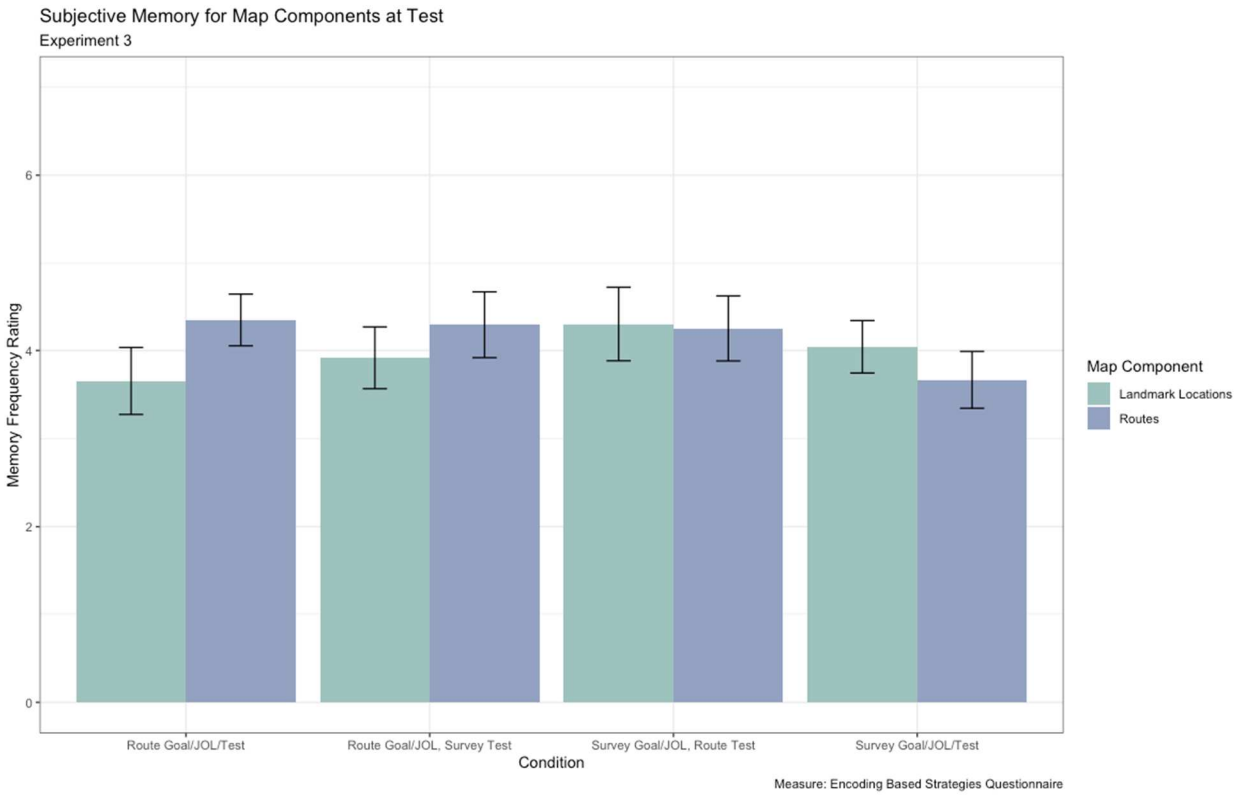


Figure 27

Mean memory rating as a function of condition and map component. Error bars show SE.



There were no significant differences between the groups on the single-item sense of direction (SOD) measure. Furthermore, the average single item SOD score was nearly equal across experiments 1 ($\mu = 4.6$), 2 and 3 ($\mu = 4.8$).

Experiment 3 Discussion

In Experiment 3, we investigated the relationship between JOL prompts and tests. Incongruency between JOL prompts and tests did not play a role in memory accuracy or metacognitive accuracy. A positive relationship between JOL ratings and memory accuracy was observed across conditions. Said differently, higher JOL ratings were associated with higher memory accuracy. This finding suggests that during map learning, even if the learning goal and

monitoring judgments are not predictive of the test conditions, learners are still able to accurately gauge their likelihood to remember the map components at test. This result is unexpected, since classic TAM results would show poor predictive accuracy among those who had a goal to learn survey information, made JOLs about their likelihood to remember survey information, but were unexpectedly tested on route knowledge. Relatively accurate monitoring judgments in incongruent conditions might suggest that during map learning, learners are forming their JOLs based on a global representation of the map, even when the JOL prompt points them to specific elements. Though some learners reported paying significantly more attention to the information that aligned with their goals, our findings suggest that the cues they used to form their judgments were reliable indicators of their future memory, independent of the specific map components they were tested on.

Exploratory Discoveries

We estimated exploratory models to gain insight on the role of map complexity on memory and metacognitive accuracy, and whether this relationship varies based on JOL-test congruency. Exploratory findings did suggest that complexity predicts memory accuracy. Differences emerged for low complexity maps. Participants taking route tests (Route Goal/JOL/Test and Survey Goal/JOL, Route Test) were more accurate on low complexity maps, irrespective of their learning goal/JOLs. We replicated exploratory findings from Experiment 2, suggesting that both map complexity and time, as indicated by learning trial, are significant predictors of JOL magnitude. Higher JOLs were observed for less complex maps and JOLs decreased across the learning span.

General Discussion

Across three experiments, we explored the role of JOLs during map learning. Experiment 1 found that JOL-reactivity does occur even for complex map-based stimuli. Participants who

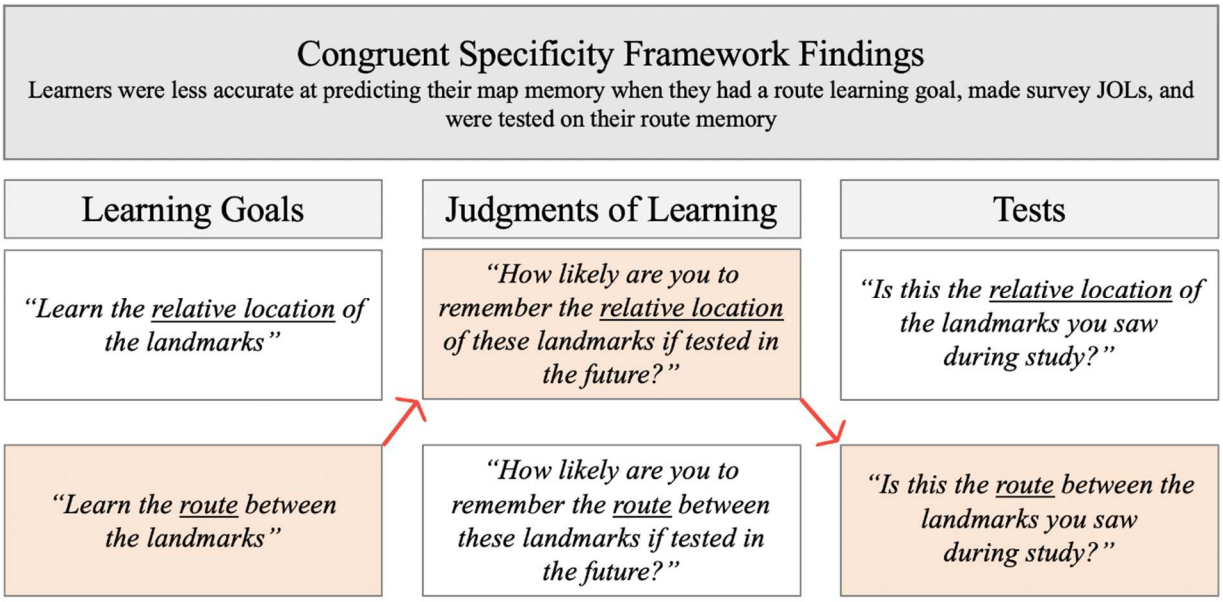
made JOLs better remembered landmark names across both map and word contexts, providing support for our novel bridge hypothesis. To our knowledge, we are the first to show evidence for JOL-reactivity during map-based learning. Both JOL and RNG groups had greater memory accuracy when there was a match between study and test contexts, such that pairs were studied and tested without maps, suggesting a TAP finding. However, contrary to traditional TAM findings, a positive relationship between monitoring judgments and memory accuracy was observed even when there was a mismatch between study and test contexts. This finding suggests that existing theories about the relationship between metacognitive judgments and tests may require reconsideration in spatial learning research.

With support for the utility of monitoring map-based learning, we turned our attention to investigate the relationship between learning goals and JOL prompts (Experiment 2), and JOL prompts and tests (Experiment 3). Experiment 2 provides novel insights on the relationship between learning goals and JOLs during map learning. Directing learners to learn one map component but to monitor their learning of a different map component resulted in poorer metacognitive accuracy compared to those who were told to learn and monitor the same map components. Critically, order mattered. While disadvantages in metacognitive accuracy were observed for those with route goals but survey JOL prompts, the same was not true for those with survey goals and route JOL prompts. It is possible that directing attention to relative landmark locations first and then focusing on routes between landmarks aligns more closely with how people naturally process map information during self-directed learning. Prior research found that learning from a survey perspective resulted in flexible representations that could be abstracted beyond the original perspective, while representations formed from route descriptions were tied to the initially experienced orientation (Brunyé et al., 2008) and were more time-consuming in nature (Brunyé &

Taylor, 2008). Our results expand on these findings to suggest that directing learners to learn route information but monitor their likelihood to remember survey components may result in ineffective metacognitive processes. This discovery supports one aspect of the proposed congruent specificity framework and offers insights on how map learning instruction may be tailored to benefit predictive accuracy among learners (Figure 28).

Figure 28

Congruent Specificity Framework findings from Experiment 2



Contrary to our hypotheses, we did not observe classic TAM effects in Experiment 3. Namely, the groups did not differ in their metacognitive accuracy. It is possible that when goals and JOLs are congruent, people learn the map in a global manner. This results in accurate predictions of future memory, irrespective of which specific map components they are tested on. These findings align with those from Experiment 1, suggesting that people may use multiple

diagnostic cues when forming JOLs during map learning. This results in relatively accurate monitoring judgments, even when the test only calls upon memory for a particular component of the learned information.

Our results from all three experiments suggest that JOLs are a significant predictor of memory accuracy. This is a striking finding as predictive accuracy tends to be poor across the verbal learning literature (Dunlosky & Lipko, 2007; Glenberg et al., 1987; Miller & Geraci, 2014). To our knowledge, we are the first to show that people are relatively accurate at predicting their memory of complex spatial map components. Furthermore, JOL rating magnitude was impacted by complexity of learned materials and time in experiments 2 and 3. Namely, people rated themselves less likely to remember more complex materials and less likely to remember materials generally as the learning span progressed. Additional work is needed to understand how JOLs function to aid learning and memory outcomes and what components influence predictive accuracy in spatial scenarios.

Conclusion

Across three experiments, we examined the role of metacognitive monitoring during map-based environment learning (EL). Our research suggests that JOLs are generally important for learning outcomes, independent of how information is learned or how visually complex the materials are. In other words, traditional JOL-reactivity findings translate to visuospatial contexts. This discovery serves as a necessary, early step in understanding metacognition's role in EL. Not all observations from the metacognitive literature with verbal materials are supported by our discoveries. Throughout multiple experiments, we observed results contrary to what would be predicted by the TAM framework. Namely, participants were relatively accurate at predicting their future memory, independent of the test context or components. Based on prior research with verbal

materials, TAM predicts better monitoring accuracy when the cognitive processes engendered during monitoring are congruent with those required by the test. However, the TAM framework was developed using relatively simple materials which may require less sophisticated evaluation of cues used to monitor learning. In contrast, monitoring EL is likely a more complex process, requiring the evaluation and weighting of more information, including landmark locations, routes between landmarks, and landmark names. Additionally, environment complexity and subjective cognitive load burden can serve as cues during learning. It is difficult to generalize predictions from the TAM framework to complex visuospatial learning contexts, where several information streams must be collectively considered during metacognitive judgment formation.

Our exploratory findings suggest that learners may globally monitor several diagnostic cues during learning, including multiple map components, map complexity, and even information accumulation across the learning span. Using several cues to generate a monitoring judgment has implications for understanding prediction accuracy and evaluating cue efficacy. Specifically, people are relatively accurate at predicting their final test performance, irrespective of the specific map components they are tested on. In sum, EL research may require a more dynamic metacognitive model that accounts for the diverse, interactive cues at play during complex learning. As described in our forthcoming review (Mason et al., 2022), several cue types can contribute to metacognitive judgment formation during EL, including external environment cues (such as routes and landmark locations), internal spatial cues (such as the cue and context landmarks in reference to the 'you've arrived' target indicator in experiments 2 and 3) and mnemonic cues (such as the subjective sense of increased cognitive load across time). Globally monitoring these interactive cues seems necessary for accurate predictions of future environment memory. Future metacognitive research with spatial materials should consider various cue types

and functions to help elucidate metacognitive judgment formation processes during complex learning.

References

Aretz, A. J. (1988). *A model of electronic map interpretation*. *32*, 130–134. https://doi.org/10.1177/154193128803200228

Baddeley, A. (1986). *Working memory*. Oxford University Press.

Blaxton, T. A. (1989). Investigating Dissociations Among Memory Measures: Support for a Transfer-Appropriate Processing Framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 657–668.

Bonura, B. M. (2011). *Age Differences in Remembering "What" and "Where": A Comparison of Spatial Working Memory and Metacognition in Older and Younger Adults*.

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*. 11.

Brunyé, T. T., Rapp, D. N., & Taylor, H. A. (2008). Representational flexibility and specificity following spatial descriptions of real-world environments. *Cognition*, *108*(2), 418–443. https://doi.org/10.1016/j.cognition.2008.03.005

Brunyé, T. T., & Taylor, H. A. (2008). Extended experience benefits spatial mental model development with route but not survey descriptions. *Acta Psychologica*, *127*(2), 340–354. https://doi.org/10.1016/j.actpsy.2007.07.002

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

Burkett, C., & Azevedo, R. (2012). The effect of multimedia discrepancies on metacognitive judgments. *Computers in Human Behavior*, *28*(4), 1276–1285. https://doi.org/10.1016/j.chb.2012.02.011

Coluccia, E. (2008). Learning from maps: The role of visuo-spatial working memory. *Applied Cognitive Psychology*, *22*(2), 217–233. https://doi.org/10.1002/acp.1357

Cuevas, H. M., Fiore, S. M., Bowers, C. A., & Salas, E. (2004). Fostering constructive cognitive and metacognitive activity in computer-based complex task training environments. *Computers in Human Behavior*, *20*(2), 225–241. https://doi.org/10.1016/j.chb.2003.10.016

Cuevas, H. M., Fiore, S., & Oser, R. (2002). Scaffolding cognitive and metacognitive processes in low verbal ability learners: Use of diagrams in computer-based training environments. *Instructional Science*, *30*. https://doi.org/10.1023/A:1020516301541

Dai, R. (2016). *The "where" and "what": Aging and visuo-spatial working memory*. ProQuest Dissertations Publishing.

Double, K. S., Birney, D. P., & Walker, S. A. (2018). A meta-analysis and systematic review of reactivity to judgements of learning. *Memory*, *26*(6), 741–750. https://doi.org/10.1080/09658211.2017.1404111

Dougherty, M. R., Scheck, P., & Nelson, T. O. (2005). Using the past to predict the future. *Memory & Cognition*, *33*(6), 1096–1115.

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A Brief History and How to Improve Its Accuracy. *Current Directions in Psychological Science*, *16*(4), 228–232. https://doi.org/10.1111/j.1467-8721.2007.00509.x

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

Flavell, J. H. (1979). Metacognition and Cognitive Monitoring: A New Area of Cognitive-Developmental Inquiry. *American Psychologist*, *34*(10), 906–911.

Ford, J. K., Smith, E. M., Consulting, A., & Weissbein, D. A. (1998). *Relationships of Goal Orientation, Metacognitive Activity, and Practice Strategies With Learning Outcomes and Transfer*. 16.

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

Glenberg, A. M., Sanocki, T., Epstein, W., & Morris, C. (1987). *Enhancing Calibration of Comprehension*. 18.

Hegarty, M. (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

Janes, J. L., Rivers, M. L., & Dunlosky, J. (2018). The influence of making judgments of learning on memory performance: Positive, negative, or both? *Psychonomic Bulletin & Review*, *25*(6), 2356–2364. https://doi.org/10.3758/s13423-018-1463-4

Kintsch, W. (1998). The Role of Knowledge in Discourse Comprehension: A Construction-Integration Model. *Psychological Review*, *95*(2), 163–182.

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

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

Kulhavy, R. W., Lee, J. B., & Caterino, L. C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology*, *10*(1), 28–37. https://doi.org/10.1016/0361-476X(85)90003-7

LaBerge, D. (1995). *Attentional Processing: The brain's art of mindfulness*. Harvard University Press.

Mayer, R. E. (2001). *Multi-media Learning*. Cambridge University Press.

Miller, T. M., & Geraci, L. (2014). Improving metacognitive accuracy: How failing to retrieve practice items reduces overconfidence. *Consciousness and Cognition*, *29*, 131–140. https://doi.org/10.1016/j.concog.2014.08.008

Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology: General*, *145*(2), 200–219. https://doi.org/10.1037/a0039923

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519–533. https://doi.org/10.1016/S0022-5371(77)80016-9

Muffato, V., Meneghetti, C., & De Beni, R. (2019). Spatial mental representations: The influence of age on route learning from maps and navigation. *Psychological Research*, *83*(8), 1836–1850. https://doi.org/10.1007/s00426-018-1033-4

Nelson, T., & Narens, L. (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 *Psychology of Learning and Motivation* (Vol. 26, pp. 125–173). Elsevier. https://doi.org/10.1016/S0079-7421(08)60053-5

Osman, M. E., & Hannafin, M. J. (1992). Metacognition research and theory: Analysis and implications for instructional design. *Educational Technology Research and Development*, *40*(2), 83–99. https://doi.org/10.1007/BF02297053

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture. *Instructional Science*, *32*(1/2), 1–8. https://doi.org/10.1023/B:TRUC.0000021806.17516.d0

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press; Oxford Oxfordshire; Clarendon Press.

Pichert, J. W., & Anderson, R. C. (1977). *Taking Different Perspectives on a Story*. 7.

Schraw, G. (2001). Promoting General Metacognitive Awareness. In H. J. Hartman (Ed.), *Metacognition in Learning and Instruction* (Vol. 19, pp. 3–16). Springer Netherlands. https://doi.org/10.1007/978-94-017-2243-8_1

Schwartz, N. H., Andersen, C., Hong, N., Howard, B., & McGee, S. (2004). The Influence of Metacognitive Skills on Learners' Memory of Information in a Hypermedia

Environment. *Journal of Educational Computing Research*, *31*(1), 77–93. https://doi.org/10.2190/JE7W-VL6W-RNYF-RD4M

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

Snow, R. E., & Lohman, D. F. (1984). Toward a Theory of Cognitive Aptitude for Learning From Instruction '. *Journal of Educational Psychology*, *76*(3), 347–376.

Starkweather, D. J. (2010). *Linear Mixed Effects Modeling using R.* 13.

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

Tuovinen, J. E., & Sweller, J. (1991). A Comparison of Cognitive Load Associated With Discovery Learning and Worked Examples. *Journal of Educational Psychology*, *91*(2), 334–341.

Valcke, M. (2002). Cognitive load: Updating the theory?ଌ. *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

Weaver, C. A., & Kelemen, W. L. (2003). Processing similarity does not improve metamemory: Evidence against transfer-appropriate monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1058–1065. https://doi.org/10.1037/0278-7393.29.6.1058

Westerbeek, H., & Maes, A. (2013). Route-external and Route-internal Landmarks in Route Descriptions: Effects of Route Length and Map Design: Route-external and route-internal landmarks in route descriptions. *Applied Cognitive Psychology*, *27*(3), 297–305. https://doi.org/10.1002/acp.2907

Appendix A.

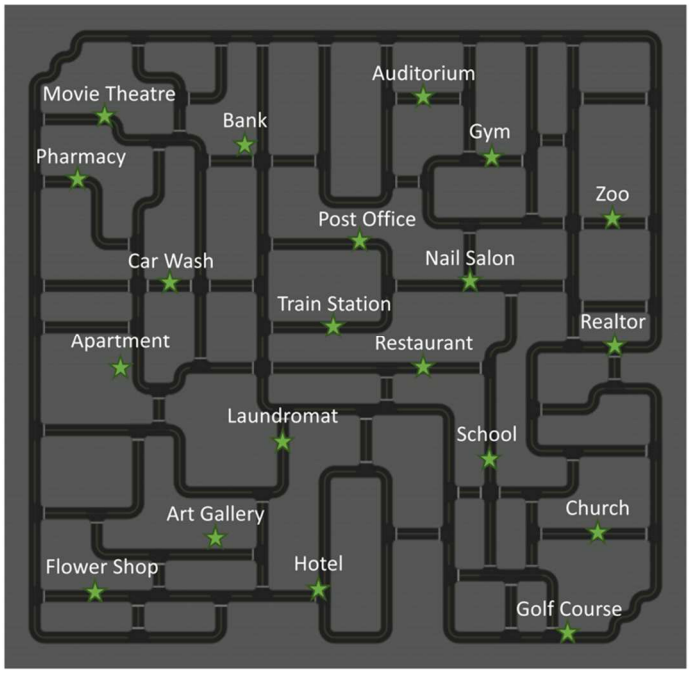
Landmark name pairings from list A and list B: Experiment 1

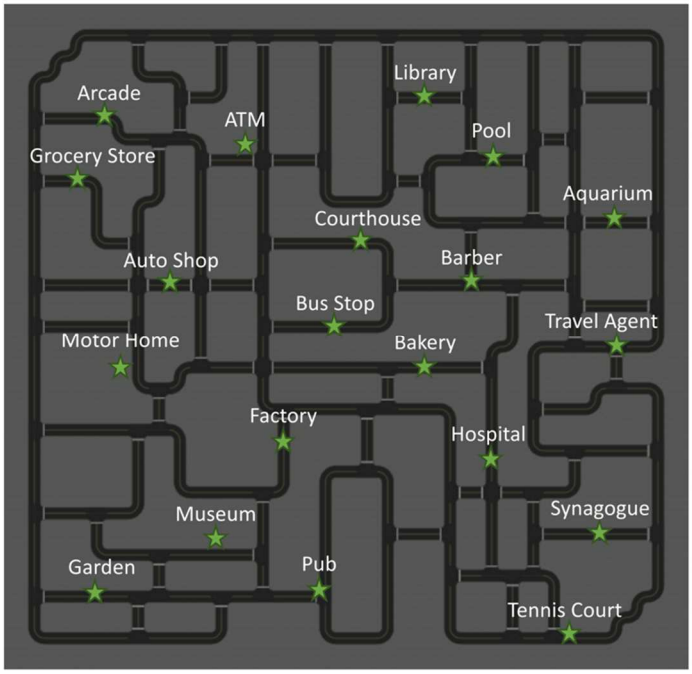
Pairs (List A)	
Movie Theatre	Zoo
Bank	Auditorium
Pharmacy	Art Gallery
Car Wash	Post Office
Train Station	Realtor
Gym	Hotel
Nail Salon	School
Restaurant	Laundromat
Golf Course	Apartment
Church	Flower Shop

Pairs (List B)	
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.

Complete Map A (Top), Map B (Bottom): Experiment 1





Appendix C.

Encoding-based strategies questionnaire: Experiment 1

Encoding-Based Strategies Questionnaire (Experiment 1)

For each of the following questions, please think about the landmark-pairs that you studied and were tested on during the task. When prompted, feel free to explain your answer in the space labeled "Explain." Press the arrow to begin.

While **viewing the full map** for 30 seconds, how closely did you pay attention to where the landmark names were placed on the map?

No Attention 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Close Attention 7 (7)

If you would like to explain your response and/or provide an example, please do so using the space below:

While **studying landmark pairs** on the maps, how closely did you pay attention to where the landmark names were placed on the map?

No Attention 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Close Attention 7 (7)

If you would like to explain your response and/or provide an example, please do so using the space below:

When tested on landmark pairs, how frequently did you remember whether the landmark names had been displayed on a map or not?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If you would like to explain your response and/or provide an example, please do so using the space below:

When tested on landmark pairs, how frequently did you think back to where the associated landmark was located on the map?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If you would like to explain your response and/or provide an example, please do so using the space below:

Did thinking about where the landmark was located on the map help you to remember the associated landmark name on the test?

Did not help 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Helped a lot 7 (7)

If you would like to explain your response and/or provide an example, please do so using the space below:

When you were studying the landmark pairs, how frequently did you try to come up with a trick to remember the associated landmark names?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If so, what trick(s) did you use to remember the landmark names?

Did using a trick help you to remember the associated landmark names on the test?

Did not help 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Helped a lot 7 (7)

If you would like to explain your response and/or provide an example, please do so using the space below:

When you were studying the landmark pairs, how frequently did you try to link the two landmark names?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If so, how did you try to link the landmark names?

Did trying to link the two landmark names help you to remember the associated landmark name on the test?

Did not help 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Helped a lot 7 (7)

If you would like to explain your response and/or provide an example, please do so using the space below:

Rate the overall difficulty of remembering the associated landmark names at test.

Not Difficult 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Difficult 7 (7)

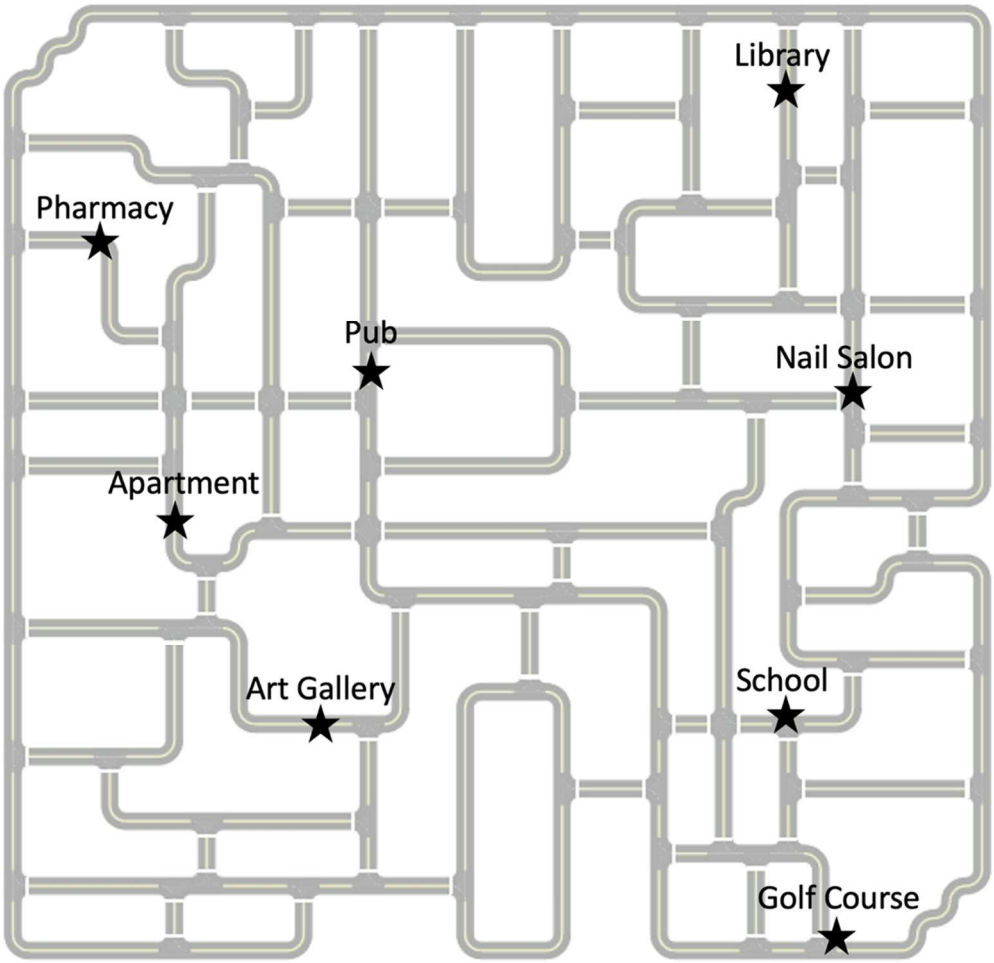
Appendix D.

Landmark name pairings: Experiment 2 & 3

Pairs	
Movie Theatre	Zoo
Bank	Auditorium
Car Wash	Post Office
Train Station	Realtor
Gym	Hotel
Restaurant	Laundromat
Church	Flower Shop
Arcade	Bakery
Grocery Store	Tennis Court
ATM	Synagogue
Auto Shop	Pool
Bus Stop	Garden
Barber	Hospital
Courthouse	Factory
Travel Agent	Museum
Motor Home	Aquarium

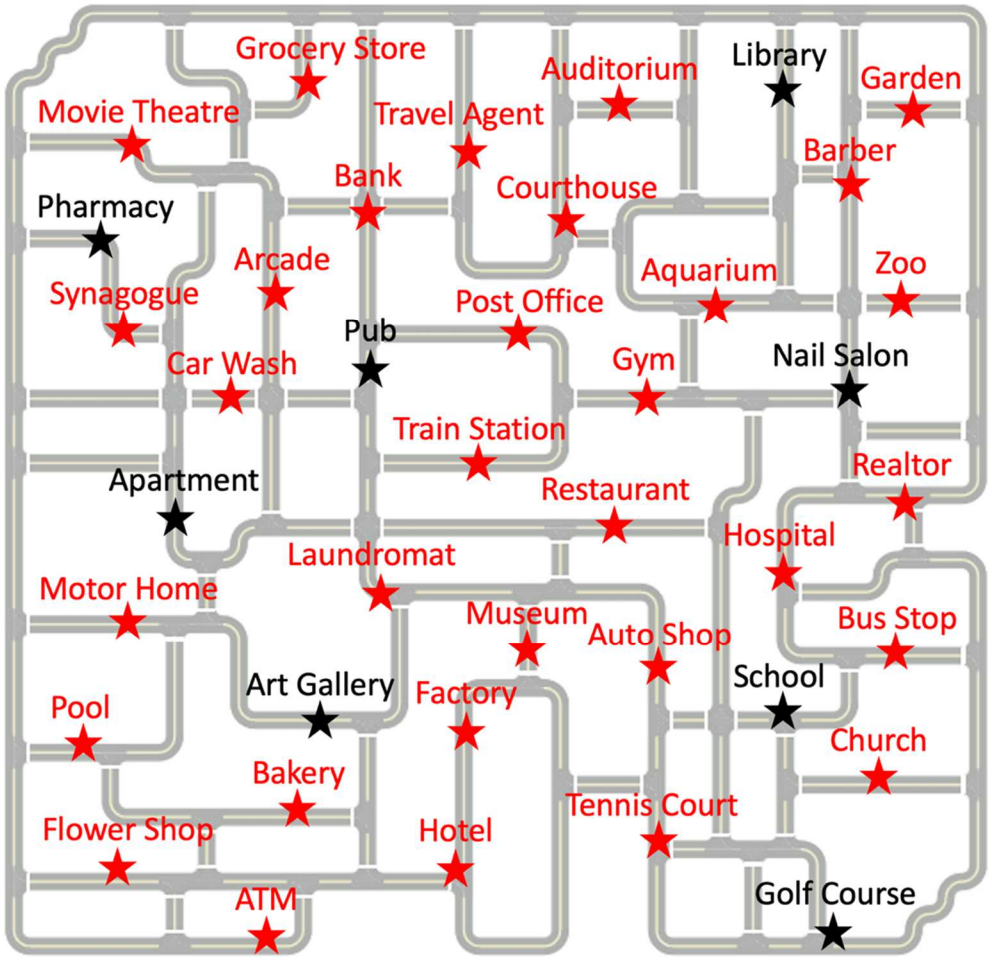
Appendix E.

Context landmarks: Experiment 2 &3



Appendix F.

Map containing all traveled landmarks (red) and context landmarks (black): Experiment 2 & 3



Appendix G.

Encoding-based strategies questionnaire: Experiment 2 & 3

Encoding-Based Strategies Questionnaire (Experiment 2 & 3)

For each of the following questions, please think about the maps that you studied and were tested on during the task. When prompted, feel free to explain your answer in the space labeled "Explain." Press the arrow to begin.

Study

While **studying the maps**, how closely did you pay attention to **where the landmarks were located** on the map?

No Attention 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Close Attention 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

While **studying the maps,** how closely did you pay attention to the **routes between landmarks** on the map?

No Attention 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Close Attention 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

Test

When **tested on your memory for map components,** how frequently did you remember **where the landmarks were located** on the map?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

When **tested on your memory for map components,** how frequently did you remember **the routes between landmarks** on the map?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

Did thinking about **where the landmarks were located** on the map help you to remember the information needed at test?

Did not help 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Helped a lot 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

Did thinking about the **routes between landmarks** on the map help you to remember the information needed at test?

Did not help 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Helped a lot 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

Open Strategies

When you were studying the maps, how frequently did you try to come up with a trick to remember **where the landmarks were located**?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If so, what trick(s) did you use to remember where the landmarks were located?

Did using a trick help you to remember the information needed at test?

Did not help 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Helped a lot 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

When you were studying the maps, how frequently did you try to come up with a trick to remember **the routes between landmarks**?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If so, what trick(s) did you use to remember the routes between landmarks?

Did using a trick help you to remember the information needed at test?

Did not help 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Helped a lot 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

When you were studying the maps, how frequently did you try to link the two landmark names?

Never 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Frequently 7 (7)

If so, how did you try to link the landmark names?

Did trying to link the two landmark names help you to remember the information needed at test?

Did not help 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Helped a lot 7 (7)

If you are willing to explain your response and/or provide an example, please do so using the space below:

Rate the overall difficulty of remembering the information needed at test.

Not Difficult 1 (1)
2 (2)
3 (3)
4 (4)
5 (5)
6 (6)
Very Difficult 7 (7)

Appendix H.

The progression of JOL ratings (left) and confidence ratings (right) as a function of the learning span across conditions in experiment 2 (top) and 3 (bottom)