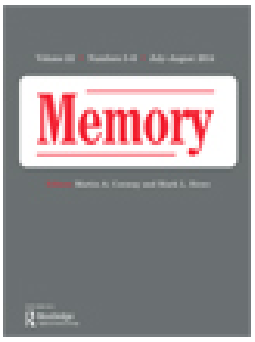


Memory

ISSN: 0965-8211 (Print) 1464-0686 (Online) Journal homepage: http://www.tandfonline.com/loi/pmem20

Verbal cues flexibly transform spatial representations in human memory

Candace E. Peacock & Arne D. Ekstrom

To cite this article: Candace E. Peacock & Arne D. Ekstrom (2018): Verbal cues flexibly transform spatial representations in human memory, Memory, DOI: 10.1080/09658211.2018.1520890

To link to this article: https://doi.org/10.1080/09658211.2018.1520890

MEMORY
https://doi.org/10.1080/09658211.2018.1520890

Routledge
Taylor & Francis Group

Verbal cues flexibly transform spatial representations in human memory

Candace E. Peacock[a,b] and Arne D. Ekstrom[a,b,c]

[a]Center for Neuroscience, University of California, Davis, CA, USA; [b]Department of Psychology, University of California, Davis, CA, USA; [c]Department of Psychology, University of Arizona, Tucson, AZ, USA

ABSTRACT
Humans possess a unique ability to communicate spatially-relevant information, yet the intersection between language and navigation remains largely unexplored. One possibility is that verbal cues accentuate heuristics useful for coding spatial layouts, yet this idea remains largely untested. We test the idea that verbal cues flexibly accentuate the coding of heuristics to remember spatial layouts via spatial boundaries or landmarks. The alternative hypothesis instead conceives of encoding during navigation as a step-wise process involving binding lower-level features, and thus subsequently formed spatial representations should not be modified by verbal cues. Across three experiments, we found that verbal cues significantly affected pointing error patterns at axes that were aligned with the verbally cued heuristic, suggesting that verbal cues influenced the heuristics employed to remember object positions. Further analyses suggested evidence for a hybrid model, in which boundaries were encoded more obligatorily than landmarks, but both were accessed flexibly with verbal instruction. These findings could not be accounted for by a tendency to spend more time facing the instructed component during navigation, ruling out an attentional-encoding mechanism. Our findings argue that verbal cues influence the heuristics employed to code environments, suggesting a mechanism for how humans use language to communicate navigationally-relevant information.

ARTICLE HISTORY
Received 18 April 2018
Accepted 31 August 2018

KEYWORDS
Memory; spatial navigation; spatial cognition; geometry; features; verbal cues

If you remember a time that you were new to a city, someone may have told you a useful means of finding your way, for example, that the streets have a grid-like structure or that there are distinctive mountains to the north (e.g., Brunye et al., 2012). While past studies demonstrate that text-based descriptions lead to similar representations as map-learning (Lee & Tversky, 2001; Taylor & Tversky, 1992), it remains unclear exactly how verbal cues influence and become integrated with spatial representations. Thus, an important and unresolved question regards how and when in spatial processing verbal cues affect memory for object positions in an environment.

Models of spatial navigation often conceptualise it as a process of binding lower-level features, like landmarks and routes, into higher-order representations that provide integrated information about object locations (Garling, Book, & Lindberg, 1984; Lynch, 1960; Siegel & White, 1975). While some have challenged these ideas (Ishikawa & Montello, 2006; Montello, 1998; Zhang, Zherdeva, & Ekstrom, 2014), arguing that many aspects of spatial processing occur in parallel, neurophysiological models of rodent navigation have typically relied on similar assumptions. Specifically, place cells and grid cells, thought to provide the neural basis for representations of position and metrics of space, respectively (Fyhn, Molden, Witter, Moser, & Moser, 2004; O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1978), assume that lower-level sensory features, like distal landmark cues, anchor these spatial representations (Moser, Kropff, & Moser, 2008; O'Keefe, Burgess, Donnett, Jeffery, & Maguire, 1998; Taube, 2007). While the importance of verbal cues is acknowledged in the attention literature (termed "top-down cues" or "task-goals") (Corbetta, Patel, & Shulman, 2008; Posner, 1980), how such cues modulate representations during navigation remains unknown. One possibility, based on the attention literature (e.g., Awh, Belopolsky, & Theeuwes, 2012), is that both salient low-level aspects (landmarks and boundaries) and task-related components (verbal cues) interact to underlie representations formed during navigation.

Cognitive and neurophysiological studies suggest that landmarks and boundaries affect spatial representations (Barry et al., 2006; Cheng, 1986; McNamara, Rump, & Werner, 2003; Morris, Hagan, & Rawlins, 1986; O'Keefe et al., 1998; O'Keefe & Nadel, 1978), and in this way, serve as powerful heuristics for organising spatial knowledge (Ekstrom & Isham, 2017; Tversky, 1981). Past studies suggest that participants are more accurate at pointing to object locations when their imagined facing angle is

CONTACT Arne D. Ekstrom adekstrom@email.arizona.edu Center for Neuroscience, University of California, Davis, CA, USA; Department of Psychology, University of California, Davis, CA 95616, USA; Department of Psychology, University of Arizona, Tucson, AZ 85721, USA

Supplemental data for this article can be accessed at https://doi.org/10.1080/09658211.2018.1520890.

aligned with environmental boundaries versus misaligned (Mou & McNamara, 2002), although both landmarks and boundaries can influence "alignment effects" (McNamara et al., 2003). For example, past work shows that in the absence of landmarks, spatial boundaries alone serve as powerful cues to extract spatial information (Mou & McNamara, 2002; Shelton & McNamara, 1997, 2001). Conversely, in the absence of boundaries, landmarks provide salient axes in which participants organise spatial knowledge (Chan, Baumann, Bellgrove, & Mattingly, 2013). The geometry versus feature literature has found that both boundaries and landmarks are used for reorientation and that this reorientation process is dependent on the relative utility of spatial boundaries and landmarks in an environment (Ratliff & Newcombe, 2008). Yet, how verbal cues modulate the binding of these lower-level features during encoding and retrieval remains understudied. Indeed, debate remains regarding whether repeated verbal interference hinders the ability to use either landmarks or the environment's shape to reorient (Hermer-Vazquez, Spelke, & Katsnelson, 1999; Ratliff & Newcombe, 2008). One largely unexplored possibility is whether verbal cues to use the boundaries or landmarks of an environment during navigation result in better learning of objects that are aligned versus misaligned to the instructed component.

In our paradigm, participants navigated and learned two virtual cities, each with a square spatial boundary and a landmark which was modified in each experiment (Experiment 1a: mountain range; Experiment 1b: single mountain; Experiment 2: coloured walls). Before navigating each city, participants were provided with a verbal cue to use either the square spatial boundary (geometry condition) or the landmark(s) (feature condition) of the virtual city to learn store locations. We then used the judgment of relative direction (JRD) pointing task to measure memory of store locations that were aligned or misaligned with the instructed component to determine how these opposing verbal cues affected spatial representations (Figures 1 and 2). Pointing errors were grouped by whether they were aligned or misaligned with the instructed component, which we term aligned-to-boundary and aligned-to-landmark contrasts throughout the text (Figure 2).

Our experimental design described above allowed us to test two hypotheses: the flexible coding versus obligatory binding hypothesis. The flexible coding hypothesis states that memory coding is largely task-dependent, and thus verbal cues to employ the landmark vs. boundaries should result in differential utilisation of either code (Eichenbaum & Cohen, 2001, 2014; Ekstrom & Ranganath, 2017) Thus, depending on the cue, memory for the location of a target relative to others (JRD performance) is better when aligned to the instructed component. This hypothesis predicts an interaction between instructed component and cue, with pointing error patterns changing based on whether participants are cued to use the boundary or landmark to encode object positions. The obligatory binding hypothesis, in contrast, conceptualises the binding of lower-level features as "obligatory" (e.g., Waller, Montello, Richardson, & Hegarty, 2002), consistent with the anchoring of the "cognitive map" to lower-level features like distal landmarks (Moser et al., 2008; O'Keefe, 1991; O'Keefe et al., 1998; Taube, 2007). Because encoding and binding are obligatory, alignment effects with reference to the boundary or landmark should not be modulated by verbal cues, thus predicting no interaction between cue and instructed component. Additionally, we analyzed navigation behaviour to assess if verbal cues simply primed participants to attend to the instructed component or if they integrated low-level and task-related features during encoding. Both measures allowed us to determine how and when during the formation of spatial representations verbal cues interacted with lower-level features.

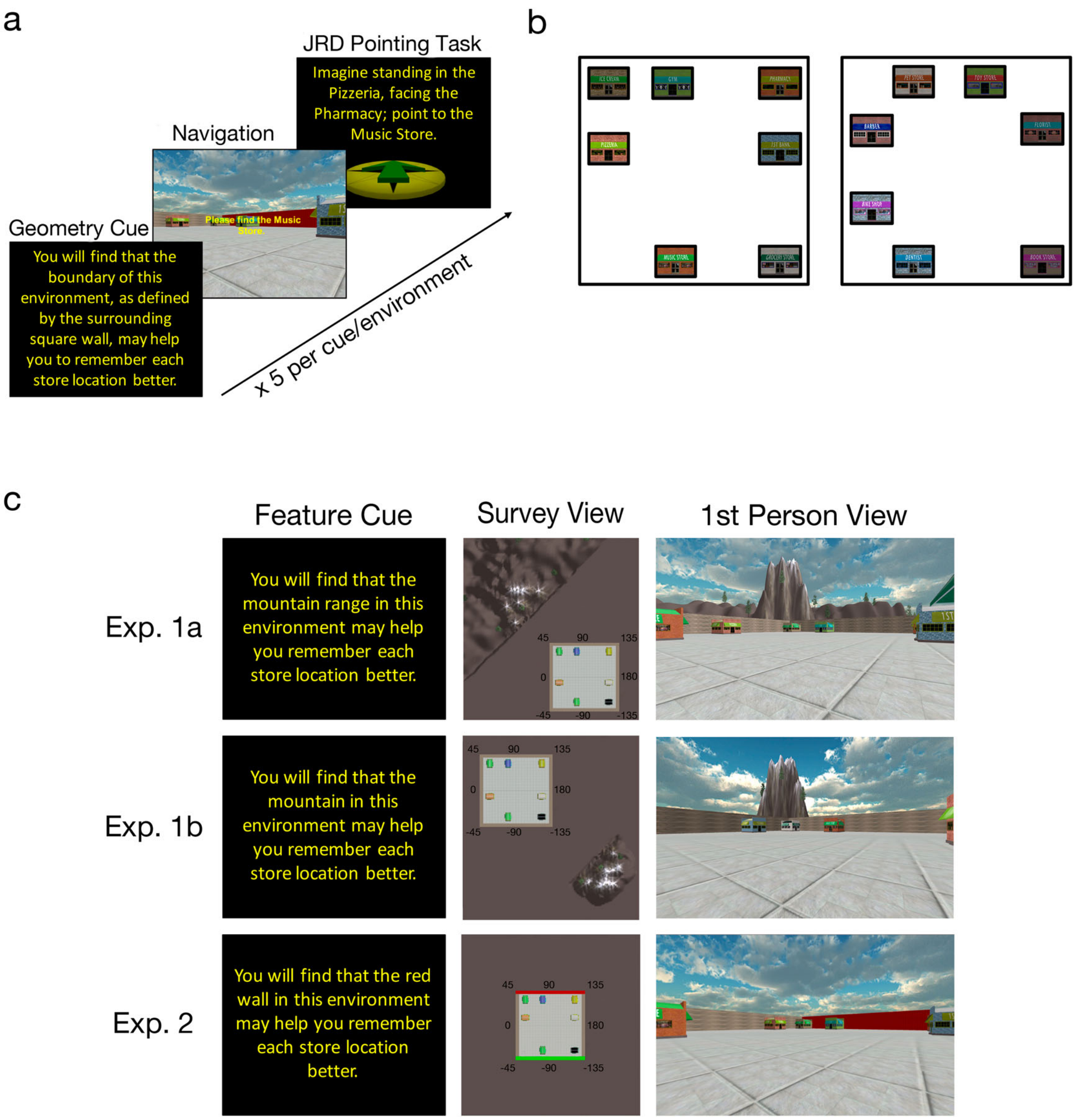
We note that the dichotomy between the flexible coding hypothesis and the obligatory binding hypothesis are at two extremes. A hybrid model may also be possible in which boundaries play a dominant role in encoding and are more obligatory than landmarks, but both can be accessed somewhat flexibly with verbal instruction. An example of this comes from the literature on reference frames which examines to what extent global reference frames (i.e., cardinal axes) and local reference frames (i.e., pairwise associations between visible landmarks in a vista space) are used to underlie spatial representations (Meilinger, Riecke, & Bülthoff, 2014; Weisberg, Badgio, & Chatterjee, 2018). This literature has found that both local and global reference frames can improve spatial memory of environments, but that local reference frames improve memory of visible landmarks encoded in vista spaces whereas global reference frames improve memory of external landmarks encoded with the cardinal axes (Meilinger et al., 2014; Weisberg et al., 2018). In the current experiments, the square boundary can be thought of as providing cardinal axes that do not change across different areas and thus is encoded globally and obligatorily regardless of verbal instruction. Conversely, the landmarks employed do not provide such a salient global frame of reference, and thus may be encoded locally, depending on verbal cue.

Method

Experiment 1a

Participants

Experiment 1a included 43 participants (32 females, mean age = 21.38). Participants were healthy volunteers from UC Davis or the surrounding community and either received payment or course credit for their participation. Participants were screened for neurological and vision impairments. The Institutional Review Board at the University of California, Davis approved all experimental protocols, with each participant giving written, informed consent. All methods were performed in accordance with the



Imagine standing in the Pizzeria, facing the Pharmacy; point to the Music Store.

You will find that the boundary of this environment, as defined by the surrounding square wall, may help you to remember each store location better.

You will find that the mountain range in this environment may help you remember each store location better.

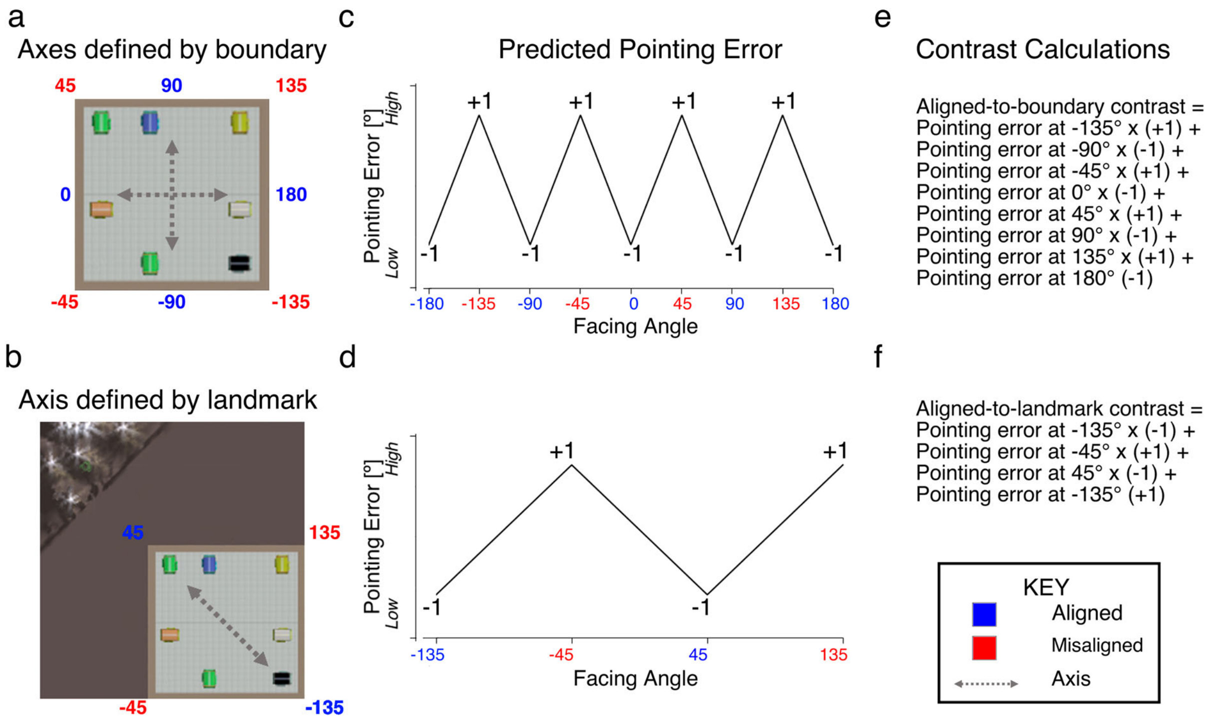
You will find that the mountain in this environment may help you remember each store location better.

You will find that the red wall in this environment may help you remember each store location better.

Figure 1. Experimental materials and procedures. During the task (a), participants were shown a verbal cue prior to navigation to emphasise either the spatial boundaries or the landmark. The geometry cue did not change across the three experiments, although we did change the landmark in Experiment 1a versus Experiments 1b & 2 to compare how different types of distal landmarks and local features affected our findings (see panel (c)). Participants then completed a navigation task ("navigation"), where they learned store locations. Finally, they performed the judgments of relative direction (JRD) task ("JRD pointing task"), in which they were asked to reference a third store in relation to two other stores. (b) Store layouts were fully counterbalanced with the geometry and feature cues. (c) The landmarks and feature condition verbal cues differed across the three experiments. In Experiment 2, the feature cue was fully counterbalanced to instruct participants to use either the red or green walls to aid navigation.

relevant guidelines and regulations. A minimum number of participants required was determined by an a priori power analysis (GPower; Faul, Erdfelder, Lang, & Buchner, 2007). Because we were looking for a novel interaction effect based on verbal cues, we did not have a clear way to estimate power. As such, we based our power analyses on aligned vs. misaligned pointing error studied in previous literature (Mou & McNamara, 2002) as we expected alignment effects to differ based upon verbal cues. The power analysis indicated each experiment required 40 participants to have 80% power to detect a medium-sized effect between aligned versus misaligned pointing error when employing an alpha of 0.05.

Participants were excluded based on the results of lower-tailed subject-specific non-parametric p-values from a permutation analysis. We chose this approach (Starrett, Stokes, Huffman, Ferrer, & Ekstrom, 2018) based on the logic that not all participants have guessing patterns that are greater 90° because each individual's guessing pattern differs and will not consistently be uniform. As such, we chose to randomly shuffle each subject's raw pointing angles with the correct angles 10,000 times to generate a subject-specific chance distribution of pointing error. Each participant's median pointing error was then compared to their chance distribution to determine if they performed better than chance performance (alpha



Aligned-to-boundary contrast =
Pointing error at -135° x (+1) +
Pointing error at -90° x (-1) +
Pointing error at -45° x (+1) +
Pointing error at 0° x (-1) +
Pointing error at 45° x (+1) +
Pointing error at 90° x (-1) +
Pointing error at 135° x (+1) +
Pointing error at 180° (-1)

Aligned-to-landmark contrast =
Pointing error at -135° x (-1) +
Pointing error at -45° x (+1) +
Pointing error at 45° x (-1) +
Pointing error at -135° (+1)

Figure 2. Contrast function calculations. The first panel depicts the survey view of each facing angle that was aligned (blue) and misaligned (red) with the square spatial boundary for all experiments (a) and Experiment 1a (b). The grey arrows depict axes that were aligned with the boundary (a) or the landmark (b). (c) and (d) depict predicted pointing error values for the aligned and misaligned facing angles. Based on this predicted pointing error, misaligned facing angles (expected to have greater pointing error) were assigned contrast weights of +1 and aligned facing angles (expected to have lower pointing error) were assigned contrast weights of −1. Facing angles that were not compared were assigned contrast weights of 0. Pointing error from each facing angle was multiplied by its respective contrast weight and summed to calculate the aligned-to-boundary (e) and the aligned-to-landmark (f) contrasts. We note that pointing error from −180° is not included in this calculation as −180° is only depicted to show the "sawtooth" shape of predicted pointing error as seen in previous studies of alignment (Mou & McNamara, 2002; Shelton & McNamara, 1997, 2001) effects.

criterion = 0.05). If a participant did not exceed criterion, their data was excluded. Here, participants were excluded as they performed worse than chance based on their subject-specific non-parametric p-value. In total, 40 participants were included in Experiment 1a analyses, as the power analysis indicated.

Stimuli and apparatus

Stimuli were presented and run on either a 27-inch iMac 2013 (3.5 GHz, 16GB 1600 MHz DDR3) or a 20-inch iMac 2008 (2.4 GHz, 6GB 800 MHz DDR2 SDRAM). Stimuli were presented with Unity 3d (Unity Technologies, San Francisco, CA). Two virtual cities were designed in Unity 3d using a modified version of Landmarks v1.0 asset pack (developed by the Human Spatial Cognition Lab with BrickOvenGames, http://humanspatialcognitionlab.org/software/). Each environment had a square boundary (90 × 90 m) and seven unique stores (each 9.5 × 6.5 × 5.8 m, length × width × height) placed within the boundary of the environment. Both environments were enriched with a tile sidewalk, sky, clouds, and a landmark, which differed in type (distal vs. local) and location for each experiment. Data from the virtual cities was combined for later analyses, as each environment was fully counterbalanced with each condition and cue. In Experiment 1a, a mountain with foothills served as the distal landmark that was located outside and misaligned to the square boundary at 45° with foothills that were visible from the 0° and 90° facing angles (Figure 1).

Procedure

Experiment 1a was completed in one session lasting between one and two hours. Before each session, participants completed a practice task, in which they navigated a novel environment to learn the instructions and computer keystrokes associated with navigation and JRD tasks. The practice environment included stores with coordinates distinct from the actual experiments. Participants were not provided with information regarding verbal cues during the practice task. After the practice task, participants began the experiment.

Using a within-subjects design, participants navigated two virtual environments (Figure 1) across two verbal cue conditions: a feature and a geometry condition. Store layouts were fully counterbalanced with the geometry and feature cues in which each subject would navigate each environment paired with one of the two cues (e.g., one subject may navigate store layout 1 paired with the feature cue and store layout 2 paired with the geometry cue or vice versa). Participants completed five blocks of intermixed learning and testing per condition/environment combination (10 blocks total). In each block, participants were first presented with a verbal cue, then completed a navigation task, and finally, were tested on store locations with the JRD task (Figure 1). Condition order was counterbalanced with the verbal cue and store layout combinations to ensure subjects were not biased towards one cue over another.

Verbal cues. Before verbal cues were presented, on-screen instructions were shown for the navigation task. Then, participants were presented with an on-screen verbal cue with a black background before each block. We used a repeated design for the instructions to ensure that participants were staying track and to remind them to use the relevant verbal cue to aid navigation. This method is often used in the adaptation literature in which subjects are readapted to a stimulus prior to the start of each learning block (Peacock & Gözenman, 2017). Before each condition, the experimenter gave an example of how the on-screen verbal cue might be used during the task (i.e., participants were instructed that they may be able to remember the positions of objects relative to one another using the axes of the relevant cue). In the geometry condition, participants were instructed to use the axes of the square environment to help them remember store locations relative to one another (10 s) (e.g., "You will find that the boundary of this environment, as defined by the surrounding square wall, may help you remember each store location better."). In the feature condition, participants were instructed to use the mountain range to help them remember store locations (10 s) (e.g., "You will find that the mountain range in this environment may help you remember each store location better.").

Navigation. At the beginning of each block, participants were placed into the centre of the environment, randomly facing one of the eight angles (0°, 45°, 90°, 135°, 180°, −135°, −90°, −45°) with replacement. They then navigated to each store location once in a randomised order per block. For each trial in which participants searched for a store, a yellow prompt in the centre of the screen indicated which store to find (10 s) (e.g., "Please find the Barber."). Participants moved to the target store using the arrow keys to control translations and the "A" and "D" keys to rotate the avatar clockwise and counterclockwise, respectively. When participants found the correct store, a yellow prompt indicated so (3 s) (e.g., "Great job; you found the Barber."). At the end of the navigation task, participants were shown instructions for the JRD pointing task.

JRD pointing task. During the JRD task, participants were presented questions with a black background on the computer monitor. A black background was used to prevent participants from using stimuli within each city to help them orient. In each trial, there was a message presented on the top half of the screen ("Imagine standing at Store A, facing Store B; please point to Store C.") and an onscreen compass and arrow in the lower half of the screen (Figure 1). Participants were instructed to use the "A" and "D" keyboard presses to rotate the arrow on the compass to point to the location of the third store relative to the location of the first two and the "return" key to advance to the next trial. When responding, participants were asked to point as accurately as possible. If participants were unsure of the correct pointing angle, they were required to guess. There was no time limit for participants to complete each trial. Participants completed 15 JRD trials per block of each condition. Because there were 5 blocks per condition, this resulted in 75 trials per condition and 150 trials total. JRD trial order was randomised for each participant.

JRD trials were created by a custom-made MATLAB script (MATLAB, Natick, MA), in which trials were generated based upon whether Store A and Store B were 1) aligned or misaligned with the mountain, 2) aligned or misaligned with the square boundary, and 3) if the mountain could be visualised at that facing angle. Fifteen trials were aligned to the mountain at 45°, 15 trials were aligned to the square boundary where the foothills were present (0°, 90°), 15 trials were aligned with the square boundary where no foothills were present (180°, −90°), 15 trials were misaligned with respect to both the mountain and square boundary where the mountains and foothills could be visualised (45°), and 15 trials were misaligned to the mountain and square boundary with no mountain visible in the at those facing angles (135°, −45°).

Data analyses

Data was processed and analyzed in MATLAB 2016a (Mathworks, Inc., Natick, MA) and RStudio (RStudio, Boston, MA).

Pointing error

We used absolute (unsigned) pointing error to assess performance for each condition on the JRD task, consistent with past work on the topic (McNamara et al., 2003; Mou & McNamara, 2002). Reaction time during the JRD task, due to the untimed nature of the task, was more variable and thus we focused on pointing error. Absolute pointing error was calculated as the absolute value of the difference between the actual angle between the facing angle (the vector between Store A and B) and the pointing target (the vector between Store A and C) versus the participant's estimate of that angle. Absolute pointing error was calculated for each JRD trial. Pointing error averages and standard deviations were calculated across participants by block, facing angle, and condition. Because the pointing error distributions were positively skewed in all experiments (Supplementary Figure S3), we removed outliers from the tails of these distributions. Pointing error from individual trials was removed by block, facing angle, and condition if they were greater than three standard deviations above the overall mean. These outlier criteria resulted in 4.08% of the total trials removed.

Pointing error was fit to aligned-to-boundary and aligned-to-landmark contrast functions (Meilinger, Strickrodt, & Bülthoff, 2016) to test the two hypotheses detailed in the introduction regarding the relative efficacy of verbal cues in spatial memory: the flexible coding hypothesis versus obligatory binding hypothesis. The aligned-to-boundary contrasts describe the sawtooth pattern of pointing error when aligned to geometry, as shown in previous studies of alignment (McNamara et al., 2003; Mou & McNamara, 2002), in which pointing error from aligned facing angles (0°, 90°, 180°, −90°) was lower than pointing error from misaligned facing angles (45°, 135°, −135°, −45°). Because the geometry cue instructed participants to use the square axis of the environment to remember each store location relative to another, we hypothesised that if there was an effect of the geometry cue, then the aligned-to-boundary contrast fit would be stronger for the geometry condition than the feature condition, as the feature cue instructed participants to use the distal landmark. This prediction is consistent with the flexible coding hypothesis, as outlined in the Introduction. To calculate the aligned-to-boundary contrast for each participant, aligned facing angles were assigned contrast weights of −1, whereas the misaligned facing angles were assigned contrast weights of +1 (Figure 2). Pointing error from each facing angle was then multiplied by its respective contrast weight and summed (e.g., −1 × average pointing error at 0° + 1 × average pointing error at 45°, etc.). Pointing error from −180° is not included in this calculation as −180° is the same as 180° and is only depicted to show the "sawtooth" shape of predicted pointing error in Figure 2. Positive contrast values indicate a good fit for the data (e.g., an alignment effect), whereas negative contrast values indicate that the inverse contrast provide a better fit the data (lower pointing error on misaligned trials than aligned trials). A contrast value of 0 predicts no difference in performance for aligned and misaligned JRD trials.

We chose the aligned-to-landmark contrast based upon theories from the reference frame literature (Chan et al., 2013; Meilinger et al., 2014; Weisberg et al., 2018) and function fits that we describe in Supplementary Material. Aligned-to-landmark contrasts were fit to the data to compare pointing error for trials that were aligned to the axes of the mountain range (45° and −135°) and misaligned (−45° and 135°) to the axes of the mountain range. Contrast weights of −1 were assigned to aligned facing angles (45°, −135°), whereas contrast weights of +1 were assigned to the misaligned facing angles (−45°, 135°). All other facing angles were assigned contrast weights of 0, as these were not comparisons of interest. To calculate individual contrast fits, each participant's pointing error was multiplied by its respective contrast weight, and summed (e.g., −1 × average pointing error at 45° + 1 × average pointing error at −45°, etc.). Positive contrast values indicate lower pointing error on trials that were aligned with the mountain range. Negative contrast values indicate the opposite fit, where pointing error was lower for trials that were misaligned with the axes of the mountain range. A contrast value of 0 indicates there was no difference in pointing error for axes aligned or misaligned with the mountain. Because the feature cue instructed participants to use the mountain to remember store locations, it was hypothesised that the aligned-to-landmark contrasts should be greater for the feature condition than the geometry condition, if there were an effect of feature verbal cue. This prediction is consistent with the flexible coding hypothesis.

Rotations during navigation and visual attention

Because navigation and the JRD task were interspersed, we were also interested to assess if the direction that participants faced during navigation differed as a function of verbal cue. Rotations during navigation were analyzed to assess whether verbal cues simply caused participants to attend to one cue over the other or if low-level and task-related features were integrated into encoding in a more complex and configural fashion than simple attention can explain. We used two measures to assess these hypotheses: average facing angle and the percentage of time spent oriented towards each facing angle.

Facing angles. Avatar rotations during navigation were assessed with the MATLAB circular statistics toolbox (Berens, 2009). The distribution of rotations (out of the 360° of rotational space) in each condition was evaluated for circular uniformity. The geometry condition cued participants to use the aligned facing angles with the boundaries of the environment, which would result in an overall uniform distribution. In contrast, facing angles selective to the distal landmark would result in a non-uniform distribution. Thus, we hypothesised that only the feature condition should be significantly skewed if participants attended selectively to the feature by looking at it longer.

Time. The 360° of rotational space was divided into 16 bins, where each bin was 22.5° wide. Bins that were centred on each of the eight facing angles (0°, 45°, 90°, 135°, 180°, −90° 135°, −135°, −45°). The percentage of time spent facing each angle was fit to modified aligned-to-boundary and aligned-to-landmark contrast functions to assess if participants were simply attending to the mountain more or if encoding was more integrated and configural. The aligned-to-boundary and aligned-to-landmark contrast functions for pointing error assigned cued facing angles contrast weights of −1 and uncued facing angles contrast weights of +1 because pointing error was expected to be *lower* for cued versus uncued facing angles. However, because a simple attentional explanation would assume that a *greater* percentage of time would be spent in a cued versus uncued facing angles, the opposite aligned-to-boundary and aligned-to-landmark contrast functions were applied to the navigation data, where cued facing angles were assigned contrast weights of +1 and uncued facing angles were assigned contrast weights of −1. The percentage of time spent facing each angle was then multiplied by its respective contrast weight and summed to calculate each subject's aligned-to-boundary and aligned-to-landmark contrast value in each condition. A positive contrast value indicated that a greater percentage of time was spent facing a cued facing angle, whereas a negative contrast value indicated that a greater percentage of time was spent facing an uncued facing angle. A contrast value of 0 indicated that there was no difference in the percentage of time spent in the cued or uncued facing angles.

Bayesian approach

Because one of our two competing hypotheses was always a null hypothesis, we also analyzed our data using a Bayesian framework with the BayesFactor package for RStudio (Morey & Rouder, 2015). For repeated measures factorial analyses, we used the ANOVA Bayes factor (BF) and for post-hoc tests and for mean comparisons we used the BF *t*-test with default settings. Because our hypotheses were prefaced on whether an interaction was present for the ANOVAs, we calculated the BF by comparing the full model that included the main effects and interactions with a reduced model in which the interaction was not included.

Experiment 1b

Participants

Forty-one new participants (25 females, mean age = 22.68) partook in Experiment 1b. One dataset was excluded based upon subject-specific non-parametric p-values. In total, 40 datasets were analyzed.

Stimuli and apparatus

All stimuli and apparatus remained identical except for the distal landmark. The mountain was rotated to−135° to ensure that the effect of verbal cue was due to the distal landmark itself, rather than its location. Moving the landmark to −135° served as a test to determine if our findings could be replicated with the distal landmark in a new location. Additionally, the foothills from Experiment 1a were removed so that only the mountain could only be visualised at −135° to remove contamination between the geometric and featural information (Figure 1).

Procedure

The procedure was identical except for the feature cue and the number of JRD trials. Because there was a single mountain, the feature cue was adjusted accordingly (10 s) (e.g., "You will find that the mountain in this environment may help you remember each store location better."). As the alignment effect literature has previously based the number of JRD trials on facing angle (McNamara et al., 2003; Mou & McNamara, 2002; Shelton & McNamara, 1997, 2001) rather than field of view as we did in Experiment 1a, we included 15 JRD trials for each of the eight facing angles to ensure we could replicate our effects using these methods. Participants completed 24 JRD trials per block of each condition. Because there were 5 blocks in each condition, this resulted in 120 trials per condition and 240 total trials in Experiment 1b. JRD trial order was randomised for each participant.

Data analyses

Data was analyzed based on those conducted in Experiment 1a. A total of 4.96% of the total pointing error trials were excluded based on the exclusion criteria described in Experiment 1a.

Experiment 2

Participants

Forty-seven new participants partook in Experiment 2 (34 females, mean age = 21.50). Seven datasets were excluded based upon below chance performance. In total, forty datasets were analyzed.

Stimuli and apparatus

All stimuli and apparatus remained identical except two coloured walls (a red wall at 90° and a green wall at −90°) served as the local landmarks that formed constituent walls of the square boundary (Figure 1).

Procedure

The procedure was identical to Experiment 1a except that the feature cue was modified and JRD trial counts differed. The feature cue was modified to account for the red wall (e.g., "You will find that the red wall in this environment may help you remember each store location better.") and green wall (e.g., "You will find that the green wall in this environment may help you remember each store location better."). The feature cues were counterbalanced in a between-subject fashion in which half of the participants were cued to use the red wall and the other half were cued to use the green wall. JRD trials were broken down into 15 trials orthogonally aligned to the red wall (90°), 15 trials orthogonally aligned to the green wall (−90°), 15 trials orthogonally aligned to the uncoloured walls (0°, 180°), and 15 misaligned trials (45°, 135°, −135°, −45°). Participants completed 12 trials per block totalling 60 trials per condition and 120 trials total. Trial order was randomised for each participant.

Data analyses

A total of 3.87% of pointing trials were excluded based on the criteria described in Experiment 1a. Analyses remained identical with the exception of two aligned-to-landmark contrasts centred on 90° and −90°: one that compared pointing performance and a second that compared the percentage of time spent in each facing angle for the coloured walls (90°, −90°) relative to the tan walls (0°, 180°). These aligned-to-landmark contrasts were calculated in the same fashion as the contrasts described in Experiment 1a. Positive contrast values indicated that pointing error was lower and that a greater percentage of time was spent in facing angles containing the coloured walls, respectively, whereas negative contrast values indicated an inverse fit.

Results

Experiment 1a

Experiment 1a examined the interface of verbal cues with spatial boundaries versus a distal landmark to test if representations could be modulated by task-related features or if they are bound in an obligatory manner. Specifically, we tested whether instructing participants to use the spatial geometry to remember the environment would result in lower pointing error for facing angles aligned with the axes defined by the spatial boundaries while instructing participants to use a distal landmark would result in lower pointing error for facing angles aligned with the axis defined by the landmark.

Pointing error

We conducted a 2 cue (geometry, feature) × 2 contrast function (aligned-to-boundary, aligned-to-landmark) repeated measures ANOVA on pointing error to evaluate the flexible coding versus obligatory binding hypothesis in Experiment 1a. A significant interaction between cue and contrast function would indicate that pointing error patterns change based on whether participants are cued to use the boundary or landmark to code object positions, thus supporting the flexible coding hypothesis. In contrast, a nonsignificant interaction would suggest that alignment effects and pointing error patterns should not be modulated by verbal cues, thus supporting the obligatory binding hypothesis. The ANOVA indeed revealed a significant interaction between cue and contrast: $F(1,39) = 13.04$, $p = 0.0008$, $\eta^2 = 0.05$. These data indicate that alignment effects and pointing error patterns changed depending on whether participants were cued to use the spatial boundary or the landmark, thus supporting the flexible coding hypothesis. We were also interested to see if verbal cues interfaced with the target facing angle (A-C facing angle). The interested reader may refer to the Supplemental Material for these analyses.

Since one of the two competing hypotheses was a null hypothesis (the obligatory binding hypothesis), we used a Bayesian framework to test for evidence for the null hypothesis. Because our hypotheses were prefaced on whether an interaction was present or not, we calculated the BF for the interaction in the repeated measures ANOVA (2 cue × 2 contrast function) by comparing the full model that included the main effects and interactions with a reduced model in which the interaction was not included. Here, we found strong evidence that the results were not obtained under the null hypothesis ($BF_{10} = 10.31$). The BF suggests that it is unlikely that there is no interaction between verbal cues and instructed components on pointing error patterns.

We also wanted to assess if the effect of verbal cue persisted when including block as a factor to assess if there were learning specific effects. Overall, considering block as an additional modelled variable did not change the overall results: a 2 cue × 2 contrast function × 5 block repeated measures ANOVA yielded a significant 2-way interaction for cue and contrast function: $F(1,39) = 4.77$, $p = 0.03$, $\eta^2 = 0.005$. For more details on findings from the ANOVA with block as a factor (including 3-way interactions, which were not part of our *a priori hypotheses* and inconsistently present across experiments), please see the Supplemental Material.

To obtain further evidence that there was an effect of verbal cue, post-hoc, pairwise t-tests were conducted in conjunction with BF *t*-tests to compare the relative fits of the contrast functions for pointing error between conditions. Overall, the aligned-to-boundary contrast for the geometry ($M = 14.26$, $SD = 58.71$) condition provided a significantly better fit to the data than the aligned-to-boundary contrast for the feature ($M = -10.81$, $SD = 44.32$) condition: $t(39) = 2.01$, $p = 0.01$, 95% CI [4.79, 45.17], with anecdotal evidence that the results were not obtained under the null hypothesis ($BF_{10} = 2.76$). Similarly, the aligned-to-landmark contrast for the feature ($M = 11.97$, $SD = 30.28$) condition provided a significantly better fit to the data than the aligned-to-landmark contrast for the geometry ($M = -2.50$, $SD = 29.48$) condition: $t(39) = -2.20$, $p = 0.03$, 95% CI [−27.77, −1.17] with anecdotal evidence that the results were not obtained under the null hypothesis ($BF_{10} = 1.47$) (Figure 3 and Figure S1). Together, the results of the post-hoc tests and Bayes factors suggest that memory for the target was significantly enhanced for objects that were aligned with the instructed component. In addition, these results were likely not obtained under the null hypothesis stating there were no changes in memory based on verbal cues. Additionally, we assessed if field of view played a role in how landmarks were represented. See the Supplemental Material for these analyses.

To assess evidence for a hybrid model, in which boundaries are encoded obligatorily but that landmarks and boundaries could be modulated somewhat flexibly via verbal cues, we conducted one-sample t-tests for each of the contrast fits for each condition against 0. Here, if the aligned-to-boundary contrasts were significantly greater than 0 regardless of condition, this would provide evidence that boundaries are encoded obligatorily. On the other hand, if the aligned-to-landmark contrasts are significantly greater than 0 for the feature condition but not the geometry condition, this suggests that verbal cues modulated the extent that the landmark was used. Overall, the aligned-to-boundary contrast was not significantly greater than 0 for the geometry ($t(39) = 1.53$, $p = 0.13$, 95% CI [−4.51, 33.04], $BF_{10} = 0.50$) or feature ($t(39) = -1.54$, $p = 0.13$, 95% CI [−24.99, 3.36], $BF_{10} = 0.50$) conditions. On the other hand, the aligned-to-landmark contrast was significantly greater

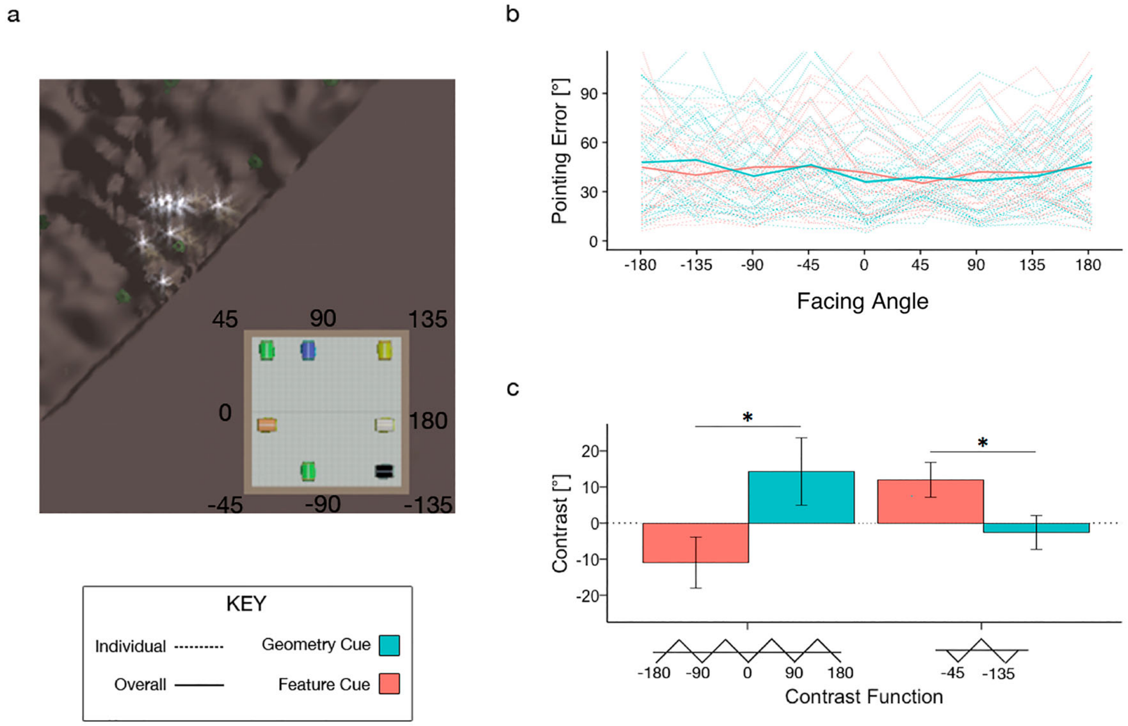


Figure 3. Experiment 1a pointing error results. (a) shows the survey view and facing angles defined in in Experiment 1a. Individual (dotted lines) and overall (solid lines) pointing error at each facing angle for each condition (feature = pink; geometry = blue). (b). Aligned-to-boundary and aligned-to-landmark contrast fits for each condition (c). Positive contrast values indicate a good fit for the data, whereas negative values indicate an inverse fit of the data. Significance stars are coded based on p-values (*: $p < 0.05$). Error bars reflect standard errors.

than 0 for the feature condition ($t(39) = -0.54$, $p = 0.59$, 95% CI [−11.93, 6.92], $BF_{10} = 0.20$) but not the geometry condition ($t(39) = 2.50$, $p = 0.02$, 95% CI [2.28, 21.65], $BF_{10} = 2.63$). This suggested that in Experiment 1a, boundaries were not necessarily encoded obligatorily but verbal cues did affect how landmarks were used.

To ensure that the effect of verbal cue was not due to significant differences in overall error for each condition, regardless of alignment, we compared overall performance between conditions. A paired t-test revealed no significant differences in overall pointing error for the feature ($M = 40.27$, $SD = 18.03$) and geometry ($M = 40.70$, $SD = 19.91$) conditions: $t(39) = -0.14$, $p = 0.89$, 95% CI [−6.57, −5.74] with moderate evidence for the null hypothesis ($BF_{10} = 0.17$). Together, these findings suggested that verbal cues directly influenced the heuristics participants used to code the spatial layout, again supporting the flexible coding hypothesis and that this difference was not due to better encoding in one condition over the other.

Rotations during navigation

Raleigh's test for circular non-uniformity was conducted on the circular distributions for each condition. If attention was simply modulated by verbal cues, only facing angles during navigation for the feature cue distribution should be significantly skewed whereas the geometry cue distribution should be uniform as the aligned facing angles in the geometry condition were uniformly distributed. A significant skew for both cue distributions would indicate that encoding was more configural. Overall, Raleigh's test showed that both distributions were significantly non-uniform (both $ps < 0.001$).

The percentage of time spent each facing angle during navigation was then fit to aligned-to-boundary and aligned-to-landmark contrast functions. A 2 cue (geometry, feature) × 2 contrast function (aligned-to-boundary, aligned-to-landmark) repeated-measures ANOVA was conducted to test if attention was simply modulated by verbal cues or if encoding was more configural and integrated than a simple attention-based explanation can provide. Here, an interaction effect would indicate that verbal cues selectively changed the percentage of time participants faced the instructed component whereas no interaction would indicate that verbal cues did not change the percentage of time spent facing the instructed component. Overall, there was no significant interaction between cue and contrast when considering the percentage of time spent facing a cued facing angle, $F(1,39) = 0.62$, $p = 0.44$, $\eta^2 = 0.003$ (Table 1). The Bayes factor for the interaction effect provided strong evidence that our results were obtained under the null hypothesis ($BF_{10} = 0.31$). These findings suggest that the effect of verbal cue on pointing error was not simply due to participants spending more time facing the instructed component during navigation. Additionally, we assessed how navigation trajectories to each store changed as a function of block and condition which is contained in the Supplemental Material (Figure S2).

Table 1. Experiments 1–2: Contrast function Means (Standard Deviations) for the percentage of time spent oriented towards each facing angle for each condition.

	Geometry cue		Feature cue	
Exp.	Aligned-to-boundary contrast	Aligned-to-landmark contrast	Aligned-to-boundary contrast	Aligned-to-landmark contrast
1a	11.02 (25.11)	9.06 (7.40)	7.82 (21.73)	10.19 (9.05)
1b	−4.75 (22.50)	15.36 (17.99)	−1.69 (22.82)	19.66 (13.98)
2	18.00 (15.62)	11.73 (20.99)	17.76 (18.31)	9.23 (17.42)

Experiments 1b and 2

Experiments 1b and 2 were designed to conceptually replicate Experiment 1a. Therefore, we grouped the two studies together for brevity. In Experiment 1b, we adjusted the location/size of the distal landmark, and in Experiment 2, we overlapped the feature and geometry axes. These modifications tested the generalizability of verbal cues to landmarks of new shapes and positions (Exp. 1b) and ones reminiscent of the local “features” (Cheng & Newcombe, 2005) from the feature and geometry literature (Exp. 2).

Pointing error

As in Experiment 1a, we conducted a 2 cue (geometry, feature) × 2 contrast function (aligned-to-boundary, aligned-to-landmark) repeated measures ANOVA to evaluate the flexible coding versus obligatory binding hypothesis in Experiments 1b and 2. The flexible coding hypothesis predicts an interaction between instructed component and verbal cue in which pointing error patterns change based on whether participants are verbally cued to use the boundary or landmark. The obligatory binding hypothesis predicts no interaction between instructed component and verbal cue in which alignment effects and pointing error patterns are not affected by verbal cues. The ANOVA indeed revealed a significant interaction between cue and contrast in both experiments (Exp. 1b: $F(1,39) = 11.84$, $p = 0.008$, $\eta^2 = 0.03$; Exp. 2: $F(1,39) = 5.72$, $p = 0.02$, $\eta^2 = 0.04$). Additionally, the BF for the interaction showed anecdotal evidence that the results were not obtained under the null hypothesis (Exp. 1b: $BF_{10} = 1.86$; Exp. 2: $BF_{10} = 2.71$). Together, these findings suggest that verbal cues enhanced pointing error performance and thus memory for the instructed component (Figure 4, Table 2, Figure S1). We were also interested to see if verbal cues interfaced with the target facing angle (A-C facing angle) and have included the relevant analyses in the Supplemental Material for the interested reader.

Table 2. Experiments 1–2: Contrast function Means (Standard Deviations) for absolute pointing error.

	Geometry cue		Feature cue	
Exp.	Aligned-to-boundary Contrast	Aligned-to-landmark contrast	Aligned-to-boundary contrast	Aligned-to-landmark contrast
1a	14.26 (58.71)	−2.50 (29.48)	−10.81 (44.32)	11.97 (30.28)
1b	33.96 (25.32)	−2.60 (14.98)	22.23 (27.49)	0.25 (13.06)
2	39.75 (45.08)	3.95 (27.35)	18.06 (44.95)	9.80 (19.67)

We included block as a variable (i.e., 2 cue × 2 contrast function × 5 block repeated measures ANOVA) to assess if block still yielded a significant 2-way interaction for cue and function in each experiment and to test if there were any learning specific effects (Exp. 1b: $F(1,39) = 8.72$, $p = 0.003$, $\eta^2 = 0.01$; Exp. 2: $F(1,39) = 9.12$, $p = 0.0006$, $\eta^2 = 0.01$). Details of other effects from this ANOVA are contained in Supplemental Material. These findings again support the conclusion that verbal cues modulated the extent to which participants used the cued component to encode the spatial layout.

Post-hoc, pairwise t-tests were conducted in conjunction with Bayesian t-tests to compare the fits of the contrast functions between conditions in each experiment. In both experiments, the aligned-to-boundary contrast for the geometry condition provided a significantly better fit to the data than the aligned-to-boundary contrast for the feature condition (Exp. 1b: $t(39) = 3.07$, $p = 0.004$, 95% CI [4.004, 19.47]; Exp. 2: $t(39) = 2.17$, $p = 0.03$, 95% CI [1.46, 41.92]). In Experiment 1b, there was moderate evidence that the results were not obtained under the null hypothesis ($BF_{10} = 9.23$). In Experiment 2, there was anecdotal evidence that the results were not obtained under the null hypothesis ($BF_{10} = 1.38$). Aligned-to-landmark contrasts were then compared which showed that the feature aligned-to-landmark contrast was numerically but not statistically better than the geometry aligned-to-landmark contrast in both experiments (Exp. 1b: $t(39) = -1.005$, $p = 0.32$, 95% CI [−8.61, 2.89]; Exp. 2: $t(39) = -1.19$, $p = 0.24$, 95% CI [−15.76, 4.06]). In both experiments, there was anecdotal evidence that the results were most likely obtained under the null hypothesis (Exp. 1b: $BF_{10} = 0.27$; Exp. 2: $BF_{10} = 0.33$). See Table 2 for these values. These findings suggest that the verbal cues modulated the extent to which participants used geometry as a heuristic but did not lead to the same direct use of feature cues, as in Experiment 1a. We return to this issue in detail in the discussion. Additionally, the interested reader should refer to the Supplemental Material for further information on how field of view played a role in how landmarks are represented.

To test for evidence in support of a hybrid model, we conducted one-sample t-tests for each of the contrast fits in each condition against 0. If the aligned-to-boundary contrasts are significantly greater than 0 regardless of condition, this provides evidence that boundaries are encoded obligatorily. Conversely, if the aligned-to-landmark contrasts are significantly greater than 0 for the feature condition but not the geometry condition, this suggests that verbal cues modulated the extent that the landmark was used. Overall, the aligned-to-boundary contrast was significantly greater than 0 for the geometry and feature conditions in both Experiments 1b (geometry: $t(39) = 8.48$, $p < 0.001$, 95% CI [25.86, 42.06], $BF_{10} = 4.85 \times 10^7$; feature: $t(39) = 5.11$, $p < 0.001$, 95% CI [13.43, 31.02], BF_{10}

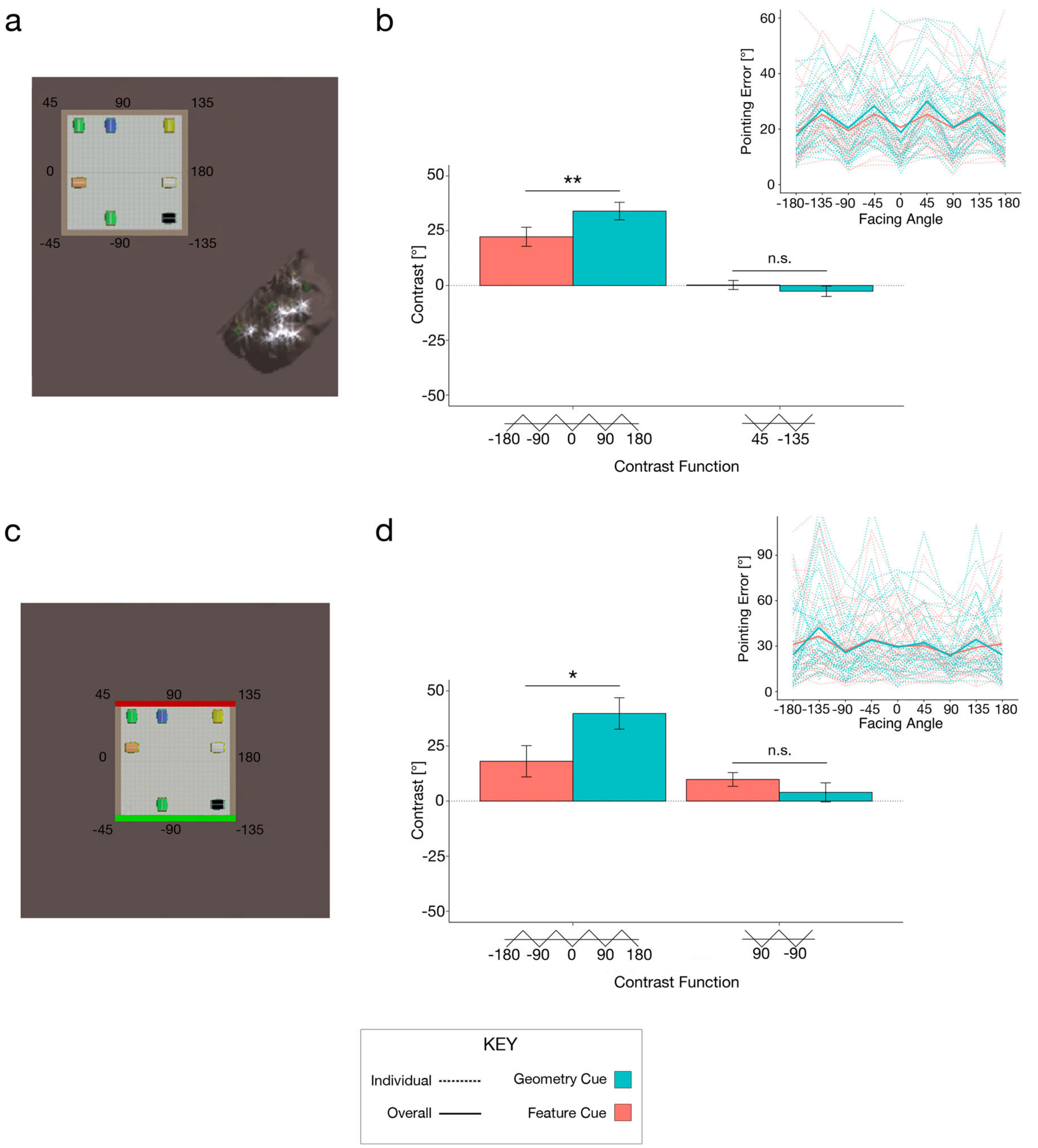


Figure 4. Experiments 1b and 2 pointing error results. Facing angles where landmarks were placed in Experiment 1b (a) and Experiment 2 (c). Contrast functions for the geometry (blue) and feature (pink) conditions for Experiment 1b (b) and Experiment 2 (d). The spaghetti plots located at the right upper inlets of (b) and (d) depict individual and overall pointing performance at each facing angle in Experiments 1b and 2, respectively. Significance stars are coded based on p-values (i.e., **: $p < 0.01$; *: $p < 0.05$; n.s.: $p > 0.05$). Error bar values reflect standard errors.

= 2271) and 2 (geometry: $t(39) = 5.58$, $p < 0.001$, 95% CI [25.33, 54.17], $BF_{10} = 8912$; feature: $t(39) = 2.54$, $p = 0.02$, 95% CI [3.69, 32.44], $BF_{10} = 2.87$). In Experiment 1b, the aligned-to-landmark contrast was not significantly greater than 0 for the feature condition ($t(39) = 0.12$, $p = 0.90$, 95% CI [−3.92, 4.43], $BF_{10} = 0.17$) or the geometry condition ($t(39) = -1.09$, $p = 0.28$, 95% CI [−7.40, 2.19], $BF_{10} = 0.30$). For Experiment 2, however, the aligned-to-landmark contrast was significantly greater than 0 for the feature condition ($t(39) = 3.15$, $p = 0.003$, 95% CI [3.50, 16.01], $BF_{10} = 11.13$) but not the geometry condition ($t(39) = 0.91$, $p = 0.37$, 95% CI [−4.80, 12.70], $BF_{10} = 0.25$). These results suggested evidence for a hybrid model in which boundaries are encoded obligatorily regardless of cue yet verbal cues may nonetheless still modulate how strongly boundaries and landmarks are employed when retrieving the target locations.

We also compared overall pointing performance between conditions to ensure that the effect of verbal cue was not simply due to significant differences in overall pointing error regardless of alignment. A paired t-test revealed no significant differences in overall pointing error between conditions in both experiments (Exp. 1b: $t(39) = -0.68$, $p = 0.50$, 95% CI [−3.76, 1.86]; Exp. 2: $t(39) = 0.85$, $p = 0.40$, 95% CI [−2.26, 5.63]). Bayes factors showed that these results were likely obtained under the null hypothesis (Exp. 1b: $BF_{10} = 0.21$; Exp 2: $BF_{10} = 0.17$). These findings suggest that the effect of verbal cue was not due to better encoding in one condition versus another (Table 3). Additionally, we conducted a one-way ANOVA to assess whether there were any overall differences in performance across the three experiments to provide insight into the differing effects of verbal cues across our three experiments. This indeed showed a significant difference between Experiment 1a ($M = 39.97$, $SD = 16.51$), 1b ($M = 22.92$, $SD = 8.45$), and 2 ($M = 27.95$, $SD = 13.65$): $F(2,117) = 17.36$, $p < 0.001$. This finding suggests that pointing error was highest in Experiment 1a, possibly one of the reasons we found the strongest evidence for the effects of verbal cue on utilisation of geometry versus features.

Rotations during navigation

Raleigh's test for circular non-uniformity showed that the facing angles during navigation for each condition were significantly non-uniform in both experiments (all $ps < 0.001$). The percentage of time spent facing each facing angle was then fit to aligned-to-boundary and aligned-to-landmark contrast functions for each condition. A 2 cue (geometry, feature) × 2 contrast function (aligned-to-boundary, aligned-to-landmark) repeated-measures ANOVA was conducted to test whether attention was simply modulated by verbal cues or if encoding was more integrated and configural. Here, an interaction effect indicates that the percentage of time facing the instructed facing angle was simply modulated by verbal cues during navigation/encoding whereas no interaction suggests that the relationship between verbal cues and instructed component have no effect on the percentage of time facing the instructed angle. As in Experiment 1a, there was no significant interaction between cue and contrast in Experiments 1b and 2 (Exp. 1b: $F(1,39) = 0.04$, $p = 0.82$, $\eta^2 = 0.0002$; Exp. 2: $F(1,39) = 0.32$, $p = 0.57$, $\eta^2 = 0.0009$) thus showing that verbal cues did not change the percentage of time spent facing the instructed component during encoding while navigating (Table 1). Using a Bayesian framework strengthened evidence for this pattern of results in both experiments by showing strong evidence that the results were likely obtained under the null hypothesis that encoding was more configural than a simple attention-based account can provide (Exp. 1b: $BF_{10} = 0.22$; Exp. 2: $BF_{10} = 0.25$).

Table 3. Experiments 1–2: Overall pointing error Means (Standard Deviations) for each condition.

Exp.	Geometry cue	Feature cue
1a	40.70 (18.91)	40.27 (18.03)
1b	23.54 (9.36)	22.59 (9.98)
2	31.60 (17.13)	33.29 (17.14)

General discussion

Our findings showed that verbal cues modulated how participants used a boundary or landmark, which in turn provided for spatially-relevant reference frames, to retrieve object locations. This was supported by a significant interaction between cue and contrast function in all experiments, an effect clearest in Experiment 1a. Based on instructions to use the boundary, participants remembered the layout when aligned with the axes defined by the boundary (−90°,0°/90°/180°) better versus instructions to use the mountain. Conversely, when instructed to use the mountain, pointing error was lowest when aligned to the single axis defined by the distal landmark (45°/−135°) versus instructions to use the boundary. A Bayesian framework, which allowed us to better assess for null evidence, supported these findings, showing that the data were most likely obtained under the alternative (rather than null) hypothesis stating that verbal cues modulated how participants retrieved object locations.

Experiments 1b and 2 showed evidence for the same pattern with some caveats. Experiment 1b involved a smaller-sized mountain (no extending foothills) in a new location. We again found an interaction effect between cue and alignment to a reference frame, showing that instructions to use the boundaries resulted in lower pointing error when aligned to the square axes versus instructions to use the distal landmark. We also found a numerically better fit to a model that considered the distal landmark when instructed to use the distal landmark versus the geometry. Using a different paradigm in Experiment 2 in which the instructed components shared axes, we again found a significant cue by reference frame interaction effect. However, post-hoc tests showed that although the fit when instructed to use the boundaries versus features was significantly higher, this was not true for the fit when instructed to use the features versus the boundary. Although the reasons for not obtaining an exact replication from Experiment 1a are unclear, perhaps this is because the landmarks had different properties (more focal in Experiment 1b and more integrated with the geometry in Experiment 2), which is supported by a lack of a clear fit to the aligned-to-landmark functions. In this way, it is possible that landmark representations were more integrated with the boundaries when instructed to use the features, and thus a clear effect of feature was not present. Additionally, the long mountain range with foothills from Experiment 1a may have acted as a directional cue used by participants for alignment, consistent with the idea that reorientation may be dependent upon the relative utility of geometry and features in an environment (Ratliff & Newcombe, 2008). Conversely, it is also possible that the spatial geometry was more salient in Experiments 1b and 2 compared to Experiment 1a due to the nature of the less prominent features employed. Because there were overall differences in performance across the three experiments, yet only the landmark changed as a function of experiment, this further supports the idea that the properties of landmarks may have played a role in how participants globally remembered the layout of each experiment. Despite these differences, each experiment supports the flexible coding over obligatory binding hypothesis because verbal cues significantly modulated how the instructed heuristic was employed to remember the environment, although experiments 1b and 2 are more consistent with a hybrid model, an issue we explore in more depth shortly.

Because the effect of verbal cue may have been attributed simply to attending to one cue over another, we also assessed participants' facing angles during navigation. An attentional account based on encoding some features at the expense of others predicts that participants looked longer at the instructed component, and thus, based on encoding this information better, retrieved it more readily. All three experiments, however, when considering the percentage of time spent facing each angle, showed no interaction between cue and contrast. Because the circular distributions were not changed by verbal cues, yet pointing error was, this suggests that low-level and task-related features are integrated into encoding in a more configural way than a simple attention-based explanation can explain. While we did not collect eye-tracking data to verify exactly where participants fixated, facing angle during navigation provided some insight into what aspects of the spatial layout participants attended to (Ekstrom et al., 2003). Together, our findings support ideas from the navigation and attentional literature that representations are updated during navigation (Burgess, 2006; Mou & Wang, 2015; Wang & Spelke, 2000) and that "top-down" and "bottom-up" mechanisms during navigation are integrated versus dichotomised (Awh et al., 2012), respectively.

Our results, particularly those in Experiments 1b and 2, are more consistent with a hybrid model of reference frame and flexible coding based on verbal cue, which incorporates the obligatory binding and flexible coding hypotheses. A hybrid model suggests that boundaries are more obligatory than landmarks, but both can be accessed somewhat flexibly with verbal cues. Given that there was always an alignment effect with respect to the boundaries in all experiments, yet this was only true for the mountain range in the first experiment, it is likely that both verbal instruction and the lower-level characteristics of competing cues affected the organisation of spatial memories. It is also possible that the boundary was used as a global reference frame (Meilinger et al., 2014; Weisberg et al., 2018) in which participants used the cardinal axes formed by the boundary obligatorily to remember the spatial configurations of stores. On the other hand, the landmarks did not provide a directional cue, and thus may have been remembered via local frames of reference (Meilinger et al., 2014; Weisberg et al., 2018) thus resulting in boundaries playing a dominant role in encoding, with both boundaries and landmarks being somewhat accessible by verbal cues. It is important to note, however, that the hybrid model still argues for the efficacy of verbal cues in promoting reference frame switches and suggests that reference frames employed during retrieval can be flexibly accessed based on task-demands.

Language studies suggest that spatial language and memory have similar properties in those who speak different languages (Munnich, Landau, & Dosher, 2001). Likewise, one study (Taylor & Tversky, 1992) noted that similar representations derive from learning routes from text- and map-based descriptions. Yet whether verbal cues directly affect how we structure spatial representations remains unclear. Our results shed light on the interface between language and navigation by suggesting that verbal cues alter what heuristics are used to access spatial representations. Our findings thus conflict with past conceptualizations of space, such as "survey representations," which are thought to arise first from the perception of distal landmarks (Lynch, 1960; Siegel & White, 1975) or that this integration process, for some components at least, is obligatory (Waller et al., 2002). To further address the efficacy of verbal cues, future studies should address whether other cues, such as visual cues, influence spatial representations in the same way as verbal cues. Because landmarks and boundaries are often viewed as inherently spatial (Bird, Capponi, King, Doeller, & Burgess, 2010; Epstein, DeYoe, Press, Rosen, & Kanwisher, 2001) while verbal cues do not possess the same low-level spatial attributes (related to navigation at least), visual cues may provide another bridge to understanding the flexibility of the cognitive map.

Studies focused on the neural basis of navigation often argue that higher-order representations, like the cognitive map, involve an "intrinsic" spatial metric anchored to distal landmarks (O'Keefe, 1991; O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1978). These models, primarily based on neural recordings in rodents, do not have an explicit means by which verbal cues can modulate representations nor how heuristics interact with them (Ekstrom, Huffman, & Starrett, 2017). Indeed, some versions of cognitive map theory suggest verbal cue storage occurs in the left hippocampus while spatial metrics are stored in the right hippocampus (Burgess, Maguire, & O'Keefe, 2002; Igloi, Doeller, Berthoz, Rondi-Reig, & Burgess, 2010). These formulations, however, make it difficult to explain how linguistic codes might flexibly modulate lower-level features like landmarks and spatial boundaries, both integral to the cognitive map (O'Keefe, 1991). For example, how is a largely verbal code in the left hippocampus translated into a purely spatial one on the right? While there may be ways of accommodating these theoretical ideas, there is not an obvious "rosetta" stone for doing so.

How do verbal cues affect spatial representations, if not via the cognitive map or survey representation? Our findings demonstrate the extent to which participants employ landmarks versus spatial geometry differs as a function of verbal cue. While a historical focus in the attention literature was on “bottom-up” versus “top-down” cues, one view is that the two are integrated in most ecological situations such that both salience and task-related goals can be differentially emphasised (Awh et al., 2012). Similarly, we suggest that during navigation (encoding), participants bind both the landmark and spatial boundaries. Then, based on verbal cues, they flexibly emphasise this information during the JRD task (retrieval). The emphasis on landmarks or geometry as a heuristic could occur at either encoding or retrieval, particularly because the two were interspersed in our paradigm, although our analysis of facing angles during navigation suggest that both landmarks and boundaries are learned then. For this reason, and the finding that schemas influence how encoded information is read-out (Moscovitch, Nadel, Winocur, Gilboa, & Rosenbaum, 2006), we favour a retrieval-related interpretation. Thus, just like schemas used to retrieve episodic memories, the heuristic used to readout memories for lower-level features encoded during navigation can differ depending on the verbal cue, with the heuristic exerting a powerful effect on memory for object positions. We note that the hybrid model is also consistent with this interpretation, with the caveat that boundaries tend to be more strongly encoded than landmarks, and thus create a stronger bias for their usage during retrieval than landmarks. Our data thus provide potentially new insight into how verbal cues can influence spatial representations via heuristics than previously appreciated, supporting models of spatial navigation that emphasise the importance of heuristics to coding a spatial layout (Ekstrom et al., 2017; Montello, 1998; Tversky, 1981).

We note that one study provided evidence for verbal cues via alignment effects in which participants were instructed to encode layouts either based on the current facing angle or surrounding axes (Mou & McNamara, 2002). The instructions resulted in a preference for the instructed component in which current facing angle was an intrinsic, egocentric cue and spatial boundaries were an extrinsic, allocentric cue. While they demonstrated that verbal cues resulted in participants favouring the extrinsic cue, our study involved two extrinsic cues, landmarks and boundaries, both that require encoding of extrinsic features used for allocentric navigation. Additionally, this study did not examine what participants viewed during encoding, leaving open the possibility that the results were due to attending to one feature over another. Because we showed that extrinsic cues can be flexibly modulated in a way that does not depend on simple attention-based explanations, our results suggest that such cues may be incorporated at earlier levels, in a more complex fashion, than previously appreciated.

Data availability

All data and materials are available by request from the corresponding author.

Acknowledgements

This research was supported by grants from the National Science Foundation (NSF BCS-1630296) and National Institutes of Health (NIH NS093052, NIH R01NS076856) to PI Ekstrom.

We thank Emily Erlenbach for assisting in data collection and Michael Starrett, Dr. Jared Stokes, Dr. Derek Huffman, Dr. Martha Forloines, Dr. Fernanda Ferreira, and Dr. John Henderson for their helpful suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by grants from the National Science Foundation (NSF BCS-1630296) and National Institutes of Health (NIH NS093052, NIH R01NS076856) to PI Ekstrom.

References

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. doi:10.1016/j.tics.2012.06.010

Barry, C., Lever, C., Hayman, R., Hartley, T., Burton, S., O’Keefe, J., … Burgess, N. (2006). The boundary vector cell model of place cell firing and spatial memory. *Reviews in the Neurosciences*, *17*(1-2), 71–97.

Berens, P. (2009). Circstat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software*, *31*(10), 1–21.

Bird, C. M., Capponi, C., King, J. A., Doeller, C. F., & Burgess, N. (2010). Establishing the boundaries: The hippocampal contribution to imagining scenes. *The Journal of Neuroscience*, *30*(35).

Brunye, T. T., Gagnon, S. A., Waller, D., Hodgson, E., Tower-Richardi, S., & Taylor, H. A. (2012). Up north and down south: Implicit associations between topography and cardinal direction. *Quarterly Journal of Experimental Psychology*, *65*(10), 1880–1894. doi:10.1080/17470218.2012.663393

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

Burgess, N., Maguire, E. A., & O’Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625–641.

Chan, E., Baumann, O., Bellgrove, M. A., & Mattingly, J. B. (2013). Reference frames in allocentric representations are invariant across static and active encoding. *Frontiers in Psychology*, *4*, 1–7.

Cheng, K. (1986). A purely geometric module in the rat’s spatial representation. *Cognition*, *23*, 149–178.

Cheng, K., & Newcombe, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin and Review*, *12*(1), 1–23.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, *58*(3), 306–324. doi:10.1016/j.neuron.2008.04.017

Eichenbaum, H., & Cohen, N. J. (2001). *From conditioning to conscious recollection: Memory systems of the brain*. New York: Oxford University Press.

Eichenbaum, H., & Cohen, N. J. (2014). Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron*, *83*(4), 764–770.

Ekstrom, A. D., Huffman, D. J., & Starrett, M. (2017). Interacting networks of brain regions underlie human spatial navigation: A review and novel synthesis of the literature. *Journal of Neurophysiology*, *118*(6), 3328–3344. jn. 00531.02017.

Ekstrom, A. D., & Isham, E. A. (2017). Human spatial navigation: Representations across dimensions and scales. *Current Opinion in Behavioral Sciences*, *17*, 84–89.

Ekstrom, A. D., Kahana, M. J., Caplan, J. B., Fields, T. A., Isham, E. A., Newman, E. L., & Fried, I. (2003). Cellular networks underlying human spatial navigation. *Nature*, *425*(6954), 184–188.

Ekstrom, A. D., & Ranganath, C. (2017). Space, time and episodic memory: The hippocampus is all over the cognitive map. *Hippocampus*, *27*, 9769. doi:10.1002/hipo.22750

Epstein, R. A., DeYoe, E. A., Press, D. Z., Rosen, A. C., & Kanwisher, N. (2001). Neuropsychological evidence for a topographical learning mechanism in parahippocampal cortex. *Cognitive Neuropsychology*, *18*(6), 481–508.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power program for the social, behavioral, and biomedical sciencesd. *Behavior Research Methods*, *39*(2), 175–191.

Fyhn, M., Molden, S., Witter, M. P., Moser, E. I., & Moser, M. B. (2004). Spatial representation in the entorhinal cortex. *Science*, *305*(5688), 1258–1264.

Garling, T., Book, A., & Lindberg, E. (1984). Cognitive mapping of large-scale environments the interrelationship of action plans, acquisition, and orientation. *Environment and Behavior*, *16*(1), 3–34.

Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: Dual-task studies of space and language. *Cognitive Psychology*, *39*(1), 3–36. doi:10.1006/cogp.1998.0713

Igloi, K., Doeller, C. F., Berthoz, A., Rondi-Reig, L., & Burgess, N. (2010). Lateralized human hippocampal activity predicts navigation based on sequence or place memory. *Proceedings of the National Academy of Sciences*, *107*(32), 14466–14471. doi:1004243107 [pii]. doi:10.1073/pnas.1004243107

Ishikawa, T., & Montello, D. R. (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, *52*(2), 93–129. doi:10.1016/J.Cogpsych.2005.08.003

Lee, P. U., & Tversky, B. (2001). *Costs of switching perspectives in route and survey descriptions*. Proceedings of the twenty-third conference of the cognitive science society.

Lynch, K. (1960). *The image of a city* (Vol. 11). Cambridge, MA: MIT Press.

McNamara, T. P., Rump, B., & Werner, S. (2003). Egocentric and geocentric frames of reference in memory of large-scale space. *Psychonomic Bulletin and Review*, *10*(3), 589–595.

Meilinger, T., Riecke, B. E., & Bülthoff, H. H. (2014). Local and global reference frames for environmental spaces. *Quarterly Journal of Experimental Psychology*, *67*(3), 542–569. doi:10.1080/17470218.2013.821145

Meilinger, T., Strickrodt, M., & Bülthoff, H. H. (2016). Qualitative differences in memory for vista and environmental spaces are caused by opaque borders, not movement or successive presentation. *Cognition*, *155*, 77–95.

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In M. J. Engenhofer, & R. G. Golledge (Eds.), *Spatial and temporal reasoning in geographic information systems* (pp. 143–154). New York: Oxford University Press.

Morey, R. D., & Rouder, J. N. (2015). Bayesfactor: Computation of Bayes factors for common designs. *R Package Version 0.9*, *9*.

Morris, R., Hagan, J., & Rawlins, J. (1986). Allocentric spatial learning by hippocampectomised rats: A further test of the "spatial mapping" and "working memory" theories of hippocampal function. *The Quarterly Journal of Experimental Psychology*, *38*(4), 365–395.

Moscovitch, M., Nadel, L., Winocur, G., Gilboa, A., & Rosenbaum, R. S. (2006). The cognitive neuroscience of remote episodic, semantic and spatial memory. *Current Opinion in Neurobiology*, *16*(2), 179–190. doi:10.1016/j.conb.2006.03.013

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

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

Mou, W., & Wang, L. (2015). Piloting and path integration within and across boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 220–234.

Munnich, E., Landau, B., & Dosher, B. A. (2001). Spatial language and spatial representation: A cross-linguistic comparison. *Cognition*, *81*, 171–208.

O'Keefe, J. (1991). An allocentric spatial model for the hippocampal cognitive map. *Hippocampus*, *1*(3), 230–235. doi:10.1002/hipo.450010303

O'Keefe, J., Burgess, N., Donnett, J. G., Jeffery, K. J., & Maguire, E. A. (1998). Place cells, navigational accuracy, and the human hippocampus. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, *353*(1373), 1333–1340. doi:10.1098/rstb.1998.0287

O'Keefe, J., & Dostrovsky, J. (1971). The Hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175.

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

Peacock, C. E., & Gözenman, F. (2017). Encoding-stage adaptation effects: Long-term memory. *Perception*, *47*(2), 216–224. doi:10.1177/0301006617739533

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(Feb), 3–25. doi:10.1080/00335558008248231

Ratliff, K. R., & Newcombe, N. S. (2008). Reorienting when cues conflict: Evidence for an adaptive-combination view. *Psychological Science*, *19*(12), 1301–1307.

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

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

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advancements in Child Development and Behavior*, *10*, 9–55.

Starrett, M. J., Stokes, J. D., Huffman, D. J., Ferrer, E., & Ekstrom, A. D. (2018). Learning-dependent evolution of spatial representations in large-scale virtual environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Taube, J. S. (2007). The head direction signal: Origins and sensory-motor integration. *Annual Review of Neuroscience*, *30*, 181–207. doi:10.1146/annurev.neuro.29.051605.112854

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, *31*, 261–292.

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology*, *13*(3), 407–433. doi:10.1016/0010-0285(81)90016-5

Waller, D., Montello, D. R., Richardson, A. E., & Hegarty, M. (2002). Orientation specificity and spatial updating of memories for layouts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1051–1063.

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

Weisberg, S. M., Badgio, D., & Chatterjee, A. (2018). Feel the way with a vibrotactile compass: Does a navigational aid aid navigation? *bioRxiv*. doi:10.1101/122994

Zhang, H., Zherdeva, K., & Ekstrom, A. D. (2014). Different "routes" to a cognitive map: Dissociable forms of spatial knowledge derived from route and cartographic map learning. *Memory & Cognition*, *42*(7), 1106–1117. doi:10.3758/s13421-014-0418-x