Psychonomic Bulletin & Review
https://doi.org/10.3758/s13423-023-02366-3

BRIEF REPORT

A common format for representing spatial location in visual and motor working memory

Sami R. Yousif[1] · **Alexander D. Forrence**[2] · **Samuel D. McDougle**[2]

Accepted: 11 August 2023

Abstract

Does the mind rely on similar systems of spatial representation for both perception and action? Here, we assessed the format of location representations in two simple spatial localization tasks. In one task, participants simply remembered the location of an item based solely on visual input. In another, participants remembered the location of a point in space based solely on kinesthetic input. Participants' recall errors were more consistent with the use of polar coordinates than Cartesian coordinates in both tasks. Moreover, measures of spatial bias and performance were correlated across modalities. In a subsequent study, we tested the flexibility with which people use polar coordinates to represent space; we show that the format in which the information is presented to participants influences how that information is encoded and the errors that are made as a result. We suggest that polar coordinates may be a common means of representing location information across visual and motor modalities, but that these representations are also flexible in form.

Keywords Spatial cognition · Spatial memory · Motor planning · Visual perception

Introduction

In animal minds and *in silico*, information is not stored indiscriminately; it must be organized – "formatted" – in some way. A classic debate in vision science, for instance, revolves around whether visual images in our minds are *depictive* (see Kosslyn, 1996; Kosslyn et al., 1995) or *propositional* (see Pylyshyn, 1973). In the domain of spatial cognition, there is debate about whether cognitive maps are *metric, Euclidean* in nature, or whether they are more "*graph-like*" (for discussions, see Peer et al., 2021; Warren et al., 2017; Yousif 2022). Then there are broader considerations about the form of mental representations in general, for example, whether they are *analog* or *digital* (see Maley, 2011, 2021). Here, we investigate a "case study" of representational format. We ask whether a common format (in this case, a common coordinate system) underlies remembered locations across visual and kinesthetic modalities.

Although there are an infinite number of ways of representing space, we typically think of two distinct formats: polar coordinates versus Cartesian coordinates. Each of these coordinate systems offers an efficient way of representing locations in two-dimensional space. Recent work has argued that the mind operates by default in polar coordinates, at least for visual representations of space: Yousif and Keil (2021b) used an "error correlation" analysis to show that, in most cases, errors between the dimensions of polar coordinates were uniquely *uncorrelated* whereas the dimensions of Cartesian coordinates were *correlated*. This is interpreted as evidence that polar coordinates are a likely candidate for the format of location representations (see Yousif & Keil, 2021b, for more information on the analysis; see also Yousif & Keil, 2021a).

Critically, these results further show that in some contexts (e.g., when the structure of the environment strongly implies a Cartesian grid), people may deploy Cartesian coordinates instead. This suggests that the mind operates spontaneously in one coordinate system but may occasionally operate in others depending on the demands of the environment. This is consistent with work arguing that multiple spatial formats may

---

✉ Sami R. Yousif
sryousif@sas.upenn.edu

[1] Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

[2] Department of Psychology, Yale University, New Haven, CT, USA

serve as the basis for representing different types of information (Hudson & Landy, 2012; Peer et al., 2021).

That humans might spontaneously rely on polar coordinates for representing location is consistent with a large body of prior work. For instance, Robinson (1972) argued that eye movements themselves may operate in polar coordinates. Huttenlocher and colleagues (1991) speculated about the use of polar coordinates for representing locations in memory, while more recently Yang and Flombaum (2018) provided some evidence that visual coordinates may operate in a polar reference frame. There is also evidence that voluntary motor actions are planned in polar coordinates (see, e.g., Baud-Bovy & Viviani, 2004; Flanders et al., 1992; Gordon et al., 1994; Krakauer et al., 2000; Messier & Kalaska, 1999). The notion of a polar format is perhaps compatible with work arguing that large-scale spatial representations are organized in a network-like or graph-like format (see, e.g., Kuipers, 1978, 1982; Warren et al., 2017) and seems even more directly compatible with work demonstrating that simple organisms like desert ants navigate home via a distance/direction vector (see, e.g., Müller & Wehner, 1988; Wittlinger et al., 2006). It has even been argued that young children are especially sensitive to distance and direction (the constituent dimensions of polar coordinates), more so than other Euclidean properties (Lee et al., 2012; but see Yousif & Lourenco, 2017). Combining this evidence, it has been proposed that location representations across domains and modalities may operate spontaneously in polar coordinates (Yousif & Keil, 2021b; see also Yousif, 2022).

While much work points towards polar coordinates as a likely candidate as the format of location representations, the evidence itself comes in many different forms. Work has relied on measures such as eye movements, reaching errors, pointing errors, various sorts of localization errors (e.g., replacing an item in a location, but also searching for discrete items in space), and more. Thus, it is challenging to compare format across modalities. A straightforward "error correlation" approach (Yousif & Keil 2021b) offers a simple way of assessing representational format across modalities and across paradigms. Here, we use this approach to evaluate localization errors in both visual and motor tasks.

In our first study, there were two distinct tasks: In one, participants visually localized objects on a computer screen; in the other, participants localized positions non-visually, based on kinesthetic information with the assistance of a robotic arm. In a second study, we more directly probed the format of remembered locations within the motor task.

Study 1

How is location information formatted in memory? Here, we addressed this question by analyzing the patterns of errors in two spatial localization memory tasks: A visual task and a reaching task. We were interested in whether the observed patterns of errors in both modalities were more consistent with the use of polar coordinates or with the use of Cartesian coordinates.

Method

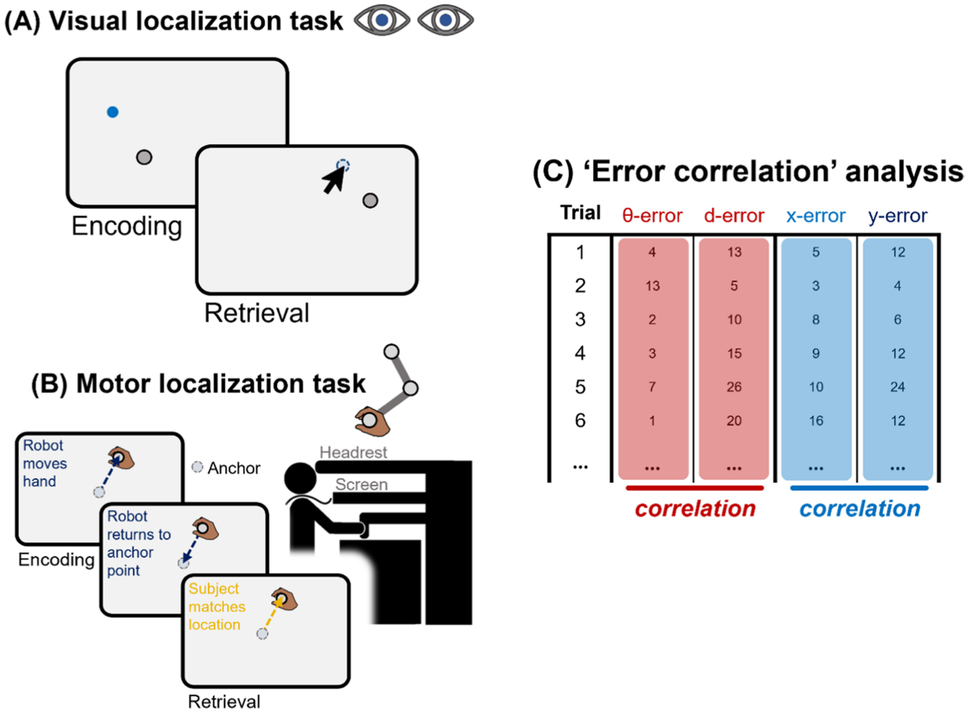
This study consisted of two separate tasks. One was a visual localization task in which participants saw dots briefly presented on a computer screen and then, after a delay, had to retrieve the location of that dot relative to a visible landmark. The other was a motor (kinesthetic) localization task in which participants were passively guided by a motorized robot from a "home" position to a location in space (with no visual input), were returned home by the robot, and then after a brief delay moved the robotic arm back to the remembered location. The pre-registration for this experiment as well as the subsequent experiment are available on the Open Science Framework (OSF) page at: https://osf.io/yeqbc/.

Participants

Forty undergraduate students participated in exchange for course credit. Half of the participants completed the visual localization task first and the other half completed the motor localization task first. Four additional participants were excluded prior to further data analysis based on predetermined exclusion criteria (three because of their responses during a debriefing survey; one because their overall accuracy was low).

Procedure and design

The visual localization task was modeled after the tasks used by Yousif and Keil (2021a, b). A simple depiction of the trial structure can be seen in Fig. 1A. The following measurements are all given in centimeters; this is so that the relative measurements for the visual and motor tasks can be more easily compared. Participants saw a blue target dot (.27 cm in diameter) presented in a random location relative to a central grey dot (.68 cm in diameter). The dots could not appear further than 3.26 cm away from the central grey dot, nor within .81 cm of the central grey dot. The dots would appear on the screen for 1,500 ms before disappearing. After another 500 ms, the grey dot would reappear in a different location and the blue dot would be absent. The participants were asked to place a new blue dot to match the location of the previous dot, relative to the current grey dot. The central grey dot would initially appear in one of the four quadrants (always 6.79 cm away from the center of the screen horizontally, and 4.07 cm away from the screen vertically); the grey dot would always reappear in the opposite quadrant from where it



Trial	θ-error	d-error	x-error	y-error
1	4	13	5	12
2	13	5	3	4
3	2	10	8	6
4	3	15	9	12
5	7	26	10	24
6	1	20	16	12
...	...	...	...	...

Fig. 1 (**A**) Depiction of the visual localization task and (**B**) the motor localization task. This is a schematic; stimuli are not to scale. (**C**) A depiction of the "error correlation" analysis used here, per Yousif and Keil (2021a, b)

had been initially. The initial position was counterbalanced so that the grey dot appeared in each quadrant an equal number of times. Once participants had clicked a single time, a blue dot would appear. However, participants could drag and drop or click additional times to replace the blue dot as they wished. They had an unlimited amount of time to respond, although they were encouraged to respond as quickly and as accurately as possible. To submit their responses, they pressed the spacebar. There were 120 trials in total. Participants completed two representative practice trials before beginning the task.

The motor localization task was designed to be as similar as possible to the localization task. A simple depiction of the trial structure can be seen in Fig. 1B. Participants sat at a desk in front of a robotic manipulandum (henceforth referred to as the "robot arm"; Kinarm End-Point, Ontario Canada). The robot arm could be moved by the participant, but it could also move autonomously (thus dragging the participants hand with it). Participants wore a black "bib" that obfuscated their vision of the robot arm and the desk itself. However, they were able to see visuals which displayed helpful information throughout the task (e.g., signals for when they could respond, start the next trial, etc.); these minimal stimuli/prompts were reflected from a horizontally mounted LCD screen onto a semi-silvered mirror positioned below it (the mirror provided further visual occlusion, thus making the full arm and hand completely invisible to participants).

Each trial began with a grey dot presented centrally on the screen. During this portion of the task only, there was a small cursor (a white dot) that corresponded to the location of the participants hand on the desk below. Participants were told to move the cursor onto the central home circle (a grey dot) to begin the trial. As soon as they did this, both the central grey dot and the cursor would disappear. At this time, the robot arm would move the participant's hand to a random location in the two-dimensional (2D) workspace. The random location could not be more than 7 cm away from the center in each x-y dimension (so that the maximum distance any point could be from the center was ~10 cm), and it had to be at least 3cm away from the center in at least one dimension. The robot arm would guide the participant's hand directly to the probe location on each trial (this passive movement was designed to always take 1,000 ms), pause for 1,000 ms for participants to commit the location to memory, then return the hand to the center. After another 500 ms, a green dot would appear on the screen at the central home location, which signaled to participants that they could respond. Participants were instructed to move immediately and directly to the point that had been indicated by the robot and hold their hand there. After the robot detected no significant movement (velocity <0.5 cm/s) for 500 ms, it would register the participant's current hand position as their response for that trial. At this point, the cursor and central grey dot would reappear, and the participant could control the cursor to return to the home location and begin the next trial.

Participants were explicitly told prior to the task that they should not rely on any special strategies or heuristics to localize the points in space. Instead, they were told to rely only on their sense of space and their memory, even if it meant they were slightly less accurate. This was done to prevent participants from surreptitiously using strategies like placing their arm against the table or pressing it against their body and trying to recreate how their arm had been positioned, rather than remembering extrinsic locations themselves. As with the visual localization task, participants completed 120 trials. They completed eight representative practice trials before beginning the task, during which they were given extensive verbal feedback (about the task itself, not their accuracy) to ensure that they understood the task.

Our main interest here was whether participants' errors are more consistent with the use of polar coordinates or Cartesian coordinates. To answer this question, we assessed the correlation between the errors in the constitutive dimensions of each coordinate system (for polar coordinates, angle/distance; for Cartesian coordinates, x/y), per the analysis used by Yousif and Keil (2021a, b). A simple visual explanation of this analysis is depicted in Fig. 1C.

Briefly, it can be shown mathematically (and demonstrated empirically) that errors generated by a system with independent Gaussian noise in the two polar dimensions will cause non-independence in the resulting x and y error components following unit conversion, leading to robust positive correlations between (unsigned) x and y error components. Analytic justification of this analysis is given in the Online Supplementary Material (OSM; see https://osf.io/yeqbc/), where we show that this is true not only in an idealized simulation, but also in a more human-like simulation (i.e., using noise parameters that resemble the human noise parameters we measured here). We further show that a distinct pattern of errors arises from a system with independent Gaussian noise in the two Cartesian dimensions. This fact allows us to make inferences about the underlying representational system used by human observers based on the patterns of errors we observe for each coordinate system in each task.

Results and discussion

First, we analyzed the accuracy in each task. In the visual localization task, participants erred by an average of .44 cm (SD = 4.27 cm); in the motor localization task, participants erred by an average of 1.5 cm (SD = 0.5 cm). Overall accuracy across tasks was significantly positively correlated $r(38) = 0.37$, $p = 0.019$ (see Fig. 2A), offering a first clue that shared spatial memory resources were deployed across modalities. We also calculated the average dispersion (aka "variable error"; Hancock et al., 1995) for each participant (i.e., the average distance between errors on each trial and the average error, or "centroid"). In the visual localization task, average dispersion was .2 cm; in the motor localization task, average dispersion was 0.6 cm. Average dispersion across tasks was also significantly positively correlated $r(38) = 0.41$, $p = 0.01$ (see Fig. 2B), again showing that performance, and perhaps spatial memory resources, were related across the two tasks.

Error correlations

Per the analysis plan outlined above (see Methods and OSM*;* see also Yousif & Keil, 2021b), we calculated the correlation between absolute error magnitudes of the dimensions of each coordinate system for each participant, resulting in a single correlation value for each individual (see Fig. 1C), to which we applied a Fisher Z-transformation (so that the values would be normally distributed). Then, we took the full sample of rho values and asked whether they differed from zero. For the visual task, memory errors computed in Cartesian coordinates showed consistent correlation ($M_r = .11$) between the constituent dimensions, $t(39) = 5.87$, $p < 0.001$, $d = 0.93$, whereas memory errors in polar coordinates ($M_r = .04$) did not, $t(39) = 1.49$, $p = 0.14$, $d = 0.24$ (see Fig. 2C). Strikingly, the same result held true for the motor task: errors for the dimensions of Cartesian coordinates were correlated ($M_r = .08$), $t(39) = 5.31$, $p < 0.001$, $d = 0.84$, whereas errors for the dimensions of polar coordinates were not ($M_r = -.02$) $t(39) = 1.93$, $p = 0.06$, $d = 0.31$ (see Fig. 2E). Note that what is relevant here is not the magnitude of the correlations, but the consistency of them across participants. Differences between the correlation values across coordinate frames were also significant in both cases, with higher correlations for Cartesian coordinates (visual: $t(39) = 6.04$, $p < 0.001$, $d = 0.96$; motor: $t(39) = 2.74$, $p = 0.009$, $d = 0.43$).

Lastly, 34/40 participants had a larger Cartesian correlation than polar correlation in the visual task (binomial test, $p < 0.001$; see Fig. 2D), and 30/40 participants had a larger Cartesian correlation than polar correlation in the motor task (binomial test, $p = 0.002$; see Fig. 2F). We note that all of the above *p*-values are prior to Bonferroni correction; given that there are four unique one-sample t-tests, the adjusted threshold for significance would be $p < 0.0125$. Thus, results that appear marginally significant should be interpreted with additional caution. In sum, that errors between the dimensions of Cartesian coordinates were correlated, and errors between the dimensions of polar coordinates were uncorrelated, suggests that location information in both spatial working memory and "motor working memory" may be formatted in polar coordinates.

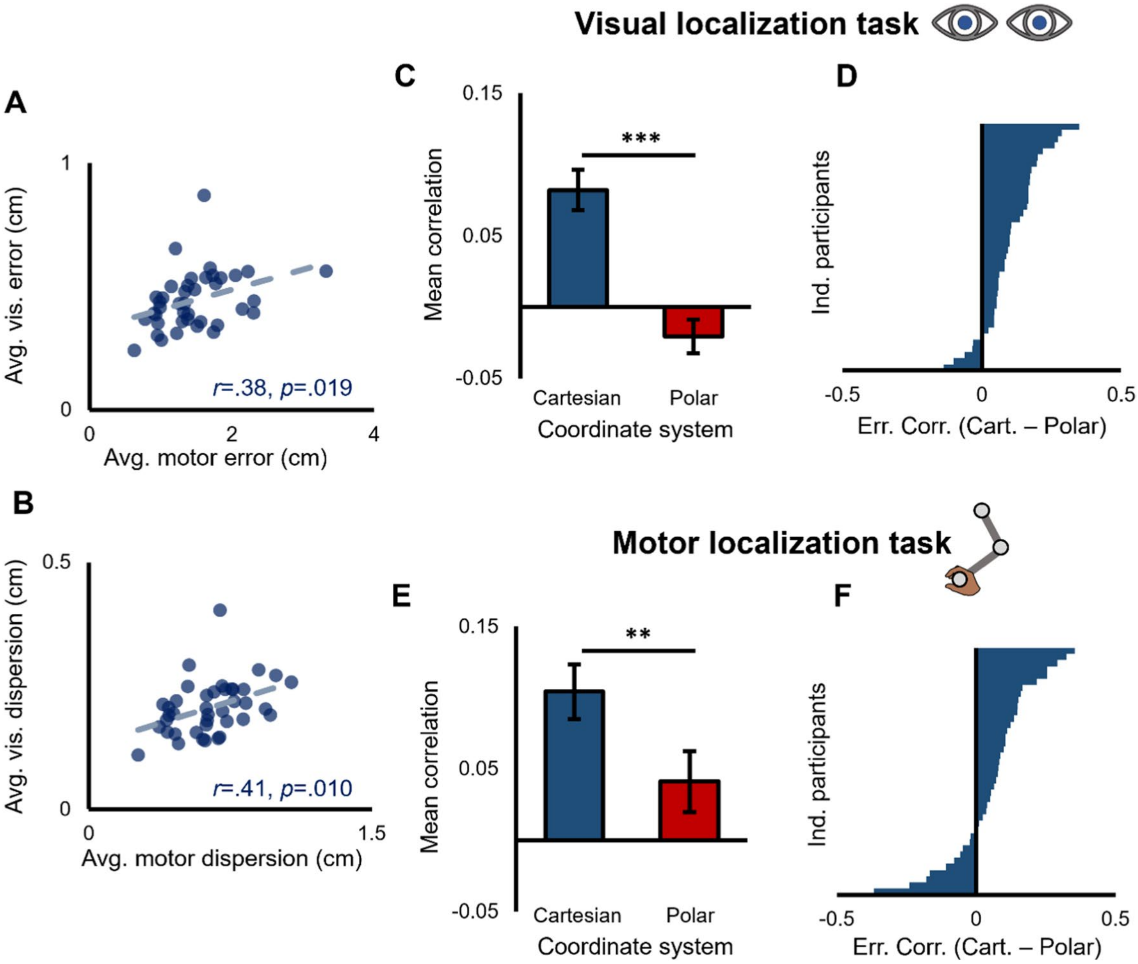


Fig. 2 (**A**) Cross task correlation for accuracy and (**B**) dispersion. (**C**) Error correlations for the visual localization task, collapsed across participants and (**D**) the difference in error correlations for each participant. (**E**) Error correlations for the motor localization task, collapsed across participants and (**F**) the difference in error correlations for each participant. Error bars represent +/- 1 SE

Oblique biases

As another metric of shared format across modalities, we also assessed bias towards the oblique regions of space. These results are displayed in Fig. 3.

There are many ways to quantify these biases. One simple metric is to count all the trials in which participants erred towards the oblique axis versus towards the cardinal axis. For the visual localization task, an average of 72% of trials ($SD = .07$) moved towards the oblique axes, $t(39) = 20.07$, $p < .001$, $d = 3.25$. For the motor localization task, an average of 59% of trials ($SD = .07$) moved towards the oblique axes, $t(39) = 8.14$, $p < .001$, $d = 1.29$. We can also quantify the magnitude of these biases: Are errors that move towards the oblique axes *larger* than errors that move towards the cardinal axes? For the visual localization task, the errors towards the oblique axes were an additional 3.91 degrees larger on average (points moving toward oblique: $M = 8.81°$, $SD = 2.29°$; points moving toward cardinal: $M = 4.91°$, $SD = 1.74°$; $t(39) = 14.71$ $p < .001$, $d = 2.33$). For the motor localization task, the errors toward the oblique axes were an additional 1.37° larger on average (points toward oblique: $M = 6.39°$, $SD = 1.23°$; points toward cardinal: $M = 5.02°$, $SD = 1.33°$; $t(39) = 6.66$, $p < .001$, $d = 1.05$). These analyses confirm what is evident from Fig. 3: Participants exhibited a robust tendency to err towards the oblique axes (akin to Huttenlocher et al., 1991; Yousif et al., 2020).

Separately, we quantified the magnitude of angular errors for points that originated near the cardinal axes versus those that originated near the oblique axes (unlike the previous analysis, which was based on where points erred towards, not where they originated). These results are shown in Fig. 3B. For the visual localization task, errors were on average 1.26° larger for points that originated near the cardinal axes, $t(39) = 4.40$, $p < .001$, $d = .70$; for the motor localization task, errors were on average 1.19° larger for points that originated near the cardinal axes, $t(39) = 6.19$ $p < .001$, $d = .98$. Combined with the previous analysis, these results

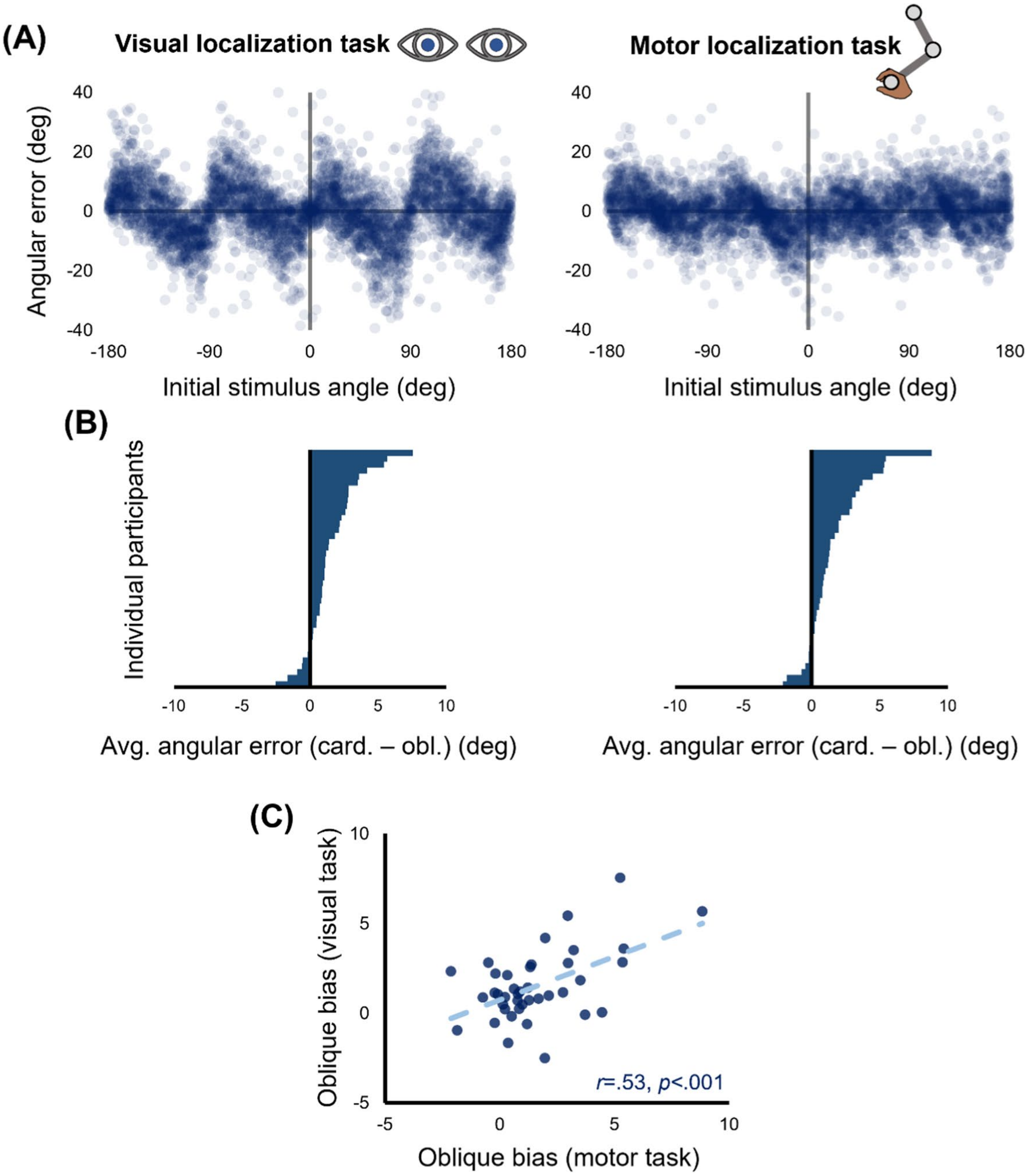


Fig. 3 Analysis of oblique biases in Experiment 1. (**A**) Angular error as a function of initial angular position. (**B**) Oblique biases, quantified as the difference in angular error for points that originated near the cardinal axes vs. the oblique axes. In other words, we took all the trials with points that originated closer to the cardinal axes and calculated the average absolute angular error for those points; then we did the same for all the trials with points that originated closer to the cardinal axes. The x-axis here reflects the difference between those two values, broken down by participant. (**C**) The correlation between the oblique biases

suggest that points originating near the cardinal axes (1) tend to move towards the oblique axes and (2) tend to move farther than points which had originated near the oblique axes.

Are the oblique biases we observed in each task related to one another? There was no correlation between the magnitude of errors that moved towards the oblique axes (Pearson's $r = .07$; Spearman's $r = .09$, $p = .55$). Crucially, however, there was a significant correlation between the magnitudes of errors that originated near one axis versus the other (see Fig. 3C; Pearson's $r = .53$, $p < .001$; Spearman's $r = .39$, $p = .014$).

Consider what it means to observe this correlation between these tasks: The values being correlated here are *differences* in angular accuracy between two different regions of space, in two different modalities and in two different spatial planes (vertical in the visual task, horizontal in the motor task). This means that participants that happen to make larger errors near the cardinal axes in a visual localization task also happen to make larger errors near the cardinal axes in a completely nonvisual motor localization task. This relation cannot be parsimoniously explained by purely visual or purely motor biases alone. It also cannot be easily explained by general inattention or inaccuracy, as there is no reason that errors due to attention or low effort should necessarily be localized to specific regions of space. Thus, these results are also indicative of a shared format underlying visual and motor spatial representations.

Study 2

Can location information be formatted in only one way? Previous work using the same approach as Study 1 in a visual localization task revealed that the use of polar coordinates is context dependent. That is, in environments that are more grid-like (perhaps implying a more Cartesian structure), patterns of error correlations flip (such that Cartesian errors become uncorrelated, and polar errors become correlated; see Yousif & Keil, 2021b). Might the same be true for information held in motor working memory?

Here we addressed this question by having participants complete a similar motor localization task as in Study 1, but with two trial types: "Direct" trials, in which participants are moved directly to the target location (directly replicating Study 1), and "Indirect" trials, in which participants are moved horizontally and then vertically to the target location, but still asked to report the remembered location by directly moving to it.

Method

This experiment was identical to the motor localization task of Study 1, except as stated below. Thirty undergraduate students participated in exchange for course credit. Three additional participants were excluded prior to data analysis based on pre-registered exclusion criteria (two because of their responses during debriefing; one because overall accuracy was low).

As in the previous task, there were 120 trials. However, unlike the previous task, these trials were of two types: "Direct" trials and "Indirect" trials. "Direct" trials were identical to those in the motor localization task in Study 1. "Indirect" trials were identical in every way, except that the robot arm would move participants first horizontally, then vertically, to the target point (before returning to center along the same path). For these trials, the arm always moved in two discrete movements, each 500 ms in duration. Direct and Indirect trial types were interleaved in a fully randomized fashion.

Results and discussion

First we analyzed the accuracy for each trial type. Participants erred by an average of 1.7 cm overall (SD = 0.5 cm). For the "Direct" trials, participants erred by an average of 1.6 cm (SD = 0.5 cm); for the "Indirect" trials, participants erred by an average of 1.8cm (SD = 0.5cm); this difference was statistically significant ($t(29) = 3.06$, $p = 0.005$, $d = 0.56$). That said, accuracy was also highly correlated across trial types, $r(28) = .86$, $p < .001$

As in Study 1, we assessed the correlation between the errors in the constitutive dimensions of each coordinate system. For the "Direct" trials, errors for the dimensions of Cartesian coordinates were correlated ($M_r = .10$), $t(29) = 3.92$, $p < 0.001$, $d = 0.72$, whereas errors for the dimensions of polar coordinates were not ($M_r = .03$), $t(29) = 1.05$, $p = 0.30$, $d = 0.19$ (see Fig. 4B). This difference was significant, $t(29) = 2.40$, $p = 0.02$, $d = 0.44$. Thus, we replicated the motor condition from Study 1. The difference between error correlations for Cartesian and polar dimensions for each participant is depicted in Fig. 4C.

For the "Indirect" trials, errors for the dimensions of Cartesian coordinates were uncorrelated ($M_r = .04$), $t(29) = 1.51$, $p = 0.14$, $d = 0.28$, as were errors for the dimensions of polar coordinates ($M_r = .05$), $t(29) = 1.50$, $p = 0.14$, $d = 0.27$ (see Fig. 4E). The difference between these two correlations was not significant, $t(29) = 0.16$, $p = 0.88$, $d = 0.03$. The difference between error correlations for Cartesian and polar dimensions for each participant is depicted in Fig. 4F.

The lack of correlations for both coordinate types is difficult to interpret, and, here, the details of the math are important. If we assume that participants rely on polar coordinates, the predicted pattern of error correlations is straightforward. Cartesian errors should be correlated and polar errors should be uncorrelated. However, if we assume that participants rely on Cartesian coordinates, the predicted pattern of error correlations is less clear. It depends on the details of the task, or the precise way in which participants err. Such patterns of errors are less reliable in simulation (see OSM; https://osf.io/yeqbc/), and humans tend to err in systematic ways that are hard to capture via simulation (e.g., as in the oblique biases discussed in Experiment 1; see also Huttenlocher et al., 1991; Yousif et al., 2020).

Still, simulations of cartesian agents with similar noise parameters as those measured here reveal the exact pattern

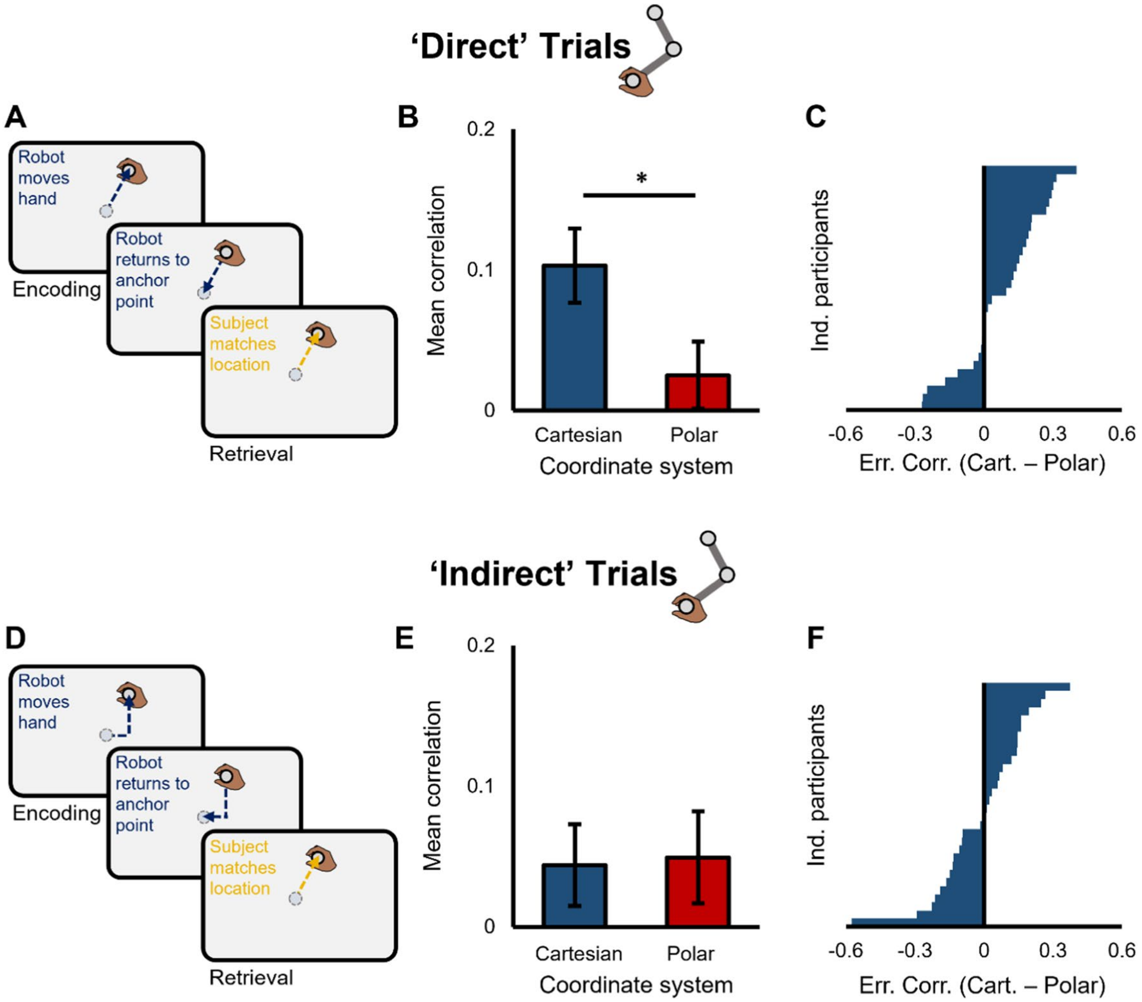


Fig. 4 (**A**) Depiction of a "Direct" trial. (**B**) Error correlations for the "Direct" trials, collapsed across participants and (**C**) the difference in error correlations for each participant. (**D**) Depiction of an "Indirect" trial. (**E**) Error correlations for the "Indirect" trials, collapsed across participants and (**F**) the difference in error correlations for each participant. Error bars represent +/- 1 SE

we see for the "Indirect' trials: a lack of correlation for both coordinate systems (see OSM). This could mean that participants are relying on Cartesian coordinates some but not all the time, or it could mean that participants are relying on some strategy or coordinate system that cannot be detected by our analyses. Therefore, while the Indirect condition supported our general predictions, these results should not be over-interpreted.

Another possibility was that participants were still representing space in polar coordinates, but relative to a different point in space. To address this, we conducted the same error correlations as before, but considered the intermediate point (i.e., the point along the x-axis that would be the corner of the right triangle formed by the three points) as the origin of the polar system (rather than the central home position). Error correlations for Cartesian coordinates are unchanged by this difference. Error correlations for polar coordinates, however, came out higher than before and significantly greater than zero, for both trial types ("Direct": $t(29) = 4.76$, $p < .001$, $d = 0.87$; "Indirect": $t(29) = 5.55$, $p < .001$, $d = 1.01$). Thus, it is unlikely that spatial memory on the Indirect trials used a polar format with the "way point" as a reference.

We see two possible ways of further interpreting the results of Study 2. One is that location information in motor working memory may be stored in multiple formats, depending on the demands of the task. The strongest support for this conclusion comes from the significant interaction we observed between the error correlations and trial types. The other way to interpret these results is that polar coordinates may be privileged in some way. This is evidenced by the fact that even when participants are moved indirectly to a location, they still exhibit low (and insignificant) error correlations for the polar dimensions.

The analytic approach taken here does not allow us to fully disentangle these possibilities. However, it was clear that the encoding context affected the spatial layout of localization recall errors.

General discussion

Here, we have proposed that representations of location encoded into memory visually or kinesthetically may operate in a common polar format. In Study 1, in both a visual and a non-visual kinesthetic location memory task, participants errors for the dimensions of polar coordinates were uncorrelated while their errors for the dimensions of Cartesian coordinates were correlated. This result supports an underlying polar representational format with independent noise sources for the angle and extent dimensions (see OSM; https://osf.io/yeqbc/). Moreover, oblique biases were correlated across tasks, further supporting a shared representation. In Study 2, when participants encoded location information in a form that was more indicative of Cartesian coordinates, the patterns of errors were distinct. Collectively, these results suggest that while people may spontaneously rely on polar coordinates in many cases, they may not rely solely on polar coordinates (especially if the environment/task implies another structure).

Implementation, computation, and "format"

Marr (1982) famously described different "levels" of analysis for cognitive systems. To bridge the gap between brain (implementation) and behavior (computation), then, one must understand the intermediate *representational formats* between them. The goal of this work is to demonstrate a format that is used across cognitive systems (e.g., action/perception) that may serve as a common means of spatial representation. Only by understanding format at the algorithmic level of computation would it be possible to understand how the building blocks of spatial cognition like place cells (O'Keefe & Dostrovsky, 1971), grid cells (Hafting et al., 2005) and head-direction cells (see Taube, 1998) combine to results in complex spatial representations.

So: What are the "formats" of spatial representation? How many are there? It is possible in principle that there exist many unique forms of location representation(s) in the mind – that there are separate systems devoted to spatial representation for perception and spatial representation for action. Indeed, popular models of the visual system describe two distinct pathways or "streams": one for perception, and one for action (Goodale & Milner, 1992; Mishkin et al., 1983). At times, however, information from perception is relevant for action (or vice versa). When catching a baseball, for instance, one must translate location information received through vision into a motor action. In such cases, it may be useful to re-code spatial information into a common format of representation, so that it may be readily translated from one system to another. The use of a general spatial code would be especially useful when memory is required, as in our tasks – a remembered location may need to be acted upon in multiple ways.

The work here has focused mainly on a contrast between polar and Cartesian coordinates. However, it is possible in principle that the mind represents locations without the use of a coordinate system at all (via "coarse" or "propositional" representations of space, for instance; see Huttenlocher et al., 1991; Pylyshyn, 1973; for further discussion, see Yousif, 2022). Thus, these results not only address a specific question about *what* coordinate system is typically employed by the mind, but a more general question about *how* the mind represents space in general – sometimes, at least, by way of a coordinate representation.

Relation to prior work

These findings are consistent with prior work showing that both visual (Huttenlocher et al., 1991; Yang & Flombaum, 2018; Yousif & Keil, 2021a, b) and non-visual (e.g., Baud-Bovy & Viviani, 2004; Gordon et al., 1994; Flanders et al., 1992; Messier & Kalaska, 1999) representations of location operate in polar coordinates. The primary limitation of prior work is that there is no agreed-upon way of evaluating representational format. Some work has focused on qualitative error patterns (e.g., Huttenlocher et al., 1991; Gordon et al., 1994; Messier & Kalaska, 1999; Yousif et al., 2020); other work has relied upon pointing errors (e.g., Warren et al., 2017); and other work has compared responses to stimuli with more polar- or Cartesian-esque properties (Yang & Flombaum, 2018).

What the current work offers is a demonstration of a common format that is dependent on a single paradigm (i.e., a location memory task) and a single analysis (i.e., 'error correlations'). Rather than relying on unique paradigms and unique dependent variables, the error correlation approach depends only on errors made in simple tasks. In the same way that we have extended this approach beyond the visual modality, future work could use the same approach to study spatial representation at the scale of the natural environment (e.g., in a real-world localization task).

One potential way of understanding the present results is in terms of "cognitive maps" (see Tolman, 1948). Some have argued that cognitive maps are roughly Euclidean (e.g., Gallistel, 1990; O'Keefe & Nadel, 1978); others have argued that cognitive maps are more network- or graph-like (Kuipers, 1978, 1982; Warren et al., 2017). Yet others have argued that location is represented in a relational, hierarchical, or categorical manner (Huttenlocher et al., 1991, 1994; Jiang et al., 2000; McNamara et al., 1989; Taylor & Tversky, 1992), perhaps relying on propositional knowledge (Pylyshyn, 1973). Interestingly, theories regarding the nature of cognitive maps often do not emphasize specific coordinate systems, but instead focus on the nature of the spatial representation itself (e.g., whether it is metric or non-metric). The present results may thus speak to an important aspect of cognitive maps. Specifically, the data here hint at the possibility that (1) there are indeed cognitive maps that are not tied to any one sensory modality, and (2) these maps may operate in a common format, or coordinate system.

Conclusion

Here, we analyzed errors made by participants in two spatial memory tasks and showed that, in both the visual and motor modality, small spatial memory errors may reveal the underlying *format* of location representations. We have argued that spatial cognitive representations serving both perception and action may partly depend on a common format: polar coordinates. Understanding the formats of spatial representations at this algorithmic level can ultimately help to bridge gaps between brain and behavior.

Data availability Both experiments were pre-registered. Those pre-registrations as well as the raw data and analyses can be found via the Open Science Framework at: https://osf.io/yeqbc/.

References

Baud-Bovy, G., & Viviani, P. (2004). Amplitude and direction errors in kinesthetic pointing. *Experimental Brain Research, 157*, 197–214.

Flanders, M., Tillery, S. I. H., & Soechting, J. F. (1992). Early stages in a sensorimotor transformation. *Behavioral and Brain Sciences, 15*, 309–320.

Gallistel, C. R. (1990). *The organization of learning*. MIT Press.

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

Gordon, J., Ghilardi, M. F., Cooper, S. E., & Ghez, C. (1994). Accuracy of planar reaching movements. *Experimental Brain Research, 99*, 112–130.

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

Hancock, G. R., Butler, M. S., & Fischman, M. G. (1995). On the problem of two-dimensional error scores: Measures and analyses of accuracy, bias, and consistency. *Journal of Motor Behavior, 27*, 241–250.

Hudson, T. E., & Landy, M. S. (2012). Motor learning reveals the existence of multiple codes for movement planning. *Journal of Neurophysiology, 108*, 2708–2716.

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

Huttenlocher, J., Newcombe, N., & Sandberg, E. H. (1994). The coding of spatial location in young children. *Cognitive Psychology, 27*, 115–147.

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 683–702.

Kosslyn, S. M. (1996). *Image and brain: The resolution of the imagery debate*. MIT Press.

Kosslyn, S. M., Thompson, W. L., Klm, I. J., & Alpert, N. M. (1995). Topographical representations of mental images in primary visual cortex. *Nature, 378*, 496–498.

Kuipers, B. (1978). Modeling spatial knowledge. *Cognitive Science, 2*, 129–153.

Kuipers, B. (1982). The "map in the head" metaphor. *Environment and Behavior, 14*, 202–220.

Krakauer, J. W., Pine, Z. M., Ghilardi, M. F., & Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *Journal of Neuroscience, 20*, 8916–8924.

Lee, S. A., Sovrano, V. A., & Spelke, E. S. (2012). Navigation as a source of geometric knowledge: Young children's use of length, angle, distance, and direction in a reorientation task. *Cognition, 123*, 144–161.

Maley, C. J. (2011). Analog and digital, continuous and discrete. *Philosophical Studies, 155*, 117–131.

Maley, C. J. (2021). The physicality of representation. *Synthese, 199*, 14725–14750.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Freeman.

McNamara, T. P., Hardy, J. K., & Hirtle, S. C. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 211–227.

Messier, J., & Kalaska, J. F. (1999). Comparison of variability of initial kinematics and endpoints of reaching movements. *Experimental Brain Research, 125*, 139–152.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences, 6*, 414–417.

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

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*, 171–175.

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

Peer, M., Brunec, I. K., Newcombe, N. S., & Epstein, R. A. (2021). Structuring knowledge with cognitive maps and cognitive graphs. *Trends in Cognitive Sciences, 25*, 37–54.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin, 80*, 1–24.

Robinson, D. A. (1972). Eye movements evoked by collicular stimulation in the alert monkey. *Vision Research, 12*, 1795–1808.

Taube, J. S. (1998). Head direction cells and the neurophysiological basis for a sense of direction. *Progress in Neurobiology, 55*, 225–256.

Taylor, H. A., & Tversky, B. (1992). Descriptions and depictions of environments. *Memory & Cognition, 20*, 483–496.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55*, 189–208.

Warren, W. H., Rothman, D. B., Schnapp, B. H., & Ericson, J. D. (2017). Wormholes in virtual space: From cognitive maps to cognitive graphs. *Cognition, 166*, 152–163.

Wittlinger, M., Wehner, R., & Wolf, H. (2006). The ant odometer: Stepping on stilts and stumps. *Science, 312*, 1965–1967.

Yang, F., & Flombaum, J. (2018). Polar coordinates as the format of spatial representation in visual perception. *Journal of Vision, 18*, Article 21. https://doi.org/10.1167/18.10.21

Yousif, S. R. (2022). Redundancy and reducibility in the formats of spatial representations. *Perspectives on Psychological Science*, *17*, 1778–1793.

Yousif, S. R., Chen, Y. C., & Scholl, B. J. (2020). Systematic angular biases in the representation of visual space. *Attention, Perception, & Psychophysics, 82*, 3124–3143.

Yousif, S. R, & Keil, F. (2021a). 'Decoding' the locus of spatial representation from simple localization errors. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*. Retrieved from https://escholarship.org/uc/item/185176jj

Yousif, S. R., & Keil, F. C. (2021b). The shape of space: Evidence for spontaneous but flexible use of polar coordinates in visuospatial representations. *Psychological Science, 32*, 573–586.

Yousif, S. R., & Lourenco, S. F. (2017). Are all geometric cues created equal? Children's use of distance and length for reorientation. *Cognitive Development, 43*, 159–169.