Spring 2022

# The Formats of Spatial Representations

Sami Yousif
*Yale University Graduate School of Arts and Sciences*, sami.yousif@yale.edu

Abstract

The Formats of Spatial Representations

Sami Ryan Yousif

2022

Mental representations are the essence of cognition. Yet, to understand how the mind works, we must understand not just the *content* of mental representations (i.e., what information is stored), but also the *format* of those representations (i.e., how that information is stored). If we want to understand how sensory information is translated into symbolic representations, if we want to know how the mind forms 'cognitive maps', if we want to know how the firing of neurons can lead to the emergent phenomenon of human cognition — all of these things require us to understand how information is organized in the mind.

In this thesis, I describe three 'case studies' of representational format in the domain of spatial cognition. I focus on spatial cognition for several reasons. First, spatial cognition is ubiquitous in the animal kingdom; thus, understanding spatial cognition in the human mind has the potential to reveal insights that generalize to *all* minds. Second, spatial cognition may be the single domain for which we know the most about the format of representations; indeed, the field was essentially founded on the premise that there exists a discernable 'cognitive map' within the mind. As such, it serves as an apt domain to study representational format. Finally, spatial representations (location representations in particular) may serve as the format of other higher-level information (e.g., numerical information, social

information, etc.). Understanding the formats of spatial representation, therefore, may shed light on how other kinds of information are represented and organized in the mind.

The first case study I describe pertains to the format of location representations. I show that, using a simple 'error correlation' analysis, we can uncover from simple spatial tasks the coordinate systems underlying spatial behavior. Using this approach, I argue that locations are spontaneously represented in polar coordinates, but flexibly in other coordinate systems (e.g., Cartesian coordinates) as needed.

The second case study I describe pertains to the format of size representations. It has been known for many decades that the perception of size is illusory; for example, larger objects are perceived as being relatively less large. However, these illusions are typically explained by vague, unfalsifiable theories of size perception. I offer a simpler (and falsifiable) explanation of size illusions: that perceived size is equal to the *sum* of an objects' dimensions rather than the *product*. Here, I focus primarily on the perception of area in adults, but this phenomenon appears to be highly general: I briefly allude to similar illusions that children experience, as well as similar illusions of volume.

The final case study I describe pertains to how spatial information is used as a format to represent other information. I show that task-irrelevant 'spatial structure' spontaneously improves working memory. This effect is specific to spatial information; color information and audio information produce no such benefit. I discuss how these findings relate to existing models of working memory, and help us to understand the relationship between space and memory more broadly.

I conclude with some final remarks about how understanding spatial behavior in light of the formats of representations can help us to understand the building blocks of cognition.

# The Formats of Spatial Representations

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy of the Degree of
Doctor of Philosophy

by
Sami Ryan Yousif
Dissertation Director: Frank C. Keil
Dissertation Chair: Samuel M. McDougle
May, 2022

# Contents

# List of Figures

# Acknowledgments

How is one meant to acknowledge all the people that have contributed to a dissertation? There are far too many people to thank; there is far too much to say.

At Emory University, two people shaped my young adult life. The first is Stella Lourenco. Stella took me on as a research assistant in the Spring semester of my sophomore year. With less than six months experience in her lab, she trusted me to act as her lab manager. That job saved my life. It was offered to me at a time when I was jobless and homeless. As I was forced to support myself through college, that job was the only way I could pursue my interests and stay in college at the same time. I couldn't have asked for better luck. For the next two years, Stella mentored me — an overconfident, overcurious kid who had no idea what they were talking about — as if I was a graduate student. She pushed me and respected me and gave me the flexibility I needed to finish my college degree. She trusted me when I needed her trust.

Here I must also thank Vlad Ayzenberg. Although Vlad and I butted heads over ideas, he was a phenomenal friend. He continued to be a good friend even when I wasn't good back. When he found out I was homeless, he dropped everything and tried to help. And on this note, let me also quickly thank Lauren Aulet, who *wasn't* in a position to impact my life so dramatically, but would have if she had been. I thank Lauren for our many days of walking around Europe talking for 14 hours non-stop; those aimless conversations came at a formative time in my academic career, and I grew from them.

The second person who shaped my life at Emory is Mahlet Endale. Mahlet acted as a role model to me. As the supervisor of the Emory Helpline, she acted

with grace and composure and wisdom every single day. I thought then, and think now, that Mahlet is the greatest leader I have ever met. I wish society would teach young people what I learned from Mahlet: that vulnerability is not mutually exclusive with leadership.

At Yale, I have a great number of people to thank. Let me first thank my two cohort-mates that have been through this whole journey with me: Rosie Aboody and Qi Lin. It has been such an honor to work with both of you. You have both challenged me and made me better. Let me also thank the members of my first lab here at Yale. Chaz, thank you for all the long conversations and the ceviche. Thank you for supporting me in spite of, you know, *everything.* Also, thank you for guilting me into learning Javascript five years ago; that decision significantly altered the course of my career. Ben, thank you for being one of the kindest people I have ever met. Thank you for entertaining all my dumb questions when we were alone in the lab on weekend mornings. Stefan, thank you for many, many laughs — and for coming up with the title of my first paper. Clara, thank you for the thousands of inside jokes, for always remembering every little detail of my work, and for being a good friend on some really tough days. Although we only overlapped in the same lab for a brief time, the term 'lab mate' will always be synonymous with you, Clara.

Let me also thank the massive cast of other colleagues who played a part in my journey: Angie Johnston, Mark Sheskin, Sam Johnson, Matt Fischer, April Bailey, Richard Ahl, Cameron Ellis, Tristan Yates, Kat Graves, and Alice Liefgreen. Thanks also to others who helped keep me sane, especially during the pandemic: Natalia Cordova, Henry Mattingly, Paul Young, and Nour Malek. Thanks to my current lab mates: Nicole Betz, Emory Richardson, Mandy McCarthy, Flora Zhang, and Maureen Gill. Thanks especially to Flora, for being such a great friend to me and such a good citizen of this department.

Let me also thank the many faculty members here and elsewhere who have 'advised' me in some capacity throughout graduate school: Brian Scholl, Frank Keil, Nora Newcombe, Darko Odic, Chaz Firestone, Monica Rosenberg, Justin Halberda, Dick Aslin, Joshua Knobe, Dave Lagnado, Julian Jara-Ettinger, Sam McDougle, Stella Lourenco, Daniel Dilks, and Julia Leonard. I thank Brian Scholl for admitting me to Yale in the first place. Despite our many differences, I learned a lot. I thank Frank for taking a chance on me — for believing in me, for trusting me, and for letting me forge my own path. I thank Frank for allowing me to cry in his office one hundred times. And I thank Frank for a million other things (truly, a million) that will not fit in this acknowledgments section. I thank Dick for his courage and his care and the constant supply of relevant anecdotes. I thank Josh for being the perfect role model for what genuine kindness and selflessness can look like in academia.

Let me thank the students who I have had the pleasure of working with here at Yale: Emma Alexandrov, Elizabeth Bennette, Sarah Joo, Caroline Reiner, and Serena Shapard. Thanks to all of you for being patient with me as I have figured out how to be a mentor. Thanks to all of you for being devoted to the work that you did. Thanks to all of you that have left for staying in touch. Thanks especially to Sarah Joo. Her first class as an undergraduate coincided with mine as a graduate student. Due to some rather extraordinary circumstances (i.e., COVID), after working together for almost five (!) years, we will get to graduate at the same time this Spring. What an honor that will be. By most metrics, you are the person with whom I have collaborated the most. I never would've guessed that my number one collaborator would be a philosophy major, but, hey — you're going to be a psychologist now, huh? Thank you for believing in me as much as I believe in you.

Let me thank my family for *kind of* supporting me, a little. Look, I know some of you had other aspirations for me. Oh well. Thanks mostly to my mother, Wendy, who had enough patience to raise me until I made it to college. I remember those years when it was often just you and I at home during some really tough times. You did an incredibly courageous thing and pursued a very difficult path so that I could have a chance at a more peaceful life. If you had not done another thing for me in all my life, I would still owe you everything. Those years of relative peace in the home that *you* built were the foundation of the rest of my life. I needed that peace to become myself. Thank you.

Thanks also to Alene and Jim Sherman for your patience, for your kindness, for your support, and for your love. I am lucky to be a part of your family. Thanks to Hannah Sherman for thinking about me when I crashed the car and for, against all odds, helping me to appreciate reality TV. Thanks to Arielle and Mark Feldman for the endless entertainment provided by Norah and Lily, even from afar. Thanks also to Arielle for giving me the best gift I have ever received in my life (it is proudly displayed in my office, right next to me, as I am writing this).

And thanks to Brynn Sherman, for holding my entire life together. I could write entire theses expressing my gratitude for you, Brynn. (Honestly, that's a good idea; I should do that.) I owe so much thanks to so many people for making it here, but I owe a thousand times as much to you alone. Thank you for trusting me, believing in me, and making me better. I want to say so much more, but, just in case someone actually reads this at some point, I don't want to embarrass you further. You're in luck: I can save these thoughts for wedding vows — soon.

# Chapter 1

# Introduction

This chapter contains text and/or materials from the following publications:

Yousif, S. R. (In press). Redundancy and reducibility in the formats of spatial representations. *Perspectives on Psychological Science*.

The central goal of psychology, cognitive science, and cognitive neuroscience is to understand how complex thought and behavior arise from neural tissue. To unite these sometimes-siloed areas of study, we canonically think of three different *levels of analysis* at which we can understand the mind and brain: the computational level, the algorithmic level, and the implementation level (e.g., Marr, 1982). Psychology is often interested in describing the computational level — understanding what the mind is trying to compute in the first place. Cognitive neuroscience, in contrast, is interested more in the implementation level — understanding how computations can be implemented in neural tissue. This thesis addresses the intermediate algorithmic level — how spatial representations for properties like location and size are *formatted* in the human mind. Here, I describe several case studies on spatial behavior that provide insight into the underlying format and nature of spatial representations.

## 1.1 What is 'format'?

For my purposes, to *represent* a piece of information means to store it symbolically, such that it can be accessed, retrieved, and updated (for extended discussions, see Brooks, 1991; Markman & Dietrich, 2000; Shea, 2018; Thelen & Smith, 1996). However, representations are not just shapeless bits of information stored in arbitrary units. Representations contain *content*, and that content must be *formatted* in some way.

This notion of format should feel familiar, given that it shapes our daily digital interactions. For example, any given document could be formatted as a .doc file, or a .pdf file, or a .tex file. Each of these formats has advantages and disadvantages: A .doc file may be better for editing, while a .pdf file may be better for standardized presentation. Ultimately, the entire computer architecture depends on the simple fact that certain programs process specific inputs and produce specific outputs. Those inputs and outputs are *representations*, and the operations that can be performed on those representations are constrained by their *format*.

Another example of format is shown in Figure 1.1. Figure 1.1A compares the classic game of tic-tac-toe with the “fifteen game”. In the fifteen game, there are 9 discs, each representing a digit 1-9. Players take turn selecting one disc from the set. The goal is to end up with a set of exactly three discs that adds up to 15. The first player to achieve that goal wins. If neither player achieves that goal, it is a draw. (I’ll assume readers are familiar with tic-tac-toe.) On their surfaces, these games seem entirely different: one involves a grid with x’s and o’s, and the other involves summing numbers to fifteen. Yet when those numbers are superimposed on the tic-tac-toe grid in a certain order (see Figure 1.1A), you can see their similarity. Just like tic-tac-toe, the “fifteen game” contains a finite number of solutions. Just like tic-tac-toe, there are exactly nine possible moves, and some are better than others: picking both “5” and the central square in the grid results in the maximum number of winning combinations (where the numbers “3”, “1”, “7”, and “2”, like the left, right, top, and bottom squares result in the fewer number of winning combinations). And, just like tic-tac-toe, participants take turns making selections, and a ‘winning’ combination requires exactly three selections. The difference between these two games could be understood as a difference in format: tic-tac-toe traffics in x’s and o’s, whereas the fifteen game traffics in numbered discs. Functionally, the inputs and outputs of the two games are equivalent, yet their different formats constrain how we interact with the information.

The history of cognitive science has been shaped by questions of format. Many of the field’s most prominent debates have centered around this exact issue. For example: is imagery depictive (see Kosslyn et al., 1995; Kosslyn, 1996) or propositional (see Pylyshyn, 1973; Pylyshyn, 2002)? What are the dimensions of 'face space' underlying face perception (e.g., Chang & Tsao, 2017)? What latent structures form the basis of human language (e.g., Traxler & Gernsbacher, 2011)? In each of these cases, understanding the formats of these representations helps us understand how the mind solves these critical problems (of imagery, face perception, and language) in the first place.

A full account of 'format' will not be provided in this thesis. Here, I will simply discuss 'format' in the broadest possible sense, as any kind of organization of information. We'll also keep in mind Marr's three levels of analysis as a reference point (Marr, 1982; see also Maley, 2021 for a useful and thoughtful discussion). The highest level, the *computational level*, describes what is being represented. The lowest level, the *implementation level*, describes the physical substrate on which that computation is implemented (e.g., in animals, neurons). And at every step in between the computational level and the implementation level — all the intermediate levels that may be collectively referred to as the *representational/algorithmic level* — information must be organized in some way. That organization, whatever it may be, is 'format'.

## 1.2 One format or many?

It may be tempting to think that all representations must be reducible to a single, ultimate format — that, to format information in more than one way would be redundant. However, this need not be the case. Indeed, one of the themes discussed throughout my work is the possibility that representations are formatted redundantly. Let's briefly consider a few reasons in principle why 'redundant' formats may be useful.

First, think of the number 'seven'. In base-10, the quantity 'seven' can be represented by the digit 7. However, in base-2, the quantity 'seven' is represented by the digits 111. In base-3, by the digits 21. We can think of these different bases as formats. It isn't as though one base or another captures the quantity 'seven' any more precisely; they just represent the quantity in different ways. Each format may have its own merits: formatting information in base-10 is intuitive to most people, but formatting information in base-2 is a necessary feature of modern computers.

Second, think about the game of chess. To most people, a chess board is nothing but an 8 by 8 grid. If you want to make a move — or explain a move — you must refer to that grid. Indeed, this is the most natural way to *see* and *play* the game of chess. But it is not the most natural way to *talk* about the game of chess. Experienced chess players and commentators use specific notation (in modern chess, 'algebraic notation') to talk about the movement of pieces (see Figure 1.1C). For example, if you have the white pieces and you want to move the pawn in front of your king two spaces forward, you could explain that in words, or you could simply say "e4". Using this notation, it requires not even one full word to convey the same move that was just described using 23 words. The algebraic notation is useful: it simplifies communication about the game of chess. That said, it would be hard to learn the game of chess from the algebraic format alone. The (beautiful) geometry of the game is most apparent when viewed as an 8 by 8 grid.

This is to say that both formats serve a different function; mastering chess requires understanding both formats. However, neither format is strictly necessary. The 'Stockfish' chess engine, for example, need not understand the game of chess as a grid at all. It must only represent the statistical value of a given move to a given tile, regardless of where that tile is relative to the other tiles. Conversely, most casual players of the game never interact with the algebraic notation. But for veteran players of the game of chess, there is value in being able to translate seamlessly between formats as needed.

This is why ‘redundant’ formats may be useful: Human minds are not machines designed to compute specific tasks optimally. The reason we have algebraic notation for chess in the first place is because it enables us to explain and describe the game (succinctly) in written form. Yet, it is not as if the development of chess notation forced chess players to abandon the board altogether; we still enjoy and appreciate the game in its physical, two-dimensional form. Unlike chess engines, it benefits the human mind to represent the chess board in multiple ways. Whereas simpler organisms (and chess engines) may benefit from having one highly specialized system for each of its (few) behaviors, more complex organisms may be better served by cognitive systems adapted to complete innumerable tasks flexibly. As such, it may benefit certain minds (in this case, chess experts) to represent information redundantly — in multiple formats, which may be called upon separately depending on the task at hand.

Here, I have discussed examples where the two formats in question are functionally equivalent (in that they both contain the same amount of overall information). That need not always be the case. When thinking about the format of mental images (e.g., Kosslyn et al., 1995; Kosslyn, 1996; Pylyshyn, 1973; Pylyshyn, 2002), for example, we may imagine that pictorial formats contain rather different information (i.e., pixels) from propositional ones (i.e., symbolic structures). One could argue that one format contains *more* or *better* information than the other. In such cases, then, one could argue that one format ought to be preferred over another. My point here is that *even in the extreme cases where two* *formats are functionally equivalent*, there may still be a benefit to relying on them both (for further discussion, see Yousif, 2022).

This is to say that the search for the 'format' of representations is not the search for one thing; it is, potentially, the search for an infinite number of things. The goal of my work — and this thesis — has been to understand the mind in light of the possibilities that (a) the *format* of representations is readily discernable and (b) the mind may format information in a multitude of ways, such that understanding complex behavior may require an understanding of how many distinct formats are stored.

### 1.3 The format frontier: Space

Though I am broadly interested in the format of representations in all disciplines of cognitive science, this thesis focuses on the format of various kinds of spatial representation as 'case studies', for a few reasons. First, space is perhaps the single domain of cognition for which the most is known about the format of mental representations, and so serves as an apt case study of 'mental representation' more broadly. Second, spatial representations may themselves serve as the format for representing other higher-level information, whether that be numerical information or social information (for review, see Peer et al., 2020). As such, understanding the format of spatial representations may in turn reveal the format of other representations. Finally, spatial representation is ubiquitous in the animal kingdom: virtually all organisms depend on representations of space in one way or another (where far fewer may possess imagery, or sophisticated face perception, or speech perception; see, e.g., Müller & Wehner, 1988).

Here, I present three 'case studies' of spatial representation, all of which reveal something about the format of a representation.

*First*, I discuss the format of location representations. I ask, simply: How does the mind represent where things are in space? There are many possibilities. Prior work, for example, has considered whether locations are represented *propositionally*, such that we represent locations relative to one another (e.g., the coffee shop is past the grocery store but before the library). I consider a different possibility: that visuospatial representations are supported by different spatial coordinate systems. Using a novel analysis technique, I show that locations are formatted by default, but flexibly, in polar coordinates. I further show how this method can be used to address other questions about spatial representation.

*Second*, I discuss the format of size representations. I ask: Is there some general rule that captures judgments of size? My work shows that both area and volume judgments can be captured by a simple heuristic — a tendency to 'add' rather than 'multiply' the dimensions of space together (such that the *perceived* area of a square of length L and width W would be equal to L+W rather than L*W).

*Finally*, I discuss how space itself may *be* the format of other representations. I ask: Does task-irrelevant spatial information spontaneously influence working memory? Using a novel paradigm that I developed, I show that task-irrelevant spatial information does indeed boost working memory, and that this effect is specific to spatial information. I discuss these findings in light of our tendency to spatialize other forms of information (e.g., social information, numerical information), and consider how large a role spatial information plays in human memory broadly.

I conclude by offering some thoughts about how each of these case studies informs our understanding of the mind at large.

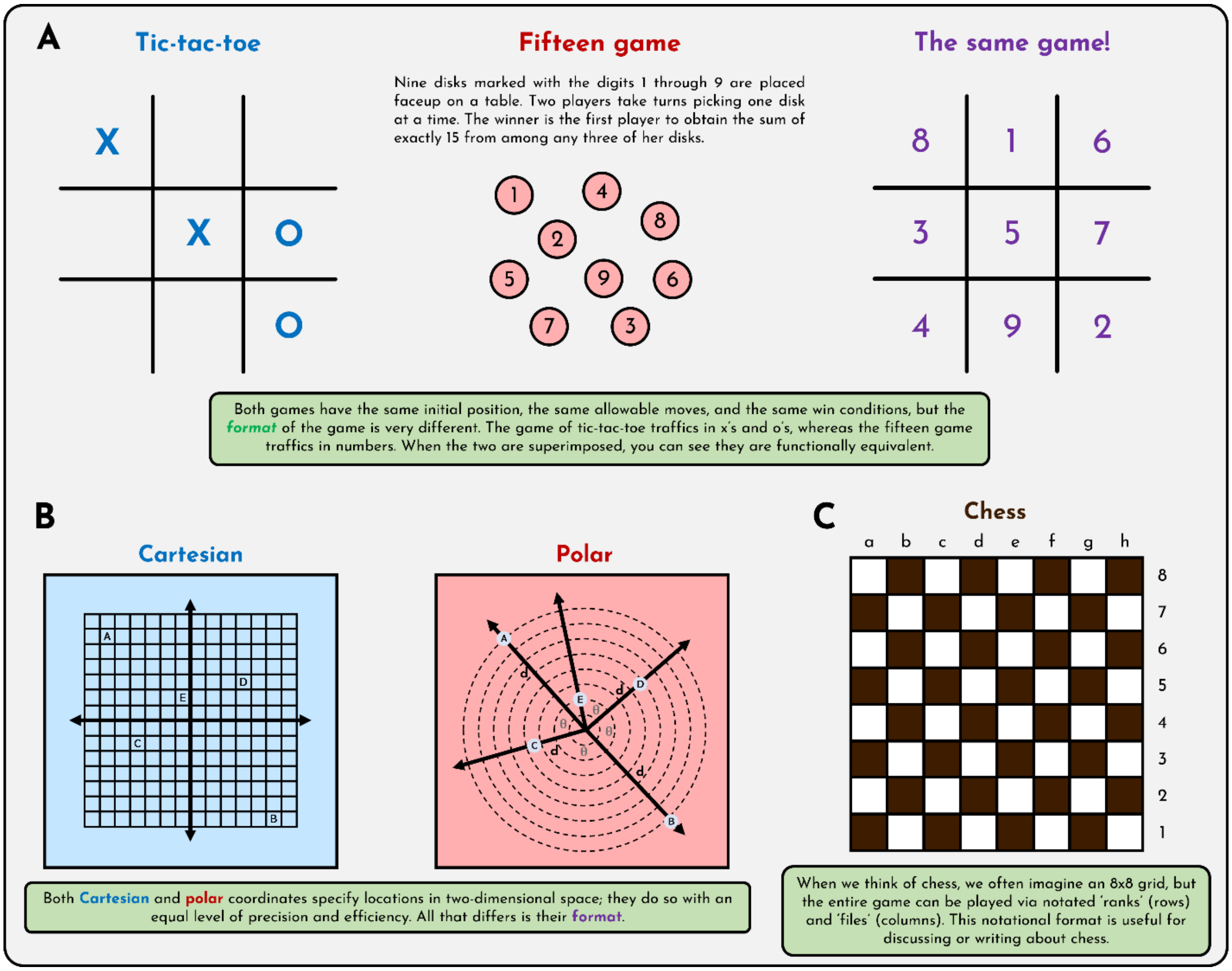



Figure 1.1. Examples of format. (A) A comparison between the classic game of tic-tac-toe and the lesser known 'fifteen game'. One may say that these games differ in their 'format'. (B) A comparison of a classic distinction between canonical spatial formats: Cartesian coordinates and polar coordinates. (C) Even the game of chess can be played in multiple 'formats'.

# Chapter 2

# The format of visuospatial representations

This chapter contains text and/or materials from the following publications:

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

## 2.1 Abstract

What is the *format* of spatial representation? In mathematics, we often conceive of two primary ways of representing two-dimensional space, *Cartesian coordinates*, which capture horizontal and vertical relations, and *polar coordinates*, which captures angle and distance relations. Do either of these two coordinate systems play a representational role in the human mind? Six experiments utilizing a simple 'visual matching' paradigm show that (1) representational format is recoverable from the errors observers make in simple spatial tasks; (2) human-made errors spontaneously favor a polar coordinate system of representation; and (3) observers are capable of using other coordinate systems when acting in highly structured spaces (e.g., grids). We discuss these findings in relation to classic work on dimension independence, as well as work on spatial representation at other spatial scales.

## 2.2 Introduction

A foundational question in the study of any mental process concerns the format of the underlying representation on which that process depends. In our daily use of computers, for example, 'file formats' shape our digital interactions: whether we use a .doc file or a .pdf file, affects how we interact with that information, what metadata is stored about that information, and what other processes (i.e., programs) can act on that information. Similarly, the format of mental representations informs where and how those representations are instantiated, whether they are domain-general or domain-specific, and what kinds of information the mind most naturally represents in the first place. Here, I address this question in the context of spatial representation: using a simple, novel approach, I reveal the latent format of our most basic visuospatial representations.

Spatial representations are foundational to a diverse array of cognitive processes — important for aesthetics (Palmer et al., 2013), for representing numbers (Dehaene et al., 1993; Zorzi et al., 2002), for working memory (e.g., Pertzov & Husain, 2014), and even for reasoning about social relationships (Parkinson & Wheatley, 2013). Yet spatial *behavior* is not always precise: Observers invariably make errors even in straightforward spatial tasks (see Hubbard, 2018; see also McCloskey et al., 1995). For example, in simple spatial memory tasks, observers tend to remember things as having been closer to the quadrant in which they originated (Huttenlocher et al., 1991; Yousif et al., 2020). And when recalling and perceiving oriented lines, observers make larger errors with diagonal lines compared to horizontal or vertical lines (Appelle, 1972; Li et al., 2003; Olson, 2013). These effects are only a few of many spatial biases, ranging from illusions of 2D area (e.g., Coren & Girgus, 1978; Yousif & Keil, 2019), to navigation errors in 3D environments (Warren et al., 2017; Yousif & Lourenco, 2017), to misperceptions of center caused by a conflation of 2D and 3D forms (Firestone & Keil, 2016). Many of these errors remain relatively mysterious, although they sometimes speak to the nature of underlying spatial representations (e.g., Ericson & McCloskey, 1996; Müller et al., 1998). Here, I exploit these spatial biases: I ask whether such mis-localizations hint at the *format* of the underlying representation.

### 2.2.1 Current Study

In several experiments, observers complete a visual matching paradigm in which they see one image (comprised of arrangements of shapes, or dots within shapes) in a corner of the screen, and a corresponding image in the opposite corner of the screen (comprised of the same arrangement of shapes, but sometimes scaled up or down in size). Observers are then instructed to place a missing shape so that the relative locations of the objects in that image exactly match the relative locations of the objects in the other image (see Figures 2.1 and 2.2).

The core conclusions of this chapter rest on an analysis of observers' errors. In short, I measure the correlation between the errors in different dimensions of space (e.g., x vs. y for Cartesian coordinates; angle vs. distance for polar coordinates). If observers represent space via any two-dimensional coordinate system, and the system is efficient, those two dimensions should be orthogonal. I propose, therefore, that *errors* in those two dimensions should also be orthogonal (see Bays, Wu, & Husain, 2011 for an example of this kind of analysis). In other words, if observers represent space via Cartesian coordinates, I expect that their errors in this coordinate system would be *independent*, or *uncorrelated*. Similarly, if observers represent space via Cartesian coordinates, I expect that errors in other coordinate systems (e.g., polar coordinates) to be *dependent*, or *correlated*.

I first validate this kind of analysis by running a simulation in specific coordinate systems (see ‘Experiment 0’, supplementary materials of Yousif & Keil, 2021). Here, I simulate the task — with parameters such as average accuracy set to match that of human participants — with models that operate in *either* Cartesian *or* polar coordinates. These simulations demonstrate how the analyses may succeed in principle. Following this, I present four experiments showing that human observers spontaneously use polar coordinates (Experiments 2.1a-c and Experiment 2.2a). I then show that, despite this default tendency, humans are capable of flexibly operating in other coordinate systems when various levels of spatial structure are imposed on the task environment (Experiments 2.2a-c).

## 2.3 Experiment 2.1a-c: Spontaneous use of polar coordinates

What is the format of human spatial representation? In three experiments, observers complete the same kind of visual matching task that I simulated in ‘Experiment 0’. In opposite corners of the screen, there were matched sets of three shapes, each containing a blue circle, a red square, and a green triangle. At first, one of the sets contained between zero and one of the three original shapes (depending on the experiment; see below). Observers had to then place the missing shapes such that the relative spatial relationships matched that of the set in the opposite corner.

### 2.3.1 Method: Experiment 2.1a

#### *2.3.1.1 Participants*

Sixteen naïve observers from the Yale community completed the experiment in exchange for course credit. This preregistered sample size was chosen before data collection began based and was fixed to be identical for each of the in-lab experiments reported here. Pre-registrations for this experiment and the following experiment are available at the following. This study was approved by the relevant Institutional Review Board.

#### *2.3.1.2 Apparatus*

The experiment was conducted with custom software written in Python with the PsychoPy libraries (Peirce et al., 2019). Observers sat without restraint approximately 60cm from a 43° × 25° display, with all spatial extents reported below computed based on this distance.

#### *2.3.1.3 Stimuli*

The display on each trial consisted of two sets of shapes, each containing three unique shapes (a blue circle, a red square, and a green triangle; each with a thin black border) on a grey (50% white, 50% black) background. The two sets appeared in opposite corners of the display (counterbalanced so that both sets appeared in each corner an equal number of times). The 'center' of each set of shapes was set to be 5.60° horizontally and vertically from the center of the display. The position of each shape within the set was randomly determined so that, for the smaller set, a point could appear within 2.24° horizontally and 2.24° degrees vertically of that set's 'center'. For the larger set, the locations were matched so that they were exactly twice the distance from their respective center (meaning points could appear anywhere within 4.48° horizontally and vertically of the set's center). Random generation of locations was constrained so that no two shapes could appear within 1.25° of one another (from one object's center to another) for the smaller set, and double that distance for the larger set. The smaller shapes were set to have a radius of .36° and the larger shapes were set to have a radius of .72° (here, radius means the distance from the center of the shape to the point along its edge *farthest* from that center). A different set of randomly generated locations was used for each observer. One set of shapes (the smaller set on half of trials; the larger set on the other half; blocked with order counterbalanced across observers as described below), was initially missing two of the three shapes (counterbalanced across trials such that each shape was missing an equal number of times in each block). No other information was visible on the screen at any point. A caricature of a representative trial can be seen in Figure 2.1 (see also Figure 2.2a).

#### *2.3.1.4 Procedure*

On each trial, observers simply had to place the missing shapes to match the relative location of their corresponding shapes, by moving and then clicking the mouse. The missing shape appeared upon mouse-click, at which point observers could click additional times or drag and drop the dot to change its location. Once observers were satisfied with the missing object's location, they pressed a key to submit their response. If a response was recorded, then the display was replaced with a blank screen for a randomly chosen interval of 0.5-1.5s, after which the next trial began. If no response was recorded within 14s, then a warning to respond more quickly appeared for 5s before the next trial began, and that trial was randomly shuffled back into the trial sequence. When that trial was reached, it would utilize the same set positions (i.e., the quadrants where the sets were located) but a different set of random locations would be generated for the objects themselves (i.e., the shapes would appear in different locations relative to one another).

#### *2.3.1.5 Design*

Each observer completed 192 trials, divided into two equal blocks: 96 small-to-large trials (i.e., with the initially missing object in the larger set), and 96 large-to-small trials (i.e., with the initially missing object in the smaller set). Between the two blocks, a message appeared encouraging observers to rest briefly before continuing. Observers completed four representative practice trials (the data from which were not recorded) before beginning the task.

### 2.3.2 Results: Experiment 2.1a

To assess representational format, I first calculated the absolute error (i.e., ignoring the direction of the error) for each subject and each trial in each of the four relevant dimensions (x, y, angle, and radial distance) relative to *where the point should have been*. For example, if the original point was at [2,2], but the second set of shapes was scaled up by in size by a factor of 2, I would calculate error relative to the point [4,4]. (This is an example of scaling in Cartesian coordinates, but the same logic would apply to polar coordinates as well.) I then correlated the dimensions of each coordinate system with each other. Then I took the absolute value of these correlations (because only the relationship, not the direction of that relationship matters). I then average those correlations across individuals and ask whether that average is significantly different from zero.

The 'coordinate system' was always imputed relative to the initially present object. For these analyses, therefore, I calculate a correlation for each person for each point (the first one placed vs. the second one placed), then average those correlations together before asking if they significantly differ from zero. Note, however, that all of the results below replicate if I analyze only the first point observers placed, or only the second point.

The results from this experiment can be seen in Figure 2.3 and Table 2.1. As shown in the table, Cartesian errors were reliably correlated ($M$=.19, $CI$: [.12,.27]; $t(15)=5.68$, $p<.001$, $d=1.42$, $CI_d=[.71,2.11]$), and polar errors were reliably uncorrelated ($M$=.02, $CI$: [-.02,.05]; $t(15)=1.11$, $p=.28$, $d=.28$, $CI_d=[-.23,.77]$). The difference between these two values was also significant ($t(15)=6.47$, $p<.001$, $d=1.62$, $CI_d=[.85,2.36]$). I can also analyze whether non-canonical dimensions are correlated as I should expect them to be. And, indeed, for all of the non-canonical coordinate systems I tested, there was a positive correlation ($p$s<.001; see Table 2.1).

Four factors strengthen the meaningfulness of this null result. *First*, I always pair a predicted null result in one dimension with a predicted positive result in another; in other words, I are more confident that a null result for polar coordinates is meaningful because Cartesian coordinates yield a positive result. *Second*, I have demonstrated via simulation (in 'Experiment 0') that this analysis functions correctly under known conditions. While I cannot perfectly simulate human behavior, these simulations were conducted in a way that mimicked human behavior (e.g., by matching average error along multiple dimensions). *Third*, I can ask about other predictions this view must make. For example, if polar coordinates are implemented in the human mind, then I should expect that any coordinate system *not* in use should have correlated errors. While Cartesian makes for an obvious comparison, there are an infinite number of non-canonical dimensions I can assess. Most straightforwardly, errors in the 'angle' and 'x' dimensions should be correlated with one another, as well as 'radial distance' and 'y' dimensions, and so on. These correlations are presented along with the other relevant correlations in Table 2.1; as you can see from the table, these non-canonical dimensions are correlated as I would expect ($p$s$<$.05). (As with the analyses above, these $p$-values are derived from a one-sample t-test conducted on the correlation values for each participant; I are asking whether on average these correlations differ from zero.) *Fourth,* I can ask about other non-existent dimensions. For example, it is possible in theory to represent 2D space in a Cartesian-esque format that is rotated, e.g., 5 degrees clockwise, 15 degrees clockwise, and so on. In other words, I imagined a Cartesian space that was rotated, e.g., 5 degrees; I then recalculated the coordinates for every error *as if* they existed in this non-existent space. Then I ask whether these non-existent dimensions properly yield positive correlations — and, indeed, they do ($p$s$<$.001). The $p$-values of these average correlations are also plotted in Table 2.1 (as 'rotated Cartesian dimensions'). Therefore, polar representations seem to underlie human errors, as errors in polar coordinates seem to be uniquely *uncorrelated* — compared not only to Cartesian coordinates, but also to other non-canonical 2D spaces.

### 2.3.3 Method: Experiment 2.1b

This experiment was identical to Experiment 2.1a, except as noted. Sixteen new observers participated, with this preregistered sample size chosen to match that of Experiment 1a. In this experiment, three shapes (as opposed to two) were initially absent from one of the sets (see Figure 2.2b). Observers had to place *all three* shapes back on each trial. To guide them, a single black dot appeared in the center of the three shapes (for both sets of shapes). Observers did not know which object would appear first when they clicked for the first time (though the objects always appeared in the same order: blue circle, green triangle, then red square). After they clicked, they could see the object and then adjust as needed. Observers could press spacebar to 'lock-in' the location of the first object, at which point clicking again would cause another shape to appear. They could then adjust the location of that object in the same way. This continued until all three objects were placed and the observer locked in their final response.

### 2.3.4 Results: Experiment 2.1b

The analyses for this experiment are identical to the analyses of Experiment 2.1a, except that there were more points to analyze, because observers placed three objects on each trial instead of two. To simplify these analyses, I will present the average values for all three points. However, the results are qualitatively identical for each of the three points (as is readily apparent in Figure 2.4). The placement error for each object on each trial was always analyzed relative to the central anchor dot. Otherwise, the analyses are identical to those in the prior experiment.

The results of this experiment can be seen in Figure 2.4 and Table 2.1. As shown in the table, Cartesian errors were reliably correlated ($M$=.12, *CI*: [.07,.16]; $t(15)=5.44$, $p<.001$, $d=1.36$, $CI_d$=[.66,2.04]), and polar errors were reliably uncorrelated ($M$=-.01, *CI*: [-.03,.02]; $t(15)=.38$, $p=.71$, $d=.10$, $CI_d$=[-.59,.40]). The difference between these two values was also significant ($t(15)=4.55$, $p<.001$, $d=1.14$, $CI_d$=[.49,1.76]). As with the previous experiment, I can also analyze whether non-canonical dimensions are correlated as I should expect them to be. And, indeed, all the non-canonical coordinate systems tested were positively correlated ($t$s>5.40, $p$s<.001, $d$s>1.30; see Table 2.1). While this task serves mostly as a replication of Experiment 1a, the experience of completing the task was quite different. The presence of only a central anchor point may have altered observers' strategies and suggests that the results of Experiment 2.1a cannot be explained by appeal to some idiosyncratic task demand (and that, on the contrary, this pattern of results may be far more general).

### 2.3.5 Method: Experiment 2.1c

This experiment was identical to Experiment 2.1a, except as noted. Sixteen new observers participated, with this preregistered sample size chosen to match that of Experiment 1a. In both prior experiments, one of the sets was 'scaled up' to be larger than the other. Such scaling does not affect Cartesian and polar coordinates equally. When spaces are scaled in size, both the dimensions of Cartesian space will change (unless a point lies directly along an axis). However, only one of the dimensions of polar space will change; the angle remains constant. Therefore, these results might occur because polar coordinates are simply a more convenient format for spatial translation. Here, I use an identical task but without the size-translation component: observers will be forced to match two identical sets of shapes. The two different sets of shapes were not scaled in size; instead, they were spatially identical. I used exactly the same parameters as Experiment 2.1a, except that the sizes of both sets was 'scaled up' to be equal to the size of the larger set in Experiment 2.1a (see Figure 2.2c).

### 2.3.6 Results: Experiment 2.1c

The analyses for this experiment are identical to the analyses of Experiment 2.1a. Again, to simplify these analyses, I will present the average values for the two placed points. However, the results are qualitatively identical for each of the points (as is readily apparent in Figure 2.5).

The results from this experiment can be seen in Figure 2.5 and Table 2.1. As you can see from the table, Cartesian errors were reliably correlated (*M*=.20, *CI*: [.15,.25]; $t(15)=8.63$, $p<.001$, $d=2.16$, $CI_d=[1.24,3.06]$), and polar errors also exhibited small correlations (*M*=.08, *CI*: [.03,.13]; $t(15)=3.22$, $p=.006$, $d=.81$, $CI_d=[.23,1.36]$). Despite the small polar correlations, the difference between these two values was significant ($t(15)=7.40$, $p<.001$, $d=1.85$, $CI_d=[1.02,2.66]$). As with the previous experiment, I can also analyze whether non-canonical dimensions are correlated as I should expect them to be. And, indeed, for all the non-canonical coordinate systems I tested, there was a positive correlation (*t*s>4.40, *p*s<.001, *d*s>1.05; see Table 2.1). The observed correlation for polar errors was driven largely by a single subject who was an outlier in terms of overall accuracy. However, I did not pre-register any exclusion criteria for accuracy for the in-lab experiments. Despite this anomaly, I note that each of the observers still exhibited a higher Cartesian correlation than polar correlation. Therefore, these results once again reveal evidence of spontaneous use of polar coordinates to represent visual space. Here, crucially, I demonstrate small correlation in polar dimensions even when observers completed no size translation task at all. This suggests that the prior results are not fully explained by an advantage of one coordinate system during spatial translation; however, the slight polar correlations observed here may suggest that the translations do impact behavior. This possibility will be further explored in Experiment 2.2a.

### 2.3.7 Discussion: Experiments 2.1a-c

Experiments 2.1a-c demonstrate that even small errors made by observers in a maximally simple task contain a wealth of information; indeed, these errors may reveal the canonical format of the spatial representations. Experiments 2.1a and 2.1b demonstrate that observers use polar coordinates when scaling spaces up or down in size, whether they are placing them relative to one another or to a single, central landmark. Experiment 2.1c demonstrates that observers use polar coordinates even when matching two spatially identical displays. Together, these three experiments provide evidence that observers spontaneously use polar coordinates to represent visual space.

## 2.4 Experiment 2.2a-c: Flexibility of representation

In tasks with minimal intervening spatial structure (Experiments 2.1a-c), observers automatically operate in polar coordinates. But how flexibly do people engage different coordinate systems across different layouts and reference frames? People's use of polar coordinates might be highly *inflexible*; i.e., regardless of the surrounding spatial environment, people will use only polar coordinates. Or, people might spontaneously use polar coordinates as a default representation but may flexibly represent space in other coordinate systems if the surrounding spatial environment strongly suggests a particular system. Here, I use the same spatial matching task as before, but in environments with varying degrees of spatial structure. In Experiment 2.2a, I replicate the findings of Experiment 2.1; in Experiment 2.2b, I impose moderate structure in the form of a bounding square; and in Experiment 2.2c, I impose strong structure in the form of a grid.

### 2.4.1 Method: Experiment 2a — Online Replication, Minimal Structure

This experiment was identical to Experiment 2.1c, except as noted. Fifty new observers participated. Of the original sample of 50, 3 were excluded for failing to complete the task, and a further 4 were excluded due to being an outlier for overall accuracy; this resulted in a final sample of 43 observers. Unlike the previous experiments, this experiment was conducted online via Amazon Mechanical Turk. (This is because these data were collected in response to a revision request that I received at around the onset of the COVID-19 pandemic. Because I was no longer able to collect data in-lab, I opted to convert the experiments to operate online. This is also why I took care to first replicate the original findings on this new platform before attempting to extend them further.) This experiment was run using custom software written in javascript.

As much as possible, I tried to match the design of the original experiments online. However, due to uncertainty about the viewing conditions of the observers (given differences in web browsers, etc.), I cannot know for sure the exact stimulus dimensions, etc. Here, observers placed three points relative to a single, central anchor point (identical to Experiment 2.1b; see Figure 2.2d). The only other substantive change I made to the task was that I had subjects complete only 48 trials compared to the original 192. Note that even though I collected fewer trials per participant, the sample size was also many times larger. Observers had 20 seconds to make a response before that trial was skipped and replaced; data from these missed trials were discarded. (The pre-registration for this experiment states that observers would have 7 seconds to respond before a trial was skipped. However, this was an error. To be consistent with all the other experiments, this should have said that observers were given roughly 7 seconds *per shape* they had to place. Because they were placing three shapes, they had 20 seconds, rounded down from 21.) To account for increased noise in online data collection, I added exclusion criteria both at the trial-level and subject-level. Any trial with an overall accuracy greater than 2.5 standard deviations from that subjects' mean was discarded; any subject with an overall accuracy greater than 2.5 standard deviations away from the group mean was discarded.

### 2.4.2 Results: Experiment 2.2a — Online Replication, Minimal Structure

The results from this experiment can be seen in Figure 2.6a. The analyses for this experiment are identical to the analyses of Experiment 2.1b. Again, to simplify these analyses, I will present the average values for all three points. However, the results are qualitatively identical for each of the three points. As you can see from the table, Cartesian errors were reliably correlated ($M$=.13, *CI*: [.09,.18]; $t(42)$=6.69, $p$<.001, $d$=1.02, $CI_d$=[.65,1.39]), and polar errors were reliably uncorrelated ($M$=.02, *CI*: [-.02,.06]; $t(42)$=1.03, $p$=.31, $d$=.16, $CI_d$=[-.15,.46]). The difference between these two values was also significant ($t(42)$=6.03, $p$<.001, $d$=.92, $CI_d$=[.56,1.27]). These results replicate the findings of Experiment 2.1a, suggesting that the pattern of results I observed in prior experiments generalizes across testing environments. Further, these results show that the lack of polar correlation is not dependent on the size translation task.

### 2.4.3 Method: Experiment 2.2b — Bounding Square, Moderate Structure

This experiment was identical to Experiment 2.2a, except as noted. Fifty new observers participated, with this preregistered sample size chosen to be identical to Experiment 2.2a. Of the original sample of 50, 1 was excluded for failing to complete the task, and a further 6 were excluded due to being an outlier for overall accuracy; this resulted in a final sample of 43 observers.

Whereas in the previous experiment observers placed three points relative to a central anchor point, observers in this experiment matched the location of only one point within a square frame (see Figure 2.2e). The goal of this experiment was to provide observers with a *moderate* level of spatial structure. Because observers were placing only a single shape (as opposed to three shapes), they only had 7 seconds to respond before a trial was skipped. Analyses in this experiment must be conducted relative to the center of the square.

### 2.4.4 Results: Experiment 2.2b — Bounding Square, Moderate Structure

The results from this experiment can be seen in Figure 2.6b. The analyses for this experiment are identical to the analyses of Experiment 2.2a, except that observers placed only a single point on each trial. As you can see from the table, I observed for the first time a case where Cartesian errors were *un*correlated ($M$=.04, $CI$: [-.03,.11]; $t(42)$=1.24, $p$=.22, $d$=.19, $CI_d$=[-.11,.49]), and polar errors were correlated ($M$=.16, $CI$: [.10,.21]; $t(42)$=6.23, $p$<.001, $d$=.95, $CI_d$=[.59,1.31]). The difference between these values was significant ($t(42)$=2.98, $p$=.005, $d$=.46, $CI_d$=[.14,.77]). This unique pattern of results suggests that spatial structure may affect observers' strategies. For the first time, I have observed uncorrelated Cartesian errors but correlated polar errors.

### 2.4.5 Method: Experiment 2.2c — Grid, Maximal Structure

This experiment was identical to Experiment 2.2b, except as noted. Fifty new observers participated, with this preregistered sample size chosen to be identical to Experiment 2.2a. Of the original sample of 50, 1 was excluded for failing to complete the task, and a further 3 were excluded due to being an outlier for overall accuracy; this resulted in a final sample of 46 observers.

Whereas in the previous experiment observers placed one point within a bounding square, observers here placed one point on top of a grid (see Figure 2.2f). The goal here was to provide observers with a *high* level of spatial structure (here meaning that, unlike the prior experiments, observers have enough spatial information to make very exact estimates of the object's position; this is why I imposed a time limit on responses; although observers could in theory respond with almost perfect accuracy, this imposed time limit is meant to force small errors).

### 2.4.6 Results: Experiment 2.2c — Grid, Maximal Structure

The results from this experiment can be seen in Figure 2.6c. The analyses for this experiment are identical to the analyses of Experiment 2.2b. As you can see from the figure, I once again observed that Cartesian errors were *un*correlated ($M$=.00, $CI$: [-.05,.06]; $t(45)$=.14, $p$=.89, $d$=.02, $\mathrm{CI_d}$=[-.27,.31]), and polar errors were correlated ($M$=.32, $CI$: [.27,.37]; $t(45)$=12.84, $p$<.001, $d$=1.89, $\mathrm{CI_d}$=[1.40,2.37]). The difference between these two values was significant ($t(45)$=7.61, $p$<.001, $d$=1.12, $\mathrm{CI_d}$=[.75,1.49]). This reversal is significant for two reasons: (1) it validates the analysis in the first place, demonstrating that this way of analyzing errors is capable of revealing different strategies that observers may take, and (2) it suggests that while may spontaneously use polar coordinates in environments with minimal spatial structure, they are capable of flexibly using different spatial representations when the environment strongly implies such representations.

### 2.4.7 Discussion: Experiment 2.2a-c

In a series of three experiments, I have shown how varying levels of spatial structure influence the *kind* of coordinate systems observers use to localize objects. In Experiment 2.2a, I replicated the findings of Experiments 2.1a-c, demonstrating that in the absence of strong spatial cues observers will spontaneously use polar coordinates. But in Experiments 2.2b and 2.2c with increasing levels of spatial structure — and, in particular, structure that may lend itself to Cartesian-esque representations — observers' patterns of errors revealed an increasing shift towards Cartesian coordinates. These results validate the previous analyses while revealing the 'boundary conditions' of the use of polar coordinates.

Another way of thinking about the results of Experiments 2.2b and 2.2c is with respect to *references frames* (see e.g., Farah et al., 1990). In Experiments 2.1a-c and 2.2a, the only possible referents (or reference frame) observers can use to situate the placement of new objects are single points in space (i.e., the already visible objects). In Experiments 2.2b and 2.2c, by contrast, observers have an entire bounded region (i.e., the square/grid) with which to situate the new object. Note here that this work does not imply that observers do or should only use one coordinate system, or one reference frame. Quite the opposite, this approach is meant to be flexible: in principle, these analyses can be conducted relative to any point in space and with respect to any reference frame. And, in practice, this is clearly the case: the fact that I observe a qualitatively different pattern of results in Experiments 2.2b and 2.2c suggests that observers are clearly capable representing space within different frames of reference.

## 2.5 General Discussion

I first demonstrated that analyses of errors have the potential to reveal representational format when that representational format is known ('Experiment 0'). I then applied this insight to six experiments with humans. In Experiments 1a and 1b I found converging evidence of polar coordinates in a simple visual matching paradigm in which no representational format was implied. In Experiment 2.1c, I showed that these prior results cannot be explained by the size translation task itself. In the following three experiments (Experiments 2.2a-c), I explored whether observers flexibly use different coordinate systems depending on the spatial context. With high levels of spatial structure (i.e., imposing the spatial matching task upon a grid), observers' pattern of errors suggested the use of Cartesian rather than polar coordinates. Collectively, these results demonstrate spontaneous, but flexible, use of polar coordinates.

These results are far from obvious: all dimensions could have been consistently correlated, or none; or, contrary to what I found, Cartesian coordinates could have been uniquely uncorrelated. Yet, the same pattern held across many variations of experiments (namely Experiments 2.1a-c and Experiment 2.2a), suggesting a robust set of findings. That said, any one of these results in isolation should be interpreted cautiously. I am comfortable interpreting the lack of correlation for polar dimensions as speaking to representational format only because (1) this result is highly replicable across many observers and several unique experiments, (2) I was able to conduct simulations indicating that this analysis could work in principle, (3) I showed data (in Experiment 2.2c) where this analysis revealed a change in representation in practice, and (4) I assessed a number of other dimensions that also could have been uncorrelated (yet never were).

Nevertheless, the present work depends on a single paradigm. While this paradigm is revealing, future work may still fruitfully seek converging evidence to support this view. For example, I now know that observers are capable of flexibly swapping between coordinate systems depending on the context. Yet further investigation may be able to address what specific context information may be sufficient to induce a change in representational format, as well as the interface between small-scale and large-scale representations (e.g., could I use this analysis to measure representational format in navigable environments?). I see the present work as a first step — one that opens the door to many other lines of inquiry.

### 2.5.1 Relation to prior work

These findings relate to prior studies on spatial (mis)-localization (Huttenlocher et al., 1991; Langlois et al., 2021; Wedell, et al., 2007; Yousif et al., 2020), some of which specifically address polar coordinates as a candidate for visuospatial representation (Huttenlocher et al., 1991; Yousif et al., 2020; see also Yang & Flombaum, 2018). Most notably, Huttenlocher and colleagues (1991) relied on similar correlation analyses to make claims about representational format. However, those results were indecisive for a few reasons: (1) the primary aim of the paper was to understand the origin of spatial biases, not to document the format of visuospatial representations; (2) their conclusions depend solely on a null result, without making predictions about or assessing other dimensions or other spatial contexts (whereas the present work tests many positive predictions and also tests many different spatial contexts); (3) they assess dimension independence only in circular spaces (whereas I specifically sought to test unbounded spaces); and (4) they assess errors in *memory* whereas all of these tasks intentionally minimize memory demands.

These results may also bear on spatial representation on larger scales or in three-dimensional environments (e.g., for purposes of navigation; see Moser et al., 2008; Moser et al., 2014). One relevant proposal suggests I use a 'network-like' cognitive graph for large scale spatial systems. These graphs especially prioritize angle and distance information between known locations (Ericson & Warren, 2020; Warren et al., 2017; but see also Gallistel, 1990; Kuipers et al., 2003; O'Keefe & Nadel, 1978). Of course, this resembles polar coordinates, which are nothing more than angle and distance vectors. Could the same highly general representational format be employed in both small-scale (i.e., visual) and large-scale (i.e., navigable) environments? Future work may shed light on the continuity of these representations across scales, or on the translation of information between small-scale and large-scale layouts (e.g., as when reading maps).

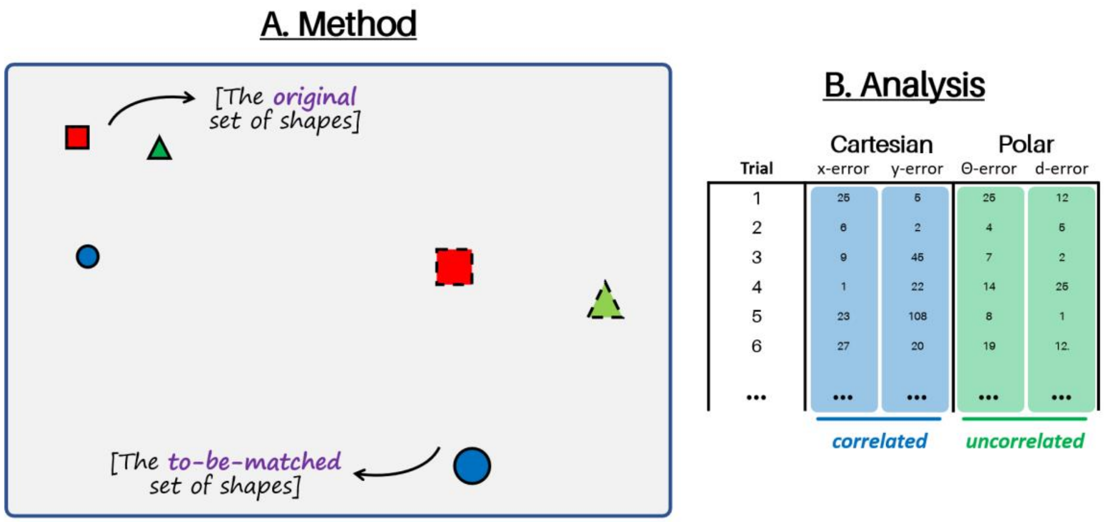
The approach here also relates to classic work on integral vs. separable dimensions (Garner & Felfoldy, 1970; for review, see Algom & Fitousi, 2016). Traditionally, studies investigate integrality/separability in one of two ways: either via Stroop effects or via speeded classification. Dimensions that interfere with one another would be considered integral; dimensions that do not would be considered separable. The error-independence analyses I use here might provide a novel method for assessing integrality vs. separability (see also Bays et al., 2011); in principle, all three analyses should yield converging results. That said, these classic paradigms would have been insufficient to address the key questions here. Space is not like other dimensions in that any point in space could be simultaneously represented in an infinite number of 2D spaces. Because this task provides a blank slate with which I can simultaneously analyze all possible dimensions at once, it provides a unique advantage over earlier tasks. Stroop and speeded classification paradigms, in contrast, require pre-commitment to the relevant dimensions. Nevertheless, future work may link this approach to the integrality vs. separability approach.

### 2.5.2 On 'format'

Throughout this chapter I have focused primarily on a contrast between polar coordinates and Cartesian coordinates (and, briefly, other non-canonical coordinate systems). However, this work also speaks to a possible contrast between the use of *some* coordinate system and *no coordinate system at all.* Indeed, it is possible that space could be represented only in coarse spatial terms (e.g., "that point was generally up and to the left"; see Huttenlocher et al., 1991). Both sets of experiments reported here, by contrast, suggest a reliance on a specific coordinate system. Even if participants flexibly rely on multiple coordinate systems, their patterns of errors have still revealed a fundamental regularity: locations in the mind are represented as variables in a two-dimensional vector. In some ways, these insights were presaged by the study of patient AH, who exhibited profound localization deficits that often involved 'mirror flipping' points in space (McCloskey et al., 1995; McCloskey & Palmer, 1996). E.g., if AH was instructed to recreate the location of a point offset to the left, they might place a point in the same relative location but offset to the right instead. Such errors suggest that space is being represented in some precise format, but one that can be manipulated (akin to flipping the sign of a variable). This work also suggests that space is represented via independent dimensions (or else it would be impossible to make an error in one dimension while acting precisely in another). The present work builds on the study of patient AH by offering — for the first time — evidence that the precise coordinate systems underlying visuospatial representations are readily recoverable, even in simple psychophysical tasks.

### 2.5.3 Conclusion

We *depend* on our ability to accurately perceive and represent space; yet, naturally, our percepts and our representations are imprecise. Here, I show how errors in the simplest possible spatial tasks contain significant clues to the underlying format of our most primitive visuospatial representations. The present work lays the groundwork for considering domain-general mechanisms that may underlie many kinds of spatial biases (e.g., those pertaining to location vs. those pertaining to orientation) across many different spatial scales (e.g., small-scale visual environments vs. large-scale navigable environments). More consequentially, the present work demonstrates that the format of spatial representations is readily accessible to empirical investigation.




| | Cartesian | | Polar | |
|---|---|---|---|---|
| Trial | x-error | y-error | Θ-error | d-error |
| 1 | 25 | 5 | 25 | 12 |
| 2 | 6 | 2 | 4 | 5 |
| 3 | 9 | 45 | 7 | 2 |
| 4 | 1 | 22 | 14 | 25 |
| 5 | 23 | 108 | 8 | 1 |
| 6 | 27 | 20 | 19 | 12. |
| ••• | ••• | ••• | ••• | ••• |

Figure 2.1. (A) A caricature of the method. In this example, observers must place a triangle and a square (which are initially absent from the display) in the bottom-right to match the spatial relationships of the three shapes in the top left. Here, the relative spatial relations of the set in the bottom-right are ‘scaled up’ by a factor of two. (B) A caricature of the analyses. Here, we are analyzing correlations between errors in Cartesian vs. polar coordinates. Each subject therefore has a unique correlation value that is used for subsequent analyses.

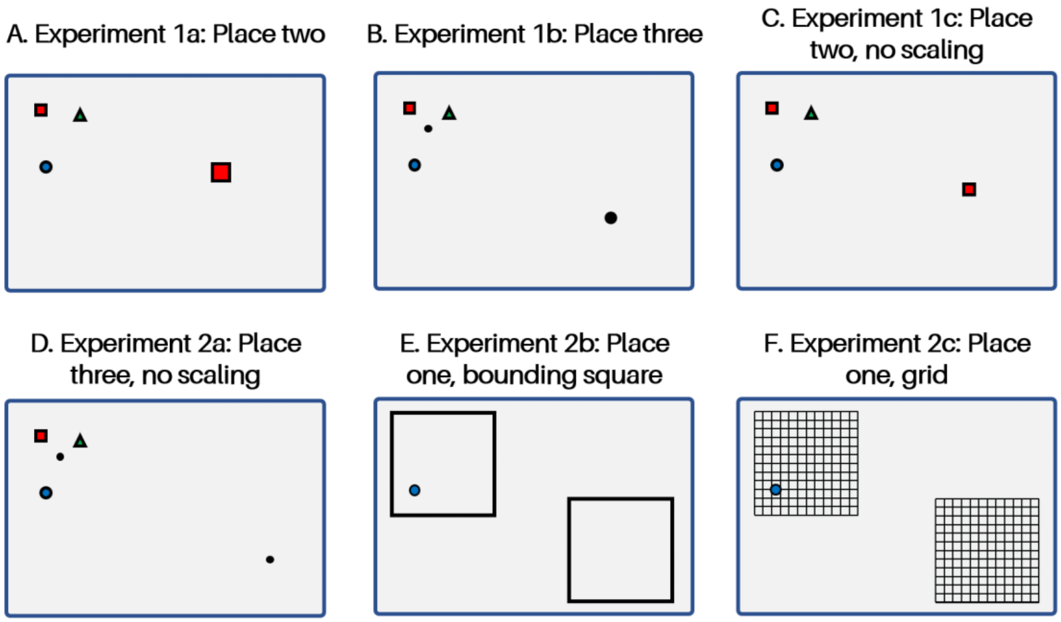



Figure 2.2. Schematic of each of the 6 experiments. Items presented here are approximately but not exactly to scale.

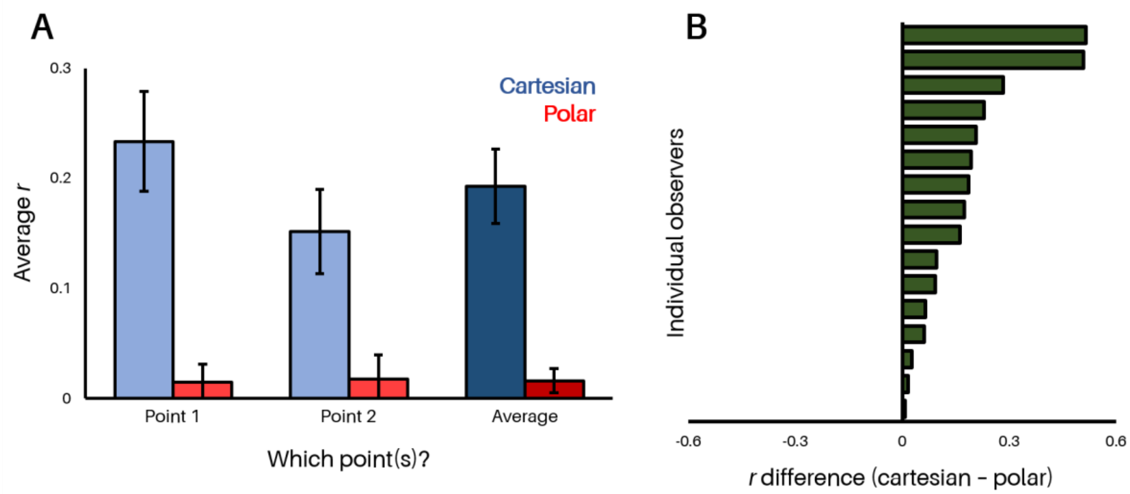



Figure 2.3. Results from Experiment 2.1a. (A) A depiction of the average correlation for Cartesian and polar errors broken down by point. (The faded bars correspond to individual points; the darker bars correspond to the average of those values.) Cartesian correlations are depicted in blue; polar correlations are depicted in red. Error bars represent +/- 1 SE. (B) The difference in correlation (Cartesian minus polar) for each observer. Bars to the right of the y-axis indicate a greater correlation for Cartesian dimensions than polar dimensions. See also Table 2.1 for additional information/statistics about these correlation values.

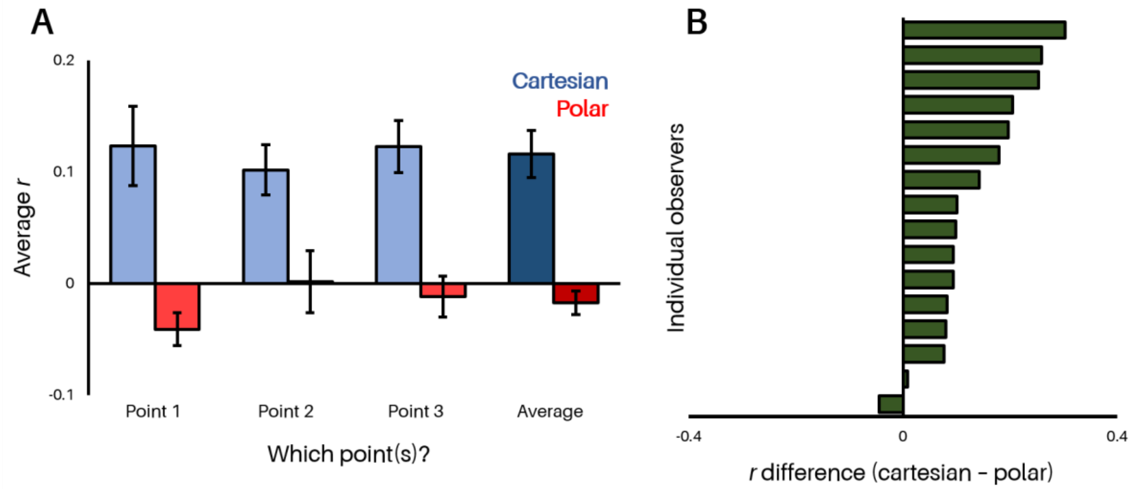



Figure 2.4. Results from Experiment 2.1b. (A) A depiction of the average correlation for Cartesian and polar errors broken down by point. (The faded bars correspond to individual points; the darker bars correspond to the average of those values.) Cartesian correlations are depicted in blue; polar correlations are depicted in red. Error bars represent +/- 1 SE. (B) The difference in correlation (Cartesian minus polar) for each observer. Bars to the right of the y-axis indicate a greater correlation for Cartesian dimensions than polar dimensions. See also Table 2.1 for additional information/statistics about these correlation values.

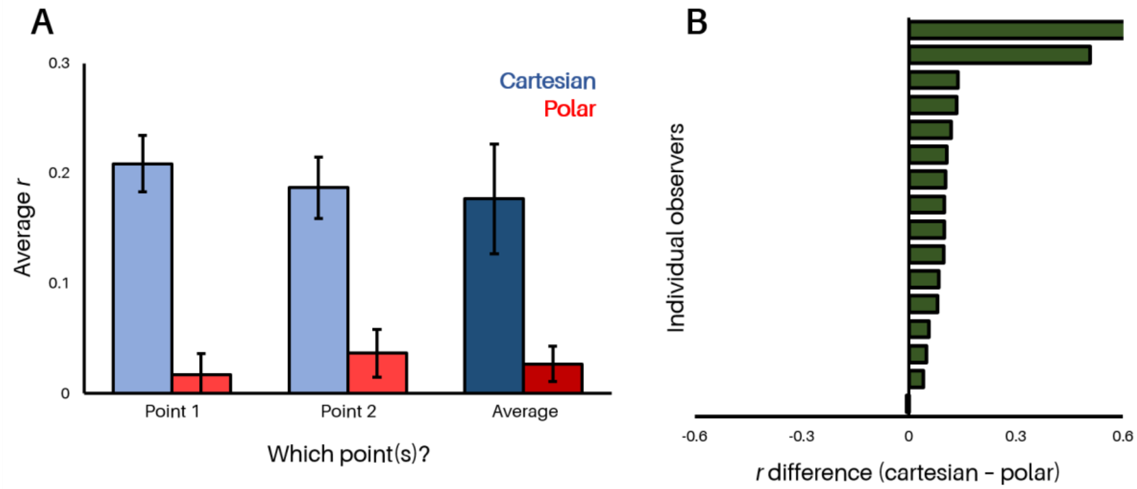



Figure 2.5. Results from Experiment 2.1c. (A) A depiction of the average correlation for Cartesian and polar errors broken down by point. (The faded bars correspond to individual points; the darker bars correspond to the average of those values.) Cartesian correlations are depicted in blue; polar correlations are depicted in red. Error bars represent +/- 1 SE. (B) The difference in correlation (Cartesian minus polar) for each observer. Bars to the right of the y-axis indicate a greater correlation for Cartesian dimensions than polar dimensions. See also Table 2.1 for additional information/statistics about these correlation values.

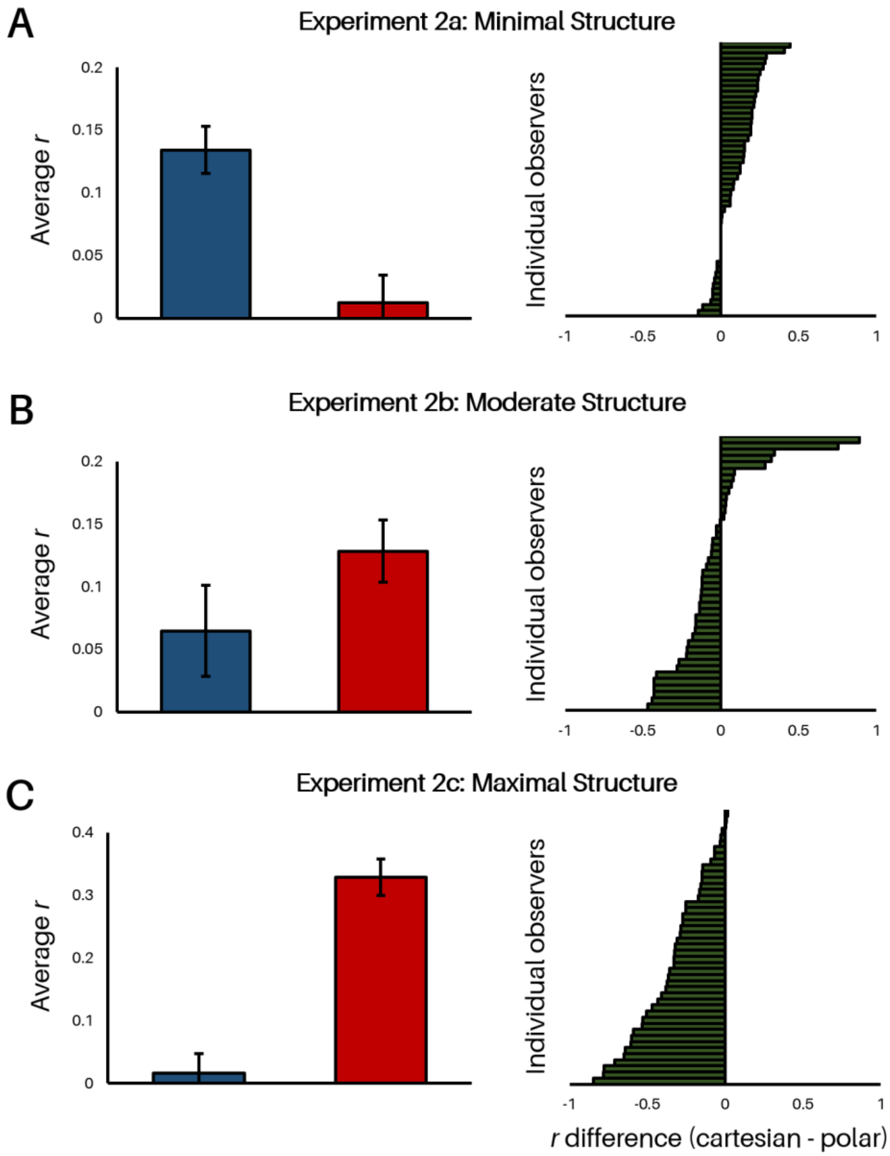



Figure 2.6. Results from Experiments 2.2a-c. On the left side are the average correlation values. Cartesian correlations are depicted in blue; polar correlations are depicted in red. Error bars represent +/- 1 SE. On the right side are the differences in correlations (Cartesian minus polar) for each observer. Bars to the right of the y-axis indicate a greater correlation for Cartesian dimensions than polar dimensions.

***p*-values**

| | Primary correlations | | | Other dimensions | | | | Rotated Cartesian Dimensions | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Expt. | Cartesian | Polar | Diff | x/angle | x/dist | y/angle | y/dist | 5 deg | 15 deg | 25 deg | 35 deg | 45 deg |
| Exp 1a | <.001 | =.162 | <.001 | <.001 | <.001 | =.024 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 |
| Exp 1b | <.001 | =.120 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 |
| Exp 1c | =.003 | =.117 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 |

***r* values**

| | Primary correlations | | | Other dimensions | | | | Rotated Cartesian Dimensions | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Expt. | Cartesian | Polar | Diff | x/angle | x/dist | y/angle | y/dist | 5 deg | 15 deg | 25 deg | 35 deg | 45 deg |
| Exp 1a | .19 | .02 | .18 | .12 | .67 | .04 | .66 | .20 | .21 | .22 | .21 | .21 |
| Exp 1b | .12 | -.02 | .13 | .10 | .66 | .11 | .67 | .18 | .20 | .22 | .24 | .24 |
| Exp 1c | .18 | .03 | .15 | .18 | .56 | .14 | .55 | .19 | .19 | .20 | .22 | .24 |

Table 2.1. *p*-values and *r* values from Experiments 2.1a-c. Each row corresponds to a separate experiment. The three most important values are the first three columns, which represent the *p* and *r* values for Cartesian errors, for polar errors, and for the difference between the two. (The 'Diff' column is the actual numerical difference in the correlation; the *p*-value corresponds to the output of the one-sample t-test conducted on those difference scores.) The following four columns ('Other dimensions') are correlations between various existing dimensions. The final five columns ('Rotated Cartesian Dimensions') are unique coordinate spaces that we created for purposes of these analyses. Here we took ordinary Cartesian space and rotated it by 5, 15, 25, 35, or 45 degrees to create new coordinate systems. The aim here is to demonstrate that the lack of correlation for polar coordinates is special — as all other combinations of dimensions yield positive correlations. Highlighted in light blue are any cells that correspond to a positive correlation; highlighted in red are any cells that correspond to the absence of a correlation. The key takeaway from this table is that the only red cells are for polar errors.

# Chapter 3

# The format of size representations

This chapter contains text and/or materials from the following publications:

Yousif, S. R., Alexandrov, E.*, Bennette, E.*, Aslin, R. N., and Keil, F. C. (2022). Children estimate area using an 'Additive-Area Heuristic'. *Developmental Science.*

Yousif, S. R. and Keil, F. C. (2021). How we see area and why it matters. *Trends in Cognitive Sciences*, *25*, 554-557.

Yousif, S. R., Aslin, R. N., and Keil, F. C. (2020). Judgments of spatial extent are fundamentally illusory: 'Additive-area' provides the best explanation. *Cognition*, *205*, 104439.

Yousif, S. R. and Keil, F. C. (2020). Area, not number, dominates estimates of visual quantities. *Scientific Reports, 10*, 1-13.

Yousif, S. R. and Keil, F. C. (2019). The 'Additive-Area Heuristic': An efficient but illusory means of visual area approximation. *Psychological Science, 30,* 495–503.

## 3.1 Abstract

How does the visual system determine 'how much' of something is present? A large body of research has investigated the mechanisms and consequences of number estimation, yet surprisingly little work has investigated *area* estimation. Indeed, area is often treated as a pesky confound in the study of number. Here, I describe the 'additive area heuristic': a means of rapidly estimating visual area that results in substantial distortions of perceived area in many contexts, visible even in simple demonstrations. I show that when controlling for additive area, observers are *unable* to discriminate on the basis of true area *per se*, and that these results cannot be explained by other spatial dimensions. These findings reflect a powerful perceptual illusion in their own right, but also have implications for other work — namely that which relies on area controls to make claims about number estimation. I discuss several extensions of these findings, as well as several areas of research impacted by these findings.

## 3.2 Introduction

When you look at a basket of oranges, or a glass of water, or a piece of cake, how do you know how much is there? A great deal of research has investigated the shared capacity of adults, infants, and nonhuman animals to estimate the *number* of objects in a set (Barth et al., 2003; Brannon & Terrace, 1998; Gordon, 2004; Nieder & Miller, 2004; Pica et al., 2004). This propensity to estimate large numerosities without counting is said to rely on an evolutionarily ancient system: the Approximate Number System (ANS; Halberda et al., 2008). Yet, to our evolutionary ancestors, estimates of number may not have been the best assessment of amount.

Imagine foraging for food. Would you select to forage from the bush with twice as many berries, or the one with berries three times in volume? In many natural settings, size estimation rather than number estimation might be most critical for survival – though only a few studies have investigated approximate area perception conjointly with approximate number (Brannon et al., 2006; Lourenco et al., 2012; Odic et al., 2013). In fact, most studies have discussed area only in an attempt to *rule out* continuous spatial dimensions (e.g., area, contour length, density) as explanations for approximate number estimation (Barth, 2008; Mix et al., 2002). Yet, area/size perception is also an autonomous area of study: models of area perception have been proposed in the context of development (e.g., Anderson & Cuneo, 1978; Gigerenzer & Richter, 1990), ensemble perception (e.g., Marchant et al., 2013; Solomon et al., 2011), perception research more broadly (e.g., Carbon, 2016; Ekman & Junge, 1961; Nacmhias, 2008; Nachmias, 2011; Teghtsoonian, 1965), and even consumer decision-making (e.g., Krider et al., 2001). Here, similarly, we demonstrate that area estimation itself reveals powerful and counterintuitive effects. Yet, unlike much of the prior work, we assess area perception a) in a context of numerous objects, and b) using displays akin to those commonly used to assess approximate number perception (e.g., Halberda et al., 2008; Odic et al., 2013). We show that, even in such displays, area estimation employs a simple heuristic that results in substantial distortions of perceived area. These distortions are not only important to understand in their own right; they also raise questions about attempts to control for area in number estimation tasks. In particular, controlling for *true* area, insofar as it dissociates from *perceived* area, may amplify a confound with numerosity in many studies.

### 3.2.1 The 'Additive Area Heuristic'

We propose that visual area estimation in simple visual displays is best captured by a single, simple heuristic: the 'Additive Area (AA) Heuristic'. Consider Figure 3.1A. Which panel looks like it has more: the left, or the right? Although the left may *look like* it has more than the right, the two are equal in cumulative area. However, they differ in one important way: *additive area* (i.e., the sum of a shape's dimensions rather than the product) is greater for the image on the left.

Five experiments were designed with the aim of manipulating either *true* area or *additive* area while holding the other constant. We find that a) humans use a simple heuristic to calculate area, b) humans often *fail to perceive* true area when accounting for this heuristic, c) this heuristic cannot be explained by appeal to other dimensions, and d) differences between *true* area and *perceived* area may have serious consequences for studies that rely on area as a manipulation or as a control.

## 3.3 Experiment 3.1: 'Additive Area' vs. True area

In a first test of the AA heuristic, observers completed an approximate area task on displays of circular discs (see Figure 3.1A). Critically, we varied these displays not only in their cumulative *true* area, but also in their cumulative *additive* area. We predicted that observers would be both slower to respond and less accurate when true area differed, and that they would be faster to respond and more accurate when AA differed.

### 3.3.1 Method

This experiment (as well as each subsequent experiment) was pre-registered. In addition to pre-registering the sample size and the basic methodology, we also pre-registered some details about how the stimuli were created (which we will reiterate here).

#### *3.3.1.1 Participants*

100 observers were recruited via Amazon's Mechanical Turk, though 3 observers were excluded because they did not complete a single trial (i.e., they accepted the HIT, but never started the task). All observers consented prior to participation, and these studies were approved by the IRB at Yale University.

#### *3.3.1.2 Materials*

All of the stimuli were generated via custom software written in Python with the PsychoPy libraries (Peirce, 2007). The aim was to create pairs of stimuli that varied in either AA or true area while the other was equated. For each stimulus pair, we randomly generated an initial set of discs (ranging from 20 pixels to 100 pixels in diameter, with a buffer of at least 10 pixels between any two discs), then pseudo-randomly generated a second set of objects based on a given AA ratio. The initial set of objects always had 7 discs. Stimulus pairs were generated randomly until a pair met both the AA criterion *and* the true area criterion, at which point that pair would be rendered another time and saved. The second stimulus (i.e., not the set with 7 discs) always had more area (whether AA or true area) than the initial stimulus. Number was unconstrained in the stimulus generation process, meaning that the number ratio is not equated across all possible AA and true area ratios (1.5 on average for AA trials; .8 on average for true area trials). All discs were rendered with a thin, black border (4-pixel stroke width). The images depicted in Figure 3.1 are representative, as they were actual images used in this experiment.

In this initial experiment, there were only two constraints: AA and true area. There were pairs where both true area was equated (to serve as a baseline), cases where true area varied while AA was controlled, and cases where AA varied while true area was controlled. While AA was controlled, area could vary in either a 1.00, 1.10, 1.20, or 1.30 ratio (and vice versa for AA while true area was controlled). As there are mathematic constraints on how much AA and true area can differ, these ratios were selected to maximize the differences between them. Because of the pseudo-random nature of stimuli creation and the mathematical constraints involved in creation such stimuli, true area was never perfectly matched with the stated ratio; it could vary +/- 1%. That is, if the true area ratio for a given trial was 1.10, then we allowed the difference in true area to fluctuate between 1.09 and 1.11.

#### *3.3.1.3 Procedure*

The task itself was administered online via Amazon Mechanical Turk, using custom software. On each trial, observers saw two spatially separated sets of lavender-colored dots, presented side-by-side in the center of the screen, with 50 pixels of space in between. Each stimulus was 400 pixels by 400 pixels. The side that contained the set with more area was counterbalanced such that half the time the left side had more cumulative area and half the time the right side had more cumulative area. Observers were instructed to press 'q' if the image on the left had more cumulative area, and 'p' if the image on the right had more cumulative area. Observers were told the following: "Your task is simply to indicate which set of circles has more cumulative area. In other words: if you printed the images out on a sheet of paper, which would require more total ink?" Later, they were told: "The sets of dots will sometimes vary in number, but the number of dots does not matter. Instead, you should answer only which has more area, regardless of number." The stimuli stayed on the screen until the observer made a response, and there was no time limit on responses. Between each trial, there was a 1000ms ITI. Observers completed 84 trials, 12 of each of 7 trial types (true area varying in a 1.10, 1.20, or 1.30 ratio while AA was held constant; AA varying in a 1.10, 1.20, or 1.30 ratio while true area was held constant; and trials where both were equated). All trials were presented in a unique random order for each participant. Observers completed two representative practice trials before beginning the actual task. Because over half of the trials had no objectively correct answer (because true area did not vary), we measured accuracy as a *propensity to choose 'more'* – whether that be more AA or more true area.

### 3.3.2 Results

The results are shown in Figure 3.2. Observers were indeed faster and more accurate in making discriminations on the basis of AA rather than true area. A repeated-measures ANOVA conducted on accuracy with two factors (condition: AA vs. true area; ratio: 1.10, 1.20, and 1.30) revealed main effects of both condition ($F[1,96]=17.80$; $p<.001$) and ratio ($F[2,95]=24.43$; $p<.001$), as well an interaction between the two ($F[2,95]=9.94$; $p<.001$). Post-hoc tests revealed that overall performance was above chance in the AA condition ($t[96]=10.88$, $p<.001$; $d=1.11$). However, surprisingly, observers were unable to make discriminations on the basis of true area alone ($t[96]=1.70$, $p=.09$; $d=.17$). Even in the trials with the biggest difference in area (1.30 ratio), observers were not above-chance in their area discriminations ($t[96]=1.93$, $p=.06$; $d=.20$). A separate ANOVA conducted on response times revealed a similar pattern, and post-hoc tests confirmed that individuals were over 120ms faster for the AA trials compared to the true area trials ($t[96]=4.88$, $p<.001$; $d=.50$).

Because we did not explicitly manipulate number, we tested whether number could potentially explain these results. A linear regression with AA, true area, and number as covariates revealed that AA did predict observer responses ($p<.001$) but that neither true area ($p=.86$) nor number ($p=.23$) did.

### 3.3.3 Discussion

AA can explain variance in area estimation and, surprisingly, observers seem unable to make area discriminations using true area when AA is controlled.

## 3.4 Experiment 3.2: Time-limited approximations

To ensure that all participants spent roughly the same amount of time assessing the displays — and to ensure that these judgments are, in fact, rapid approximations — I replicated Experiment 3.1, except that observers had only 700ms to view the stimuli.

### 3.4.1 Method

100 observers were recruited via Amazon's Mechanical Turk, though 3 observers were excluded because they did not complete a single trial (i.e., they accepted the HIT, but never started the task). All observers consented prior to participation, and these studies were approved by the IRB at Yale University. This experiment, like the previous experiment, was pre-registered. All of the details of this experiment were exactly identical to Experiment 3.1, except that the stimuli appeared for only 700ms before disappearing. Observers still had an unlimited amount of time to make their responses.

### 3.4.2 Results

Results from this manipulation can be seen in Figure 3.3A. Observers were more accurate in the AA condition than the true area condition ($t[96]=6.40$, $p<.001$; $d=.65$), though there were no differences in response times (which is to be expected, because participants were rushed; $t[96]=1.60$, $p=.112$; $d=.50$). Observers were unable to make discriminations on the basis of true area alone ($t[96]=1.26$, $p=.21$; $d=.13$).

Once again, linear regression revealed that AA significantly predicted observer responses ($p<.005$), whereas number did not ($p=.51$). True area was also a significant predictor in this model ($p$=.035), though in the opposite direction (i.e., participants were *less* likely to choose the option with more area). This latter result is likely driven by all the ratios where true area was zero and additive area varied, and thus should not be over-interpreted. Note that overall performance for the true area trials did not significantly differ from chance ($p$=.21).

### 3.4.3 Discussion

Experiment 3.2 further supports the 'heuristic' nature of this phenomenon: in addition to replicating Experiment 3.1, these results show that AA is used for rapid approximation of visual displays.

## 3.5 Experiment 3.3: Rectangles

Most studies on the ANS have relied on displays of discs. However, to ensure that the AA heuristic was not specific to such displays, we replicated the results of the prior experiment, but with rectangles instead of discs (see Figure 3.1B).

### 3.5.1 Method

100 observers were recruited via Amazon's Mechanical Turk, though 1 observer excluded because they did not complete a single trial (i.e., they accepted the HIT, but never started the task). All observers consented prior to participation, and these studies were approved by the IRB at Yale University. This experiment, like the previous experiments, was pre-registered. All of the details of this experiment were exactly identical to Experiment 3.2, except that the stimuli were rectangles instead of discs. The aspect ratio of the rectangles varied from 1.0 to 5.0 (the minimum dimension length was 20 pixels; the maximum was 100). (And, to rule

out minor differences caused by how those borders are rendered, we rendered these rectangles without borders.)

### 3.5.2 Results

Once again (see Figure 3.3b), observers were more accurate in making discriminations on the basis of AA rather than true area ($t[98]=2.61$, $p=.01$; $d=.26$). Observers were above chance at making true area discriminations ($t[98]=4.88$, $p<.001$; $d=.49$), though they only made the correct selection 57% of the time. A linear regression revealed that while both AA ($p<.005$) and true area ($p<.05$) significantly predicted observer responses, number did not ($p=.61$).

### 3.5.3 Discussion

In general, observers had more trouble making area discriminations with rectangles – both in AA and true area trials. We suspect this is due to one dimension being over-weighted relative to the other, meaning that a slightly more complex model might best explain area approximations in such cases. Although the effects in this experiment are weaker than those in prior experiments, AA still outperforms true area as a model of area approximation.

## 3.6 Experiment 3.4: Number control

Might these results be explained by a confound with number? The creation of the stimuli in Experiments 3.1-3 was constrained in such a way that number was partially confounded with AA. In all three cases, differences in AA predicted accuracy, while differences in number did not. In a stronger test, we constructed stimuli for which we could independently manipulate number.

### 3.6.1 Method

100 observers were recruited via Amazon's Mechanical Turk, though 2 observers were excluded because they did not complete a single trial (i.e., they accepted the HIT, but never started the task). All observers consented prior to participation, and these studies were approved by the IRB at Yale University. This experiment, like the previous experiments, was pre-registered. All of the details of this experiment were exactly identical to Experiment 3.1, except as otherwise noted.

There were stimuli 60 pairs, 4 of each of 15 types (5 AA/true area ratios x 3 numerosities each). Additionally, it should be noted that it is not possible to use the exact same number ratios across of the AA/true area ratios. The goal, instead, was merely to have three different levels of numerosity at each AA/true area ratio. This way we could independently assess the role of number at each level. To determine what numbers ought to be chosen in the first place, we ran an initial simulation to see how number would naturally vary (if unconstrained) for each AA/true area ratio. From these initial simulations, we picked three of the possible numerosities. We purposefully chose numerosities that would maximally overlap across conditions (to minimize the impact of any unforeseen confound). The default number of items in each display was set to 10.

### 3.6.2 Results

The results from this experiment can be seen in Figure 3.4. Replicating previous results, observers were both faster and more accurate in the AA trials compared to the true area trials (accuracy: $t[97]=5.02$, $p<.001$, $d=.51$; response time: $t[97]=3.31$, $p=.001$, $d=.33$). Regression once again revealed that AA ($p<.001$) but not true area ($p$=.14) significantly predicted observers' responses. The same regression revealed that number *did* significantly predict responses ($p$=.004), such that greater numerosity resulted in a *decreased* likelihood to indicate an item had more area. However, this effect was specific to the true area trials ($p$=.004). If we analyze only the trials in which AA varied, there is no effect of number ($p$=.92). In other words, although number may be used as a cue in certain contexts, it has no apparent effect on area judgments when perceived area does differ.

### 3.6.3 Discussion

These results once again reveal the use of an AA heuristic. However, there was an effect of numerosity whereby the presence of additional discs in the display *decreased* the likelihood that an observer would indicate that display had more area. This is in contrast to previous results that suggest correspondences between number and continuous magnitudes such as area (Hurewitz et al., 2006). Thus, it seems that many past studies reporting influences of numerosity in these sorts of tasks may have been detecting variation caused by AA instead. Importantly, when perceived area does vary, observers do not rely on number as a cue.

## 3.7 Experiment 3.5: Perimeter control

More than most continuous spatial dimensions, perimeter has been a dimension of interest in these sorts of displays (e.g., DeWind et al., 2015; McCrink & Wynn, 2007), and there is some evidence that perimeter may actually explain number approximation (see DeWind et al., 2015; Mix et al., 2002). While perimeter-based approximations may not serve as feasible models of area perception (see 'coastline paradox'; Mandlebrot, 1967), they should be addressed. In a final experiment, we used a new stimulus to fully dissociate perimeter from perceived area: ellipses.

### 3.7.1 Method

100 observers were recruited via Amazon's Mechanical Turk. 1 observer was excluded for failing to complete the task. All observers consented prior to participation, and these studies were approved by the IRB at Yale University. This experiment, like the prior experiments, was pre-registered. All of the details of this experiment were exactly identical to Experiment 3.2, except substituting perimeter for true area. To do this, ellipses were used in place of discs (see Figure 3.1C). In other words, AA varied while perimeter was held constant, and perimeter varied while additive area was held constant (and both varied in 1.0, 1.1, 1.2, and 1.3 ratios). No specific limits were imposed on area or numerosity (meaning that, in practice, they varied much more than either AA or perimeter). The default number of stimuli was 15. These stimuli were rendered without borders. The aspect ratio of the discs ranged from 1.0 to 2.20.

### 3.7.2 Results

Overall, observers were better at making discriminations on the basis of AA rather than perimeter ($t[98]=10.31$, $p<.001$; $d=1.04$). Observers were unable to make discriminations on the basis of perimeter alone ($t[98]=.99$, $p=.33$; $d=.10$).

### 3.7.3 Discussion

Cumulative perimeter seems unable to explain the approximation of area while AA alone is able to do so.

## 3.8 General Discussion

Not only does the AA heuristic account for a high proportion of the variance in area judgments, but, also, observers seem to be insensitive to differences in true area under certain conditions. These results have implications for many different research programs in cognitive science. We highlight four areas of active research likely to be influenced by these findings.

### 3.8.1 Visual perception

Many papers have addressed the question of size perception (Ekman & Junge, 1961; Teghtsoonian, 1965). Some have addressed illusions of visual size (Coren & Girgus, 1978). Others have discussed the continuous dimensions of space that influence not only the perception of size, but also the perception of density, numerosity, and texture (e.g., Anobile et al., 2014; Anobile et al., 2016; Durgin, 1995). In all of these cases, the additive area heuristic offers a simple, powerful, low-dimensional means of area estimation. This finding may clarify and unify various prior studies on the perception of area (e.g., Carbon, 2016), while also raising questions about links between the perception of *size* (of a single object), *area* (of a set of objects), density, and texture.

### 3.8.2 Approximate Area

The study of approximate area is not nearly as pervasive as the study of approximate number, yet several prominent papers have studied the two in tandem (Lourenco et al., 2012; Odic et al., 2013). Both approximate number estimation and approximate area estimation are proposed to independently contribute unique variance to mathematical competence (Lourenco et al., 2012). However, this work involved manipulation of mathematical area rather than perceived area. Thus, number discrimination could have been influenced by AA, even though mathematical area was controlled.

### 3.8.3 Approximate Number

The subject of hundreds of papers and cumulatively tens of thousands of citations, the ANS has dominated the field of numerical cognition for the past decade (Barth et al., 2003; Halberda et al., 2008; Lourenco et al., 2012; Lourenco & Bonny, 2017). Much attention has been given to the continuous spatial dimensions that are confounded with numerosity (e.g., Barth, 2008; DeWind et al., 2015; Mix et al., 2002). Of these, area is by far the most common control (e.g., Halberda et al., 2008; Lourenco et al., 2012; Xu & Spelke, 2000). Yet, if *true* area is different from *perceived* area, variance in perceived area might well explain performance on these tasks.

### 3.8.4 General Magnitude

Several studies have investigated the link between number and other magnitudes. One prominent theory suggests that representations of time, space, quantity, and other magnitudes rely on similar cortical processes (Lourenco & Longo, 2010; Sokolowski et al., 2017; Walsh, 2003). In support of this theory, many have pointed to Stroop-like errors between area and number (Brannon et al., 2004; Hurewitz et al., 2006; Rousselle et al., 2004). Though the present results do not bear on all facets of this extensive literature, they do relate to the tendency to use number as a cue to approximate area and vice versa (e.g., Hurewitz et al., 2006). This could be the result of shared mechanisms, but it might also be the result of a simple confound. The bias to select the set with more number might instead be a bias to select the set with more perceived area. These findings suggest either a) that number has an adverse effect on area estimation – exactly the opposite the general magnitude account– or b) observers are using some other heuristic to make their responses in these cases (e.g., choosing the display with the single largest object).

### 3.8.5 Extensions of these findings

In subsequent work, I have shown that this illusion of area is robust in several ways. First, these illusions cannot be explained by variation in number; when sets of items are perfectly equated in terms of number, this illusion persists (Yousif, Aslin, & Keil, 2020). Second, an 'additive heuristic' appears to capture volume judgments as well as area judgments (Bennette, Keil, & Yousif, 2021). Third, children appear to (at least sometimes) rely on an 'additive heuristic' to make 2D area judgments (Yousif et al., 2022). Fourth, these illusions are significant in that they meaningfully influence and are related to the perception of number (Yousif & Keil, 2020). And, finally, despite objections that these effects arise from statistical anomalies across the stimuli, these illusions manifest even when observers witness only a single trial (as in Yousif & Keil, 2021c, as well as Bennette, Keil, & Yousif, 2021).

### 3.8.6 The *illusion* of approximate area

There have been many careful attempts to capture approximate number acuity by modeling a nearly exhaustive list of continuous dimensions of the stimuli (DeWind et al., 2015), concluding that continuous dimensions of space influence the approximation of number. What does this approach reveal that is not already captured by existing models? Consider the Ebbinghaus illusion, whereby one disc, surrounded by many smaller discs, appears greater in size than an equal-sized disc surrounded by many larger discs. Modeling ANS performance by exhaustively characterizing every continuous dimension of a display is akin to explaining the Ebbinghaus illusion by measuring every continuous dimension of the two discs being compared. No measurements collected on the relevant discs could explain the Ebbinghaus illusion because they are exactly the same; it can be explained only by appeal to perception.

By contrast, we make an explicit prediction about what drives area approximation and manipulate that specific dimension in order to *eliminate* differences in perceived area. This does *not* mean that AA fully explains area perception: there may be context- or task-dependent interactions among many continuous variables (e.g., density, convex hull, average element size) that contribute to the perception of size. Yet, controlling AA *eliminated* the ability to distinguish displays in most cases, providing strong evidence that this factor is directly linked to area perception. Further, this heuristic offers a simple solution that may be easily implemented in ANS studies.

### 3.8.7 Conclusion

This chapter documents the 'additive area heuristic': a simple, low-dimensional heuristic that accounts for substantial variability in area approximation. The explanatory power of this heuristic persists despite variance in other salient dimensions (e.g., true area, perimeter, and number), and may bear on the interpretation of many seminal papers in the field of numerical cognition, as well as work on area estimation, general magnitude, and various aspects of visual perception. The notion of *perceived area* helps explain other findings in many diverse fields of cognitive science, while advancing those fields both theoretically and methodologically.

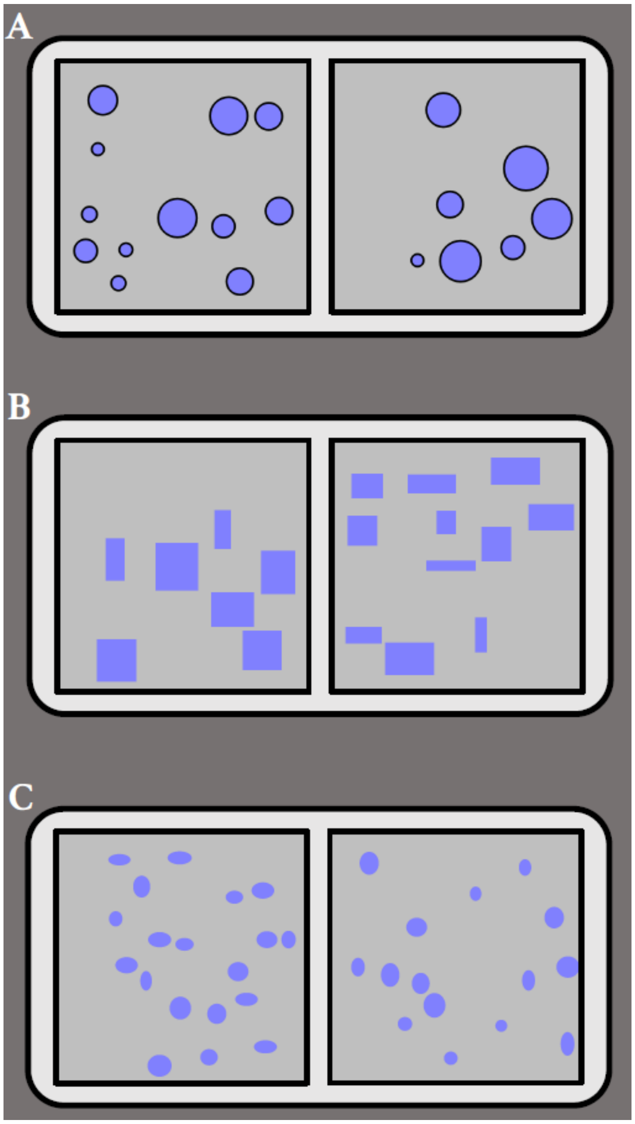



Figure 3.1. Depiction of example displays from Experiments 3.1, 3.2, and 3.4 (A), Experiment 3.3 (B), and Experiment 3.5 (C). True area is equated for each pair in (A) and (B). However, additive area is 30% greater in the left panel of (A), and 30% greater in the right panel of (B). Perimeter is equated for each pair in (C). However, additive area is 30% greater in the left panel of (C). The stimuli appear here exactly as they would have to observers in the task. Additive area in each case is equal to the sum of the objects' cumulative heights and widths. For circles, additive area for each shape is equal to twice the diameter (which can be simplified to just diameter). For the rectangles, additive area for each shape is equal to height + width. For ellipses, additive area for each shape is equal to height + width (also the sums of the lengths of the major and minor axes).

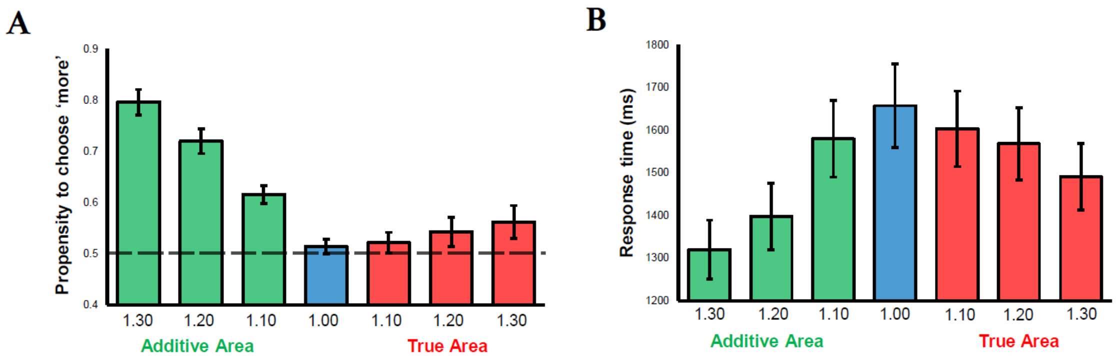



Figure 3.2. Results from Experiment 3.1: (A) The proportion of trials for which observers select the option with 'more' – whether that was more true area or more additive area – for each of the seven additive area/true area ratios tested. The dashed lined represents at-chance performance. (B) Response times for each of the seven ratios tested. In both graphs, the x-axis represents the ratio. While additive area varied, true area remained constant. While true area varied, additive area remained constant. Thus, green bars correspond to additive area trials, red bars correspond to true area trials, and the blue bar represents trials where both were equated. Error bars represent +/- 1 SE.

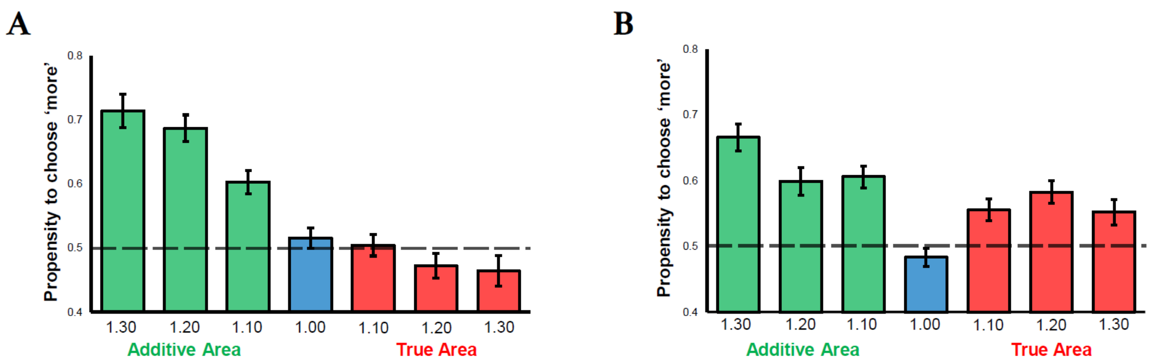



Figure 3.3. Results from Experiment 3.2 (A) and Experiment 3.3 (B). The proportion of trials for which observers select the option with 'more' – whether that was more true area or more additive area – for each of the seven additive area/true area ratios tested. The dashed lined represents at-chance performance. The x-axis represents the ratio. While additive area varied, true area remained constant. While true area varied, additive area remained constant. Thus, green bars correspond to additive area trials, red bars correspond to true area trials, and the blue bar represents trials where both were equated. Error bars represent +/- 1 SE.

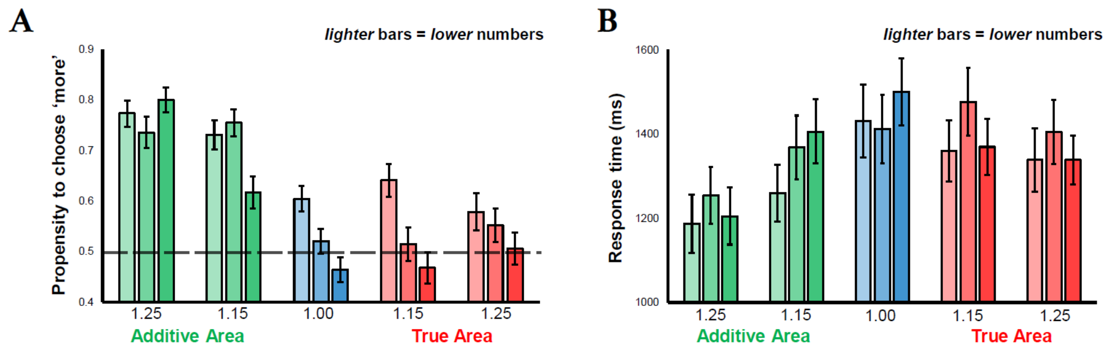



Figure 3.4. Results from Experiment 3.4: (A) The proportion of trials for which observers select the option with 'more' – whether that was more true area or more additive area – for each of the seven additive area/true area ratios tested. The dashed lined represents at-chance performance. (B) Response times for each of the seven ratios tested. In both graphs, the x-axis represents the ratio. While additive area varied, true area remained constant. While true area varied, additive area remained constant. Thus, green bars correspond to additive area trials, red bars correspond to true area trials, and the blue bar represents trials where both were equated. Three different numerosities were tested for each area ratio; lighter bars correspond to lower numerosities within each set. Error bars represent +/- 1 SE.

# Chapter 4

# Spatial information *as* format: A 'case study' from working memory

This chapter contains text and/or materials from the following publications:

Yousif, S. R., Rosenberg, M. D., and Keil, F. C. (2021). Using space to remember: Short-term spatial structure spontaneously improves working memory. *Cognition*, *214*, 104748.

## 4.1 Abstract

Spatial information plays an important role in how we remember. In general, there are two (non mutually exclusive) views regarding the role that space plays in memory. One view is that objects overlapping in space interfere with each other in memory. For example, objects presented in the same location (at different points in time) are more frequently confused with one another than objects that are not. Another view is that spatial information can 'bootstrap' other kinds of information. For example, remembering a phone number is easier when one can see the arrangement of a keypad. Here, building on both perspectives, I test the hypothesis that task-irrelevant *spatial structure* (i.e., objects appearing in stable locations over repeated iterations) improves working memory. Across 7 experiments, I demonstrate that (1) irrelevant spatial structure improves memory for sequences of objects; (2) this effect does not depend on long-term spatial associations; (3) this effect is unique to space (as opposed to features like color); and (4) *spatial structure* can be teased apart from *spatial interference*, and the former drives memory improvement. I discuss how these findings relate to and challenge 'spatial interference' accounts as well as 'visuospatial bootstrapping'. More broadly, these findings speak to how spatial information may be leveraged *as* the format of higher-level representations.

## 4.2 Introduction

Spatial representations are implicated in a diverse array of cognitive processes — from aesthetics (Palmer et al., 2013), to numerical cognition (Dehaene et al., 1993; Zorzi et al., 2002), to social cognition (Parkinson & Wheatley, 2013). For millennia, there has also been a notion that space plays some pivotal role in how we remember, or else can be used to improve memory (as in the 'method of loci'). Here, I investigate the role of space in working memory to better understand the flexibility of spatial representations in memory and to address the ways in which spatial representations influence memory in the first place.

Here, I focus on working memory, which refers to the short-term maintenance and manipulation of information in the mind (Baddeley, 1992). I consider two distinct possibilities regarding the relationship between space and working memory. The first I call 'spatial interference'. This view suggests that items appearing in the same location in space interfere with one another (e.g., Treisman & Zhang, 2006). The second is 'visuospatial bootstrapping'. Work on visuospatial bootstrapping suggests that items presented with stable spatial mappings (e.g., as in the digits on a keypad) are better remembered than items without such mappings (for review, see Darling et al., 2017). While these views are not mutually exclusive, they stem from different approaches. I will briefly highlight typical 'spatial interference' and 'visuospatial bootstrapping' effects.

### 4.2.1 Spatial interference

One theory is that space supports the binding of features to objects, and therefore that objects overlapping in space interfere with one another (e.g., Treisman & Zhang, 2006). That is, if you see a blue circle and a green triangle in the same location, you will be more likely to experience a feature binding error, misremembering having seen either a green circle or a blue triangle. Similar effects are observed across a wide range of paradigms (cf. Jiang et al., 2000; Rajsic & Wilson, 2014; Woodman et al., 2012). For example, participants tasked with remembering the orientation of a line perform worse when multiple lines appeared in the same location (Pertzov & Husain, 2014). Such effects are also specific to space: lines with overlapping colors do *not* result in the same kind of interference (see also Rajsic & Wilson, 2014).

### 4.2.2 Visuospatial bootstrapping

Suppose you are memorizing a phone number; perhaps you would try to visualize where each number is located on a keypad. Indeed, people are better at remembering verbal information when it is mapped onto consistent spatial locations: digits are better remembered when they are presented in a keypad formation, as opposed to being presented in a single location, or even in a line (Darling & Havelka, 2010; Race et al., 2015). However, other work suggests this is only true when those mappings exist in long-term memory; people do *not* better remember digits that are displayed in a rearranged keypad formation (Darling et al., 2012). This phenomenon of *'visuospatial bootstrapping'* is said to speak to communication between verbal and visual information systems in working memory (for review, Darling et al., 2017).

A related body of work suggests that items in working memory are automatically spatialized (Abrahamse et al., 2014; Aulet et al., 2017; Guida & Campitelli, 2019; Guida et al., 2016; Guida et al., 2018; van Dijck & Fias, 2011; van Dijck et al., 2014). In other words, imagine a task in which individuals must remember a sequence of objects: 'orange', 'apple', 'pear', 'banana', 'cherry'; after memorizing this sequence, individuals respond relatively faster to earlier items with their left hand compared to their right hand, and relatively faster to later items with their right hand compared to their left hand. Such lateralization suggests that the sequence had been mapped onto space in some way, and perhaps that this mapping was functionally involved in the maintenance of that information in working memory. This is known as the Spatial–Positional Association of Response Codes ('SPoARC').

So far, I have discussed working memory broadly as a system for maintaining and manipulating information online. However, there are many distinct models of the working memory system. For example, earlier working memory models differentiated discrete subsystems of working memory: a visuospatial sketchpad, which manipulates visual information, a phonological loop, which manipulates verbal information, and a central executive (Baddeley & Hitch, 1974). Later, a fourth subsystem was proposed, the episodic buffer, to explain the interactions of information across modality and across memory systems (i.e., short-term and long-term memory; Baddeley, 2000). Other views characterize working memory more continuously and do not segregate visual and verbal working memory (and thus have no need for a fourth system to moderate between them; e.g., Cowan, 1998). The exact nature of working memory remains a topic of ongoing debate.

Note, however, that visuospatial bootstrapping and related phenomena such as the 'SPoARC' effect address interactions between short-term and long-term memory (e.g., Darling et al., 2012; Race et al., 2015) as well as interactions between verbal and visuospatial information in working memory (see also Alloway et al., 2006). As such, this body of work speaks not only to how spatial information influences memory, but also to the organization of our memory systems in the first place (in that information in long-term memory influences the retention of information in short-term memory, and in that visuospatial information influences the retention of verbal information). In this way, the present results may bear on debates regarding the extent to which working memory is modality specific (see, e.g., Allen et al., 2015; Morey, 2018).

### 4.2.3 Current Study

Here, I test the hypothesis that *short-term spatial structure* supports working memory maintenance, even in tasks that pose no explicit spatial requirements. Spatial structure could take many forms, and here I operationalize it as a consistent mapping between objects and space (as in 'visuospatial bootstrapping'; Darling et al., 2017). For example, shopping in a grocery store whose aisles are rearranged every time you visited would be an experience with low spatial structure, whereas attending a meeting in which participants always sat in the same seats would be an experience with high spatial structure. In the current set of studies, I ask whether objects repeatedly appearing in the same location (although on a shorter time scale) are better remembered than objects repeatedly appearing in different locations.

Based on prior work suggesting a role of spatial representations in a broad range of cognitive processes (both in the domain of working memory, e.g., Darling et al., 2017; Pertzov & Husain, 2014; and beyond the domain of working memory; e.g., Dehaene et al., 1993; Parkinson & Wheatley, 2013; Zorzi et al., 2002), I hypothesize that task-irrelevant spatial information will benefit visual working memory more than matched non-spatial visual information (i.e., color) and non-visual information (i.e., audio information). I will also further probe *how* space is special, specifically contrasting the idea of 'spatial structure' with both 'spatial interference' and 'visuospatial bootstrapping' effects.

These broad goals can be broken down into a few specific aims. *First*, I asked about the interaction of space, long-term memory, and working memory. Prior work has suggested that spatial mappings facilitate memory only when those spatial mappings are held in long-term memory (Darling et al., 2012). However, another possibility is that visuospatial bootstrapping does not *depend* on mappings in long-term memory, but that mappings in long-term memory *interfere* with short-term mappings. Here, I test short-term spatial mappings (i.e., established over a span of 6-10 seconds) in a case where there are no long-term mappings. *Second*, I asked *how* spatial information influenced working memory (assuming an effect of space in the first place). For example, what if objects are mapped onto stable locations, but other objects are mapped onto those same locations? Work on neither 'spatial interference' nor 'visuospatial bootstrapping' makes clear predictions about such cases. *Finally*, I want to understand how spatial structure and verbal rehearsal compete when participants can rely on both. The task used here allowed for both visual rehearsal (see Awh et al., 1998) and verbal rehearsal of the relevant information through memorizing sequences of common shapes (i.e., circles, pentagons, and diamonds). Those shapes appeared in different locations and in different colors, but participants were specifically instructed to remember only (a) what shapes they saw and (b) what order they saw them in. Participants were free to verbally rehearse the sequences (which they frequently did) — but this was not a requirement. Therefore, akin to visuospatial bootstrapping tasks (Darling et al., 2017), the task here may speak to interactions between verbal and visual working memory.

To address these questions, I present seven experiments, all with the same essential components. (1) Participants always remembered sequences of shapes that they were free to verbally rehearse, and (2) the sequences were often structured such that either location information or color information (and in one case audio information) covaried with the different shapes. E.g., in a 'space-structured' condition, any shape that appeared multiple times in a sequence always appeared in the same location, but the colors of those shapes are random (and vice-versa for a 'colored-structured' condition). Combined, these features allow us to probe when and how we use space to remember and how these effects inform current research on working memory.

## 4.3 Experiments 4.1a and 4.1b: Space vs. Color

In a first set of experiments, participants completed the simplest version of this paradigm: they saw a series of 5-7 shapes (comprised of three unique shapes appearing at least once each, and in a random order) and had to recall (a) what those shapes were and (b) the order that they saw them in. Crucially, participants completed two blocks of trials, each of which was structured in a unique way. In the 'space-structured' block, any shape appearing multiple times appeared in the same location; no other shapes appeared in that location, and the colors of all shapes were randomized. In the 'color-structured' block, any shape appearing multiple times appeared in the same color; no other shapes appeared in that color, and the locations of all shapes were randomized. There were two key questions: (1) Does either location-based or color-based structure uniquely influence working memory, and (2) If so, does structure influence working memory *even when* participants report verbal rehearsal strategies?

### 4.3.1 Method

This experiment, and all subsequent experiments, were pre-registered. Experiments 1a and 1b were identical, except for their sample sizes and one change to the instructions (as noted below).

#### *4.3.1.1 Participants*

24 undergraduate students participated in Experiment 4.1a and 16 undergraduate students participated in Experiment 4.1b in exchange for course credit. The sample sizes were chosen in advance based on pilot data and were pre-registered. The sample size of Experiment 4.1b was chosen based on repeatedly sub-sampling data from Experiment 4.1a and finding that 16 participants were sufficient to demonstrate the primary effect. This study was approved by the relevant Institutional Review Board.

#### *4.3.1.2 Apparatus*

The experiment was conducted with custom software written in Python with the PsychoPy libraries (Peirce et al., 2019). Participants sat without restraint approximately 60cm from a $49° \times 29°$ display, with all spatial extents reported below computed based on this distance.

#### *4.3.1.3 Stimuli*

The display consisted of four black squares ($5.10° \times 5.10°$) on a grey (50% black; 50% white) background (Figure 4.1). The squares were located in each of the four quadrants of the screen, each 7.66° horizontally and 7.66° vertically displaced from the center of the screen. The shapes themselves (a circle, a pentagon, and a diamond) were all just shorter than 5.10° in height and appeared centered within one of the four black squares. They appeared in one of four colors (the default 'red', 'green', 'blue', or 'yellow' in PsychoPy).

#### *4.3.1.4 Procedure & Design*

On each trial, participants watched as shapes appeared one at a time in different locations and in different colors. The shapes appeared for 1000ms, with 500ms between presentations. Any given trial had either 5, 6, or 7 shapes (see more on how the sequences were constructed below). After all shapes were presented, the three shapes appeared in white side-by-side in the center of the screen, in a random order. The four black squares in each quadrant remained on the screen during this time. Participants then had to click on the shapes in the order that they saw them. Even though the shapes varied in color and location, participants knew that they would only have to report what shapes they saw and what order they saw them in. They were specifically told that they could only click one time for each shape that they saw (e.g., if they saw seven shapes, they were instructed to click seven times; they were allowed to click the same shape multiple times), and that the next trial would automatically start when they had pressed the correct number of shapes. The purpose of this was to ensure that for each trial there were a number of responses equal to the number of items in the sequence, thus simplifying the measure of performance. There was no counter or any other indicator reminding them how many shapes they had seen. Each time a shape was selected, it briefly flashed yellow (as a form of visual feedback, so that participants would know their response was recorded). Once a certain number of shapes were selected (equal to the number of items that had been in the previous sequence), the experiment automatically moved onto the next trial (after a 1.5s delay).

The sequences of shapes were constructed in a few important ways. First, there were two distinct trial types, divided into two unique blocks. In the 'space-structured' block, any shape appearing multiple times appeared in the same location; no other shapes appeared in that location, and the colors of all shapes were randomized. In the 'color-structured' block, any shape appearing multiple times appeared in the same color; no other shapes appeared in that color, and the locations of all shapes were randomized. The number of colors and locations were matched (4). The first three shapes of each sequence were always unique; in other words, participants always saw all three shapes within the first three. The remaining two to four shapes were random in every respect (except that they adhered to the relevant structure, depending on the block).

Each participant completed 48 trials, divided into two equal, counterbalanced blocks: a space-structured block, and a color-structured block. In each of these blocks, participants completed 24 trials (3 difficulties [5, 6, or 7 shapes] × 8 unique trials). Between the two blocks, a message appeared encouraging participants to rest briefly before continuing. Participants completed one representative practice trial (the data from which were not recorded) before beginning the task. Including instructions and practice trials, the total task duration was about 25 minutes.

In Experiment 4.1a, but not in Experiment 4.1b or subsequent experiments, participants were explicitly cued to the relevant structure. They were specifically told that, although both color and location information were irrelevant to their task, they were free to use this information if it benefited them. The instructions explained the way that color and location would be structured, in general, and told them that the two blocks of trials would be distinct in this way. However, participants were also reminded that they could disregard or ignore this information as they wished.

### 4.3.2 Results & Discussion

Results from Experiment 4.1a can be seen in Figure 4.2 (panels A & B). Accuracy was generally higher for space-structured trials ($M$=.84, $SD$=.08) compared to color-structured trials ($M$=.79, $SD$=.11), and this effect was most pronounced at higher set sizes. Indeed, a repeated measures ANOVA revealed a main effect of set size ($F[2,46]=20.30$, $p<.001$, $\eta_p 2=.47$), a main effect of trial type ($F[1,23]=13.10$, $p=.001$, $\eta_p 2=.36$), and an interaction between the two ($F[2,46]=7.87$, $p=.001$, $\eta_p 2=.26$). Post-hoc tests confirmed that accuracy was higher for space-structured trials than color-structured trials ($t[23]=3.62$, $p=.001$, $d=.74$), and that accuracy was higher for set size 5 than 6 ($p=.003$), and higher for set size 6 than 7 ($p=.002$).

For all experiments, I calculated Bayes factors for the key experimental contrasts (i.e., between the space-structured and color-structured trials) to assess null effects. I report Bayes factors for significant results (such as ones here) for consistency. Bayes factors are reported as measure of relative evidence for an alternative hypothesis (here, a difference in accuracy between experimental conditions) relative to a null hypothesis (no difference between conditions). Whereas Bayes factors greater than 3 are considered substantial evidence in favor of the alternative hypothesis, Bayes factors less than 1/3 are considered substantial evidence in favor of the null hypothesis (see Wetzels et al., 2011). Bayes factors for the space-structured vs. unstructured comparison provided substantial evidence for the alternative hypothesis (BF=25.47).

I also coded participants' responses during debriefing to identify whether they spontaneously identified either a verbal rehearsal (e.g., "I said the shapes in the order in my head") or spatial (e.g., "If I forgot the pattern, I would try to remember the locations") strategy. Of the 24 participants, 15 explicitly indicated the use of a verbal rehearsal strategy, whereas only 2 participants indicated the use of a spatial strategy. Therefore, these results are unlikely to results from an explicit spatial strategy.

Results from Experiment 4.1b can be seen in Figure 4.2 (panels C & D). As is evident from the figure, accuracy was generally higher for space-structured trials ($M$=.86, $SD$=.07) compared to color-structured trials ($M$=.78, $SD$=.10), and this effect was equally pronounced at all set sizes. Indeed, a repeated measures ANOVA revealed a main effect of set size ($F[2,30]=18.71$, $p<.001$, $\eta_p2=.56$), a main effect of trial type ($F[1,15]=13.82$, $p=.002$, $\eta_p2=.48$), and no interaction between the two ($F[2,30]=.45$, $p=.64$, $\eta_p2=.03$). Post-hoc tests confirmed that accuracy was higher for space-structured trials than color-structured trials ($t[15]=3.75$, $p=.002$, $d=.94$, BF=21.83), and that accuracy was higher for set size 5 than 6 ($p=.01$), and higher for set size 6 than 7 ($p=.005$).

Once again, I coded participants' responses during debriefing to identify whether they spontaneously identified either a verbal rehearsal or spatial strategy. Of the 16 participants, 16 explicitly indicated the use of a rehearsal strategy (though one of these 16 reported rehearsing musical notes rather than verbal information), whereas only 1 participant indicated the use of a spatial strategy.

These experiments provide converging evidence that spatial structure benefits working memory even compared to another, matched type of structure (in that the color-structured condition, like the space-structured condition, had four options). Experiment 4.1b demonstrates that this is true even when participants are not cued to think about the structure at all. In fact, only 7 of the 16 participants reported noticing anything about the structure of the sequences during debriefing, and only 3 of those 7 believed that structure had anything to do with what was being tested (while 14/16 participants in the experiment showed an effect of spatial structure).

Notably, this task allows participants to verbally rehearse. Although from some perspectives this could defeat the point of studying *visual* working memory (but see 'visuospatial bootstrapping'; Darling et al., 2017), this is a strength of the present task. Given that participants *could* rehearse verbally (and they clearly did), an effect of spatial structure is especially notable. This spatial information is affecting visual working memory in spite of its irrelevance and in spite of participants' explicit engagement with verbal working memory, theorized to be a different sub-system (see, e.g., Allen et al., 2015; Morey, 2018). This pattern of results suggests one of two things: (1) our minds are capable of recruiting visual and verbal working memory simultaneously, as needed, or (2) visual working memory is *automatically* engaged (at least when there is salient, even if task-irrelevant, spatial information), and spatial structure boosts working memory even when participants are not explicitly relying on this information.

Like 'visuospatial bootstrapping', the results here speak to communication between verbal and visual information systems; however, *unlike* visuospatial bootstrapping, here there are effects of short-term spatial mappings that are not stored in long-term memory (in contrast with 'bootstrapping' effects; see Darling et al., 2012). The relation between this these effects and visuospatial bootstrapping will be explored further in the following experiments.

## 4.4 Experiments 4.2a and 4.2b: Space vs. Color vs. Unstructured

The previous results may be understood in one of several ways. For example, the results could be explained as a *benefit* of spatial structure or as a *decrement* of color structure for shape working memory. Alternatively, it could be that *both* spatial and color structure benefit working memory for shapes, but that spatial structure benefits working memory *more*. Here, I tested space and color structure vs. an unstructured baseline where both location and color were randomized.

### 4.4.1 Method

These experiments were identical to Experiment 4.1 except as noted. 12 participants completed Experiment 4.2a in exchange for course credit; this sample size was chosen based on sub-sampling of data from Experiments 4.1a and 4.1b and was pre-registered. However, anonymous reviewers raised concerns about the small sample size of Experiment 4.2a. As such, in Experiment 4.2b, I collected usable data from 158 participants via Amazon Mechanical Turk to reach the desired pre-registered sample size of 120 participants who met all of the inclusion criteria; an additional 19 participants were excluded for failing an attention check or failing to complete the correct number of trials (see below). Unlike Experiments 4.1a and 4.1b, these experiments included a third condition, in which both space and color were unstructured. As a result, Experiment 4.2a had 54 trials, 18 trials (3 difficulties [5, 6, or 7 shapes] × 6 unique trials) in each block. Experiment 4.2b had exactly half that many trials, to accommodate constraints imposed by the online format of the task. In Experiment 4.2b, the shapes appeared for 500ms, with 1000ms between presentations.

For the online experiment (4.2b), filters and checks were included to ensure high-quality data. At the outset, participants were eligible to complete the task if they (a) had an approval rate on Mechanical Turk greater than 98%, (b) lived in the United States, and (c) had completed at least 500 tasks. Participants were excluded prior to data analysis based on an attention check at the end of the task, in which participants were asked which shapes they saw (of 6 options). They had to select all that applied. Participants were excluded if they missed two or more items. Participants were also excluded if they failed to complete the task correctly (i.e., if they did not finish, or if they restarted partway through). The pre-registered analysis plan also stated that I would analyze the data before and after excluding participants with at least 50% accuracy overall; this was to ensure that there was a high-powered sample with performance comparable to what was observed in a laboratory setting.

### 4.4.2 Results & Discussion

Results from Experiment 4.2a can be seen in Figure 4.3 (panels A & B). Accuracy was generally higher for space-structured trials ($M$=.88, $SD$=.09) compared to color-structured trials ($M$=.80, $SD$=.14) and unstructured trials ($M$=.77, $SD$=.15). Indeed, a repeated measures ANOVA revealed a main effect of set size ($F$[2,22]=4.60, $p$=.02, $\eta_p 2$=.30), a main effect of trial type ($F$[2,22]=8.15, $p$=.002, $\eta_p 2$=.43), and no interaction between the two ($F$[4,44]=1.80, $p$=.15, $\eta_p 2$=.14). Post-hoc tests confirmed that accuracy was higher for space-structured trials than both color-structured trials ($t$[11]=2.40, $p$=.04, $d$=.69, BF=2.16) and unstructured trials ($t$[11]=3.93, $p$=.002, $d$=1.13, BF=19.43), whereas color-structured trials and unstructured trials did not differ ($t$[11]=1.43, $p$=.18, $d$=.41, BF=.65). Of the 12 participants, 11 explicitly indicated the use of a verbal rehearsal strategy, whereas only 1 observer indicated the use of a spatial strategy.

Results from Experiment 4.2b can be seen in Figure 4.3 (panels C & D; results shown are from the final sample of 120 participants, after exclusion based on accuracy). Per the pre-registered analysis plan, I separately analyzed the data including and excluding participants with overall task accuracy greater than 50%; this was to account for the fact that Amazon Mechanical Turk pilot data revealed worse overall performance than the in-lab sample. First, I report analyses on the set of 158 participants who passed the attention checks, prior to the accuracy exclusion. Accuracy was generally higher for space-structured trials ($M$=.70, $SD$=.21) compared to color-structured trials ($M$=.66, $SD$=.21) and unstructured trials ($M$=.67, $SD$=.20). A repeated measures ANOVA revealed a main effect of set size ($F$[2,314]=45.00, $p$<.001, $\eta_p^2$=.22), a main effect of trial type ($F$[2,314]=6.64, $p$=.002, $\eta_p$2=.04), and no interaction between the two ($F$[4,628]=2.34, $p$=.05, $\eta_p$2=.02). Replicating Experiment 4.2a, post-hoc tests confirmed that accuracy was higher for space-structured trials than both color-structured trials ($t$[157]=3.80, $p$<.001, $d$=.30, BF=77.66) and unstructured trials ($t$[157]=2.43, $p$=.016, $d$=.19, BF=1.55), whereas color-structured trials and unstructured trials did not differ ($t$[157]=.89, $p$=.38, $d$=.07, BF=.13).

Next, I report analyses for the final set of 120 participants who met the accuracy inclusion criteria (>50%). Accuracy was generally higher for space-structured trials ($M$=.80, $SD$=.14) compared to color-structured trials ($M$=.76, $SD$=.15) and unstructured trials ($M$=.76, $SD$=.14). A repeated measures ANOVA revealed a main effect of set size ($F$[2,238]=68.91, $p$<.001, $\eta_p$2=.37), a main effect of trial type ($F$[2,238]=6.58, $p$=.002, $\eta_p$2=.05), and no interaction between the two ($F$[4,476]=1.37, $p$=.24, $\eta_p$2=.011). Again, post-hoc tests confirmed that accuracy was higher for space-structured trials than both color-structured trials ($t$[119]=3.53, $p$<.001, $d$=.32, BF=33.13) and unstructured trials ($t$[119]=2.79, $p$=.006, $d$=.25, BF=4.02), whereas color-structured trials and unstructured trials did not differ ($t$[119]=.39, $p$=.70, $d$=.04, BF=.11). Online participants were not asked about their strategies in the task.

These experiments provide converging evidence with Experiment 4.1 that spatial structure benefits working memory. Here, these results clarify what kinds of structure matter. For example, it could have been the case that both space-structure and color-structure improve shape working memory but that space-structure does so to a larger extent. Alternatively, it could have been that space-structure does not benefit shape working memory, but that color-structure somehow interferes with shape working memory. However, it seems that neither of these accounts are true. Instead, spatial structure benefits working memory whereas there is no evidence of a color-structure benefit: although performance in the color-structure condition was numerically higher than the unstructured condition in Experiment 4.2a, it was actually *lower* in Experiment 4.2b (which had a sample size 10 times greater). This coheres with other working suggesting a privileged status of spatial information in working memory (e.g., Pertzov & Husain, 2014).

## 4.5 Experiment 4.3: Space vs. Sound vs. Unstructured

The previous results establish that spatial structure benefits working memory — but is space special? One possibility is that *many* kinds of structure (i.e., repetition) benefit working memory, and that color simply isn't a salient or valuable kind of structure. Here, I compared spatial structure to *audio* structure. In other words, the relevant block featured ‘audio-structured’ trials in which any repeating shape was paired with the same tone each time. Is there still a greater benefit to spatial structure?

### 4.5.1 Method

This experiment was identical to Experiment 4.2 except as noted. 18 participants completed this experiment in exchange for course credit. This sample size was pre-registered and was chosen to be approximately identical to Experiment 4.1b (but rounded to a different number to account for a difference in the number of conditions). Instead of a color-structured condition, there was an audio-structure condition in which each shape was paired with a tone of a specific note. To match the number of locations, there were four possible notes: ‘A’, ‘C’, ‘E’, or ‘G’. Matching Experiment 4.2a, there were 54 trials, 18 trials (3 difficulties [5, 6, or 7 shapes] × 6 unique trials) in each block. Due to the difficulty of administering audio experiments online (e.g., an inability to ensure that participants have their audio turned on, etc.), an online replication was not conducted.

### 4.5.2 Results & Discussion

Results from Experiment 4.3 can be seen in Figure 4.3 (panels E & F). As is evident from the figure, accuracy was generally higher for space-structured trials ($M$=.84, $SD$=.08) compared to audio-structured trials ($M$=.79, $SD$=.12) and unstructured trials ($M$=.80, $SD$=.12). Indeed, a repeated measures ANOVA revealed a main effect of set size ($F$[2,34]=18.89, $p$<.001, $\eta_p 2$=.53), a main effect of trial type ($F$[2,34]=3.47, $p$=.04, $\eta_p 2$=.17), and no interaction between the two ($F$[4,68]=2.46, $p$=.05, $\eta_p 2$=.13). Post-hoc tests confirmed that accuracy was higher for space-structured trials than both audio-structured trials ($t$[17]=2.81, $p$=.01, $d$=.66, BF=4.49) and unstructured trials ($t$[17]=2.28, $p$=.04, $d$=.54, BF=1.89), whereas audio-structured trials and unstructured trials did not differ ($t$[17]=.35, $p$=.73, $d$=.08, BF=.26). I also coded participants' responses during debriefing to identify whether they spontaneously identified either a verbal rehearsal or spatial strategy. Of the 18 participants, 16 explicitly indicated the use of a verbal rehearsal strategy, whereas none indicated the use of a spatial strategy.

This experiment provides converging evidence with Experiments 4.1 and 4.2 that spatial structure selectively benefits spatial working memory and further demonstrates that this benefit of structure is *unique*: neither equivalent color structure nor audio structure yielded similar benefits. Note that some other paradigms, such as those used in 'visuospatial bootstrapping' experiments (see Darling et al., 2017), are not readily adaptable to comparing spatial structure with other kinds of structure. In this way, the present set of studies are an extension of that research program. For example, this clarifies that the 'visuospatial' aspect of visuospatial bootstrapping is uniquely important because there seems to be no evidence for an effect of 'audio' bootstrapping. Prior work supporting the 'spatial interference' view also employed primarily visual controls; in this way, the audio control of this experiment extends that research program, as well, by showing that spatial information is uniquely beneficial in working memory, not just compared to other visual cues, but compared to information in other modalities. Finally, this control is an especially strong one. Prior work has shown that tones can themselves be spatialized (Lidji et al., 2007); therefore, one may have expected audio structure to be more useful than color structure. Nevertheless, spatial structure is unique: audio structure did not improve retention in working memory.

## 4.6 Experiment 4.4: What structure matters?

Experiments 4.1-3 demonstrate a benefit of short-term spatial, but not color or audio, structure on working memory. But why? In previous work investigating the role of space in working memory, spatial overlap often results in memory interference (Pertzov & Husain, 2014; Treisman & Zhang, 2006). In other words, items appearing in the same location were remembered *worse* than items that appeared in different locations (but had some other overlapping feature, like color). However, many previous paradigms were unable to decouple spatial interference from spatial structure. The present paradigm has several features that enable decoupling. (1) Participants remember sequences comprised of a small set of recurring items, and (2) These items belong to distinct categories (as opposed to something like oriented lines). So, here I asked: is the effect of spatial structure caused by the *presence* of structure (i.e., the fact that any given object appears in a consistent location) or the *absence* of overlap (i.e., the fact that no two objects appear in the same location)?

To test this difference, I created two opposing conditions — an 'overlapping' condition in which different items (e.g., circle and pentagon) always appear in consistent locations but may overlap with each other, and a 'separate' condition in which different items may appear in multiple locations but will *never* overlap with each other (see Figure 4.4, panels A & B). According to interference accounts, memory performance should be higher in the separate condition; although shapes appear in many unique locations, no two shapes ever overlap with one another (and thus never interfere with each other). Alternatively, the opposite might be true: the presence of structure might drive memory performance, and what matters is not whether items overlap with each other, but whether they are consistent with themselves (i.e., whether the circles always appear in one location, regardless of where the other shapes appear). This pattern of results may would be more consistent with visuospatial bootstrapping, although such studies have never tested different items overlapping in one location.

### 4.6.1 Method

This experiment was identical to Experiment 4.2 except as noted. 18 participants completed this experiment in exchange for course credit. The color-structured and space-structured conditions were replaced with two new conditions. In a 'separate' condition, shapes could appear in any location, but no two unique shapes ever appeared in the same location (on a given trial). Although the locations were partially constrained by the shapes, there was no 'spatial structure' because the shapes did not appear in stable locations across presentations. Conversely, in an 'overlapping' condition, each shape *always* appeared in the same location, and two of the three shapes *always* overlapped with each other. To maximize the difference between the 'separate' and 'overlapping' conditions, the display was altered so that there were 6 locations (black squares) instead of 4. They were arranged in a hexagonal structure, all roughly 10.23° from the center of the screen. To account for the two new locations, there were two new colors: the default 'purple' and 'orange' in PsychoPy.

### 4.6.2 Results & Discussion

Results from Experiment 4.4 can be seen in Figure 4.4. As is evident from the figure, accuracy was generally slightly higher for 'overlapping' trials ($M$=.85, $SD$=.06) compared to both 'separate' trials ($M$=.80, $SD$=.10) and unstructured trials ($M$=.80, $SD$=.10). A repeated measures ANOVA revealed a main effect of set size ($F[2,34]=22.56$, $p<.001$, $\eta_p 2=.57$), a main effect of trial type ($F[2,34]=4.46$, $p=.02$, $\eta_p 2=.21$), and no interaction between the two ($F[4,68]=.04$, $p=.99$, $\eta_p 2=.003$). Post-hoc tests confirmed that accuracy was higher for 'overlapping' trials than both 'separate' trials ($t[17]=2.25$, $p=.038$, $d=.53$, BF=1.80) and unstructured trials ($t[17]=3.11$, $p=.006$, $d=.73$, BF=7.73), but no difference between 'separate' and unstructured trials ($t[17]=.22$, $p=.83$, $d=.05$, BF=.25). Similar to the previous experiments, most of the participants (15/18) explicitly reported a verbal rehearsal strategy, and no participants explicitly reported a spatial strategy.

In the previous experiments, there was a robust effect of spatial structure (as compared to color structure, audio structure, and no structure). Here, I asked what *kind* of structure matters. Specifically, I asked whether the effects of spatial structure were caused by the *presence* of structure (as in the 'overlapping' condition) or the *absence* of overlap (as in the 'separate' condition). I found that performance was better in the 'overlapping' condition, suggesting that the benefit seen in prior experiments may have been due to the presence of structure rather than the absence of overlap.

The findings in this experiment are different from both 'spatial interference' and 'visuospatial bootstrapping' effects. For example, the spatial interference account predicts that objects overlapping in space should lead to memory impairments; however, this is not what was observed. Exactly the opposite, these results show that the condition in which shapes overlapped had the best memory performance. Similarly, visuospatial bootstrapping makes no specific predictions about what should be expected in the overlapping vs. separate conditions. The predictions of this view depend on what is being bootstrapped, and what it is being bootstrapped to. One possible prediction could have been that each shape needs to be bootstrapped to a single, unique location. In this case, one may have predicted equal performance across all three conditions (because in the overlapping condition, multiple shapes are bound to the same location, and in the separate condition, individual shapes are bound to multiple locations). However, this is not what is observed. Instead, the present results clarify the process of visuospatial bootstrapping: memory benefits from binding information to *specific* locations, but not necessarily to *unique* locations.

However, the key difference in this experiment (between the overlapping vs. separate conditions) could be explained by a difference in the number of locations used across conditions. Notably, the 'overlapping' condition only ever utilized 2 of the 6 locations, whereas the 'separate' condition could have utilized up to 6 locations. Thus, the effect of spatial structure could be explained by *attention* to that structure (i.e., participants can focus on a subset of locations in the 'overlapping' condition, thus reducing attentional demands), rather than the underlying structure *per se*. Regardless of the underlying mechanism, these results run counter to predictions of interference accounts (e.g., Pertzov & Husain, 2014) and are unexplained by accounts that emphasize long-term spatial associations (as in visuospatial bootstrapping, e.g., Darling et al., 2017; see also Darling et al., 2020). Future work could adopt a similar approach to the one taken here so as to further probe the inference account and better understand the scope of visuospatial bootstrapping.

## 4.7 Exp 4.5: How robust is the effect of spatial structure?

Experiment 4.4 address two different kinds of spatial structure. However, one key difference between conditions was the number of locations participants had to attend to: in the 'overlapping' condition, in which accuracy was highest, participants had to attend to only 2 of the 6 locations, whereas in the 'separate' condition, they had to attend to all of the possible locations. The same is true, though to a lesser degree, of the previous experiments. Given the nature of the spatial structure manipulation, participants noticing the structure could realize they need only to attend to 3 of the 4 locations. It is possible that this enhanced attention to 3 of the 4 locations explains the effect of spatial structure observed so far. Here, I address this possibility (as well as other methodological details) to provide a stronger test of the 'spatial structure' account. I make three key changes to the task: (1) the number of locations and colors was reduced to 3, to match the number of shapes; (2) the first three items on each trial always had a unique color, unique location, *and* a unique shape (previously, only the shapes had to be unique, except in the space-structured and color-structured conditions in which location and color would also be unique); and (3) the locations were arranged in a line, rather than in a grid format (to account for the fact that spatial information contained two dimensions, perhaps providing an advantage over the other information types).

If the previously observed effects in the space-structured conditions were caused by one or more of these three factors, then we should not expect to observe an effect of spatial structure here. Similarly, if the lack of effects in the color structure conditions were caused by the lack of these advantages (i.e., the predictability of the first few items, or the fewer locations one needed to attend to), then we should expect to observe an effect of color structure.

### 4.7.1 Method

This experiment was identical to Experiment 4.2b except as noted. Usable data was collected from 158 participants via Amazon Mechanical Turk (to reach the desired pre-registered sample size of 120 participants who met all of the inclusion criteria); another 63 participants were excluded for failing an attention check or failing to complete correct number of task trials. Note: the majority of these exclusions came from participants who completed *extra* trials, ostensibly because they refreshed the task halfway through. Because there is not a way to know *why* they refreshed the task, all such participants are excluded. There were three key changes to the task, detailed below. All three of these changes were made to better equate the information presented across conditions.

First, the number of locations and colors was reduced to 3, to match the number of shapes. This means that, across all three conditions, participants would now only have to attend to 3 locations. Previously this was not the case. In the space-structured conditions, participants would have to attend to only 3 locations, whereas in the color-structured and unstructured conditions, participants would have to attend to up to 4 locations.

Second, the first three items on each trial always had a unique color, unique location, *and* a unique shape. Previously this was not the case. In Experiments 4.2a and 4.2b, the first three shapes were always unique, but the first three colors/locations differed across conditions. In the space-structured condition, for example, the first three locations would be unique, but the first three colors would be random (and could include repeats). The opposite was true for the color-structured condition. However, this meant that certain items appeared in slightly more predictable locations. Consider a trial in the space-structured condition. If a shape appeared in Location #1, the participant would then know that they need only to attend to Locations #2, #3, and #4 to see where the next shape will appear. If the second shape appeared in Location #2, then the participant would know that they need only to attend to Locations #3 and #4 to see where the third shape will appear. By contrast, in the color-structured condition, any object could appear at any location at any time. By ensuring that shape, location, and color were unique for the first three items, the location of each shape was equally predictable across conditions. However, this also means that space and color structure were partially confounded (because the first three items are always 'structured'), meaning that we may expect reduced effects overall.

Third, the three targets were arranged in a line, rather than in a grid format. This is to account for the fact that the spatial information contained two-dimensions, whereas the color information did not.

### 4.7.2 Results & Discussion

Results from Experiment 4.5 can be seen in Figure 4.5 (results shown are from the final sample of 120 participants, after exclusion based on accuracy). Per the pre-registered analysis plan, I separately analyzed the data including and excluding participants with accuracy greater than 50%; this was to address the fact that Amazon Mechanical Turk pilot data revealed worse overall performance than the in-lab sample. First, I report analyses on the set of 158 participants who passed the attention checks and completed the correct number of trials prior to the accuracy exclusion. Accuracy was generally higher for space-structured trials

($M$=.69, $SD$=.21) compared to color-structured trials ($M$=.67, $SD$=.21) and unstructured trials ($M$=.65, $SD$=.22). A repeated measures ANOVA revealed a main effect of set size ($F$[2,314]=41.99, $p$<.001, $\eta_p 2$=.02), a main effect of trial type ($F$[2,314]=5.00, $p$=.007, $\eta_p 2$=.003), and no interaction between the two ($F$[4,628]=.54, $p$=.71, $\eta_p 2$<.001). Post-hoc tests confirmed that accuracy was higher for space-structured trials than unstructured trials ($t$[157]=3.14, $p$=.002, $d$=.25, BF=9.70). However, space-structured trials and color-structured trials ($t$[157]=1.51, $p$=.13, $d$=.12, BF=.27) as well as color-structured trials and unstructured trials ($t$[157]=1.67, $p$=.10, $d$=.13, BF=.34) did not differ.

Next, I report analyses for the final set of 120 participants who met the accuracy inclusion criteria (>50%). Accuracy was generally higher for space-structured trials ($M$=.78, $SD$=.14) compared to color-structured trials ($M$=.76, $SD$=.14) and unstructured trials ($M$=.74, $SD$=.17). A repeated measures ANOVA revealed a main effect of set size ($F$[2,238]=42.98, $p$<.001, $\eta_p 2$=.27), a main effect of trial type ($F$[2,238]=3.76, $p$=.025, $\eta_p 2$=.03), and no interaction between the two ($F$[4,476]=.59, $p$=.67, $\eta_p 2$=.005). Post-hoc tests confirmed that accuracy was higher for space-structured trials than unstructured trials ($t$[119]=2.76, $p$=.007, $d$=.25, BF=3.78), however space-structured trials and color-structured trials ($t$[119]=1.07, $p$=.29, $d$=.10, BF=.18) as well as color-structured trials and unstructured trials ($t$[119]=1.63, $p$=.11, $d$=.15, BF=.37) did not differ. Thus, Bayes factors provided substantial evidence in favor of the alternative hypothesis for the space-structured vs. unstructured contrast, substantial evidence in favor of the null hypothesis for the space-structured vs. color-structured contrast, and moderate evidence in favor of the null hypothesis for the color-structured vs. unstructured contrast. Online participants were not asked about their strategies.

This experiment was designed to provide a strong test of the spatial structure account, by equating the information presented across conditions as much as possible. Despite this, there is nevertheless a robust effect of spatial structure, but no corresponding effect of color structure.

Interestingly, the spatial-structure and color-structure conditions did not differ from one another. There are two ways to interpret this null effect. First, color structure does benefit working memory performance, despite the null effect in Experiments 4.2a, 4.2b, and here, and that this study underpowered to detect such an effect. Yet this experiment and Experiment 4.2b have a combined 316 participants, and there was no reliable effect. This would be extremely unlikely if color-structure had a true effect. Second, color structure does *not* benefit working memory performance, but performance in the Experiment 4.5 color-structured condition benefits from the spatial structure of the first three items. One of the key changes made in this experiment, in contrast with Experiments 4.2a and 4.2b, was the fact that the first three items always have a unique color and location. In this way, there will be a non-trivial number of trials in the color-structured condition, which have spatial structure (because, by random chance, shapes later in the sequence may appear in their initial locations; this will be more likely for the lower set sizes). This could cause smaller differences between conditions overall (compared to Experiment 4.2b). Note, however, that Experiments 4.2a and 4.2b *were* explicitly designed to better de-confound these conditions. In those cases, effects between conditions were much larger. Looking ahead, future work can independently test the effect of each factor (the number of spatial dimensions, the number of possible locations, the predictability of each item's location) on working memory maintenance.

## 4.8 General Discussion

These experiments investigated the interactions between spatial and verbal information in working memory, as well as the specific role of spatial information in working memory. Across all seven experiments, 'spatial structure' improved memory. This enhancement was true both when participants were cued to this structure (Experiment 4.1a) and when they were not (Experiment 4.1b, 4.2a, 4.2b, and 4.5). The benefit of space is unique (compared to other visual features, like color, and also auditory features, like tones; Experiment 4.3). Further, these effects may be best understood as caused by the *presence* of structure, rather than an *absence* of overlap (Experiment 4.4). Finally, these effects are generally robust to various changes in the experimental design (Experiment 4.5).

All told, these data inform several aspects of working memory. *First*, these experiments illustrate that although people seem to have the default tendency to explicitly engage in verbal rehearsal (i.e., 83% of participants who were queried reported a rehearsal strategy), they nevertheless benefit from spatial structure; this may be seen as broadly consistent with the fact that visuospatial bootstrapping also does not depend on 'executive resources' (Calia et al., 2019). *Second*, these experiments support the view that visuospatial and verbal information interact to facilitate memory. Despite the irrelevance of spatial information, it influenced memory (as in visuospatial bootstrapping; Darling et al., 2017). In this way, these findings also constrain working memory models. For example, they are potentially at odds with the view that working memory is supported by two independent subsystems for verbal and visual working memory (e.g., Baddeley, 1992). *Third*, these experiments show that space is special compared to seemingly equivalent *visual* features, like color (see also Pertzov & Husain, 2014), and also compared to equivalent non-visual features, like auditory tones. *Fourth*, these experiments shed light on *how* space influences memory. Whereas previous work emphasized either 'spatial interference' and the role that space plays in binding features to objects (Pertzov & Husain, 2014; Treisman & Zhang, 2006), or 'visuospatial bootstrapping' and the way that spatial information facilitates memory (Darling et al., 2017), the present work asks specifically about spatial *structure*. Here, the presence of spatial structure, not the absence of overlap, influences working memory (seemingly in contrast with prior accounts). The results presented here are not mutually exclusive with either spatial interference accounts or visuospatial bootstrapping, but they do present some conflicting results. For example, an interference account should predict that memory performance in the two key conditions in Experiment 4.4 (separate vs. overlapping) was equal, yet that is not what was observed; clearly the absence of overlap alone is not all that influences memory. In contrast, the data here imply that stability (or, structure) benefits working memory, regardless of overlap. Furthermore, visuospatial bootstrapping has been said to depend on associations in long-term memory (Darling et al., 2012; but see Darling et al., 2020); however, the present effects could be seen as effects of visuospatial bootstrapping that do not depend on long-term spatial associations. The findings in prior work may not be because bootstrapping depends on long-term associations, but rather than long-term associations interfered with the ability to make new associations over that same configuration (as when rearranging a keypad, for instance; Darling et al., 2012).

There is one key difference between this account and a spatial interference account. In Experiment 4.4, items overlapping with one another (in the 'overlapping' condition) are better remembered than items that never overlap (in the 'separate' condition). At first blush, these results seem directly at odds with a spatial interference account. Consider, for example, the view that features are bound to objects via space (e.g., Treisman & Zhang, 2006). What should be made of better memory for overlapping objects? Perhaps both accounts are correct and the 'interference' predicted in this task does not involve the shapes themselves but instead involves the binding of color and shape. For example, if there had been a separate measure of color memory performance, participants would have worse color memory (but better shape memory) in the overlapping condition. Yet even if that were true, a spatial interference account would not necessarily predict better memory for shapes in the overlapping condition. In this way, some of the results presented here highlight gaps in our understanding of classic location binding results (see also Pertzov & Husain, 2014; Rajsic & Wilson, 2014) — raising questions about the underlying mechanisms of feature binding and motivating future work.

Another possibility is that these results do not contradict the basic idea of spatial interference, but instead speak to many different *kinds* of possible interference. Various types of interference have been conceptualized, each with unique consequences. For example, interference by feature overwriting predicts that similar items are more likely to interfere with one another, whereas interference by superposition predicts that different items are more likely to interfere with one another (Oberauer et al., 2016). In other words, it is possible that the seemingly different patterns of results observed in these studies compared to previous interference studies may come down to the details of the stimuli themselves. For example, maybe "circle" and "pentagon" are more similar to the mind than "blue" and "green", resulting in less interference for shape than color. This highly speculative possibility can be addressed by future work quantifying the relative similarity of different stimuli.

These results are generally more consistent with the idea of visuospatial bootstrapping (Darling & Havelka, 2010), although with a few key differences. First, unlike classic bootstrapping designs, the patterns here do not depend on long-term spatial mappings; the 'spatial structure' defined here is confined to, at most, a ~10s trial (but see Darling et al., 2020). Second, here there are test cases where items not only possess stable spatial mappings, but also *share* stable spatial mappings with other objects (Experiment 4.4). While these results are not necessarily at odds with visuospatial bootstrapping, they provide unique insight not obvious in the canonical bootstrapping designs. Third, whereas visuospatial bootstrapping studies often focus on a contrast between spatial structure vs. no structure, these studies *compare* spatial structure to two other forms of structure (color structure and audio structure). In so doing, it is clear that the effect of spatial structure is highly specific. Concretely, this means that the term 'visuospatial bootstrapping' may be an apt name; these results demonstrate, e.g., that visual information alone (in the form of color information) does not result in the same form of bootstrapping. Finally, the results show that the notion of 'spatial structure' (whether in the form of visuospatial bootstrapping, or otherwise) is highly robust: the key effect is replicated seven separate times, after numerous critical manipulations to the basic paradigm.

### 4.8.1 Possible Explanations

Three possible explanations for these data come to mind. The first is *simplicity*: perhaps performance in the space-structured condition was better than performance in the unstructured and color-structured condition because the task was easier. One may note, for example, that in Experiments 4.1-3, all four of the locations were never used in a single space-structured trial (because there were fewer shapes than locations), whereas all four locations could potentially be engaged in a single color-structured or unstructured trial (because no constraints were imposed on spatial location). However, Experiment 4.5 rules out this possibility: when the number of shapes and locations was equated, such that it all three conditions participants would always have to attend to all of the visible locations, there was nevertheless a benefit of spatial structure relative to no structure. Although questions remain about how exactly spatial structure benefits working memory, Experiment 4.5 demonstrates that the primary result reported here (a difference between spatial structure and no structure) cannot be explained solely by a difference in the number of locations.

The second possible explanation is *predictability*: could these results be explained by some difference in the predictability of the sequences? One may note, for example, that the location of the second and third shapes in the space-structured conditions were partially predictable; if spatial information is structured, and participants are aware of that structure, then they know that the second and third shapes cannot appear where the prior shapes had. Although this possibility cannot be ruled out for the earlier experiments, Experiment 4.5 was specifically designed to equate predictability. There, the second and third items were equally predictable across conditions; nevertheless, there is a benefit of spatial structure relative to no structure.

The third possible explanation is *eye movements*. In this task, participants are free to move their eyes around the screen as they wish. Spatial location *per se* might not influence working memory but instead eye movements may lead directly to differential encoding. Indeed, eye movements do influence working memory, at least at retrieval (e.g., Awh et al., 1998; Theeuwes et al., 2009). Yet, even controlling eye gaze (and thus overt shifts of attention) by having subjects fixate would not rule out the possibility that covert shifts of spatial attention nevertheless could explain the present results. Thus, future work might characterize the effect of eye movements on the benefits of spatial structure for working memory maintenance.

### 4.8.2 Conclusion

Space may be *foundational* to working memory: not only does spatial structure benefit working memory, it does so even when that information is task-irrelevant, and even when participants rely on distinctly non-spatial strategies (e.g., verbal rehearsal). Further, spatial structure seems unique in its influence; neither non-spatial visual structure (color) nor non-visual auditory structure benefitted memory, even compared to an unstructured baseline. These results raise questions about the nature of working memory, its subsystems, and their interactions, whilst emphasizing the importance of spatial *structure* in working memory.

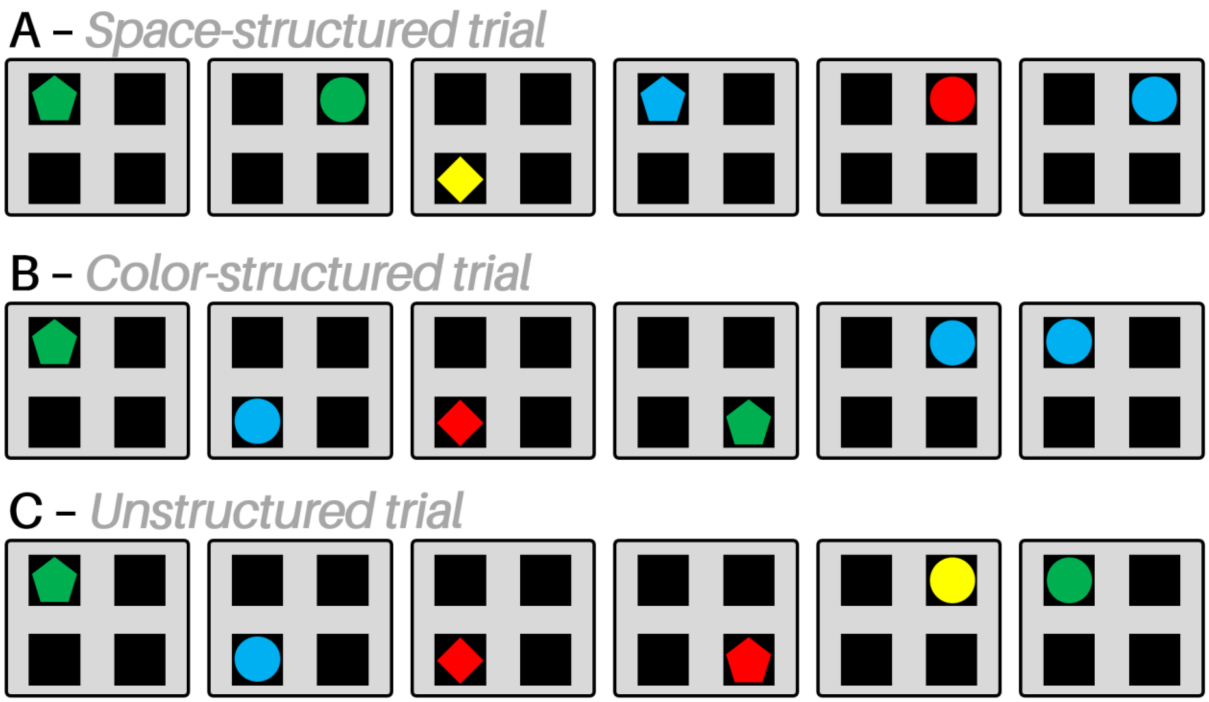



Figure 4.1. Schematic of the trial structure and three unique trial types. Shapes appeared one at a time with a brief ISI between shapes. (A) An example of a 'space-structured' trial. In this example, circles always appear in the top right corner, but their color is random. This condition was used for Experiments 4.1-3 and Experiment 4.5. (B) An example of a 'color-structured' trial. In this example, circles always appear in blue, but their location is random. This condition was used for Experiments 4.1-3 and Experiment 4.5. (C) An example of an unstructured trial. In this example, both colors and locations are random. This condition was used in Experiments 4.2-3 and Experiment 4.5.

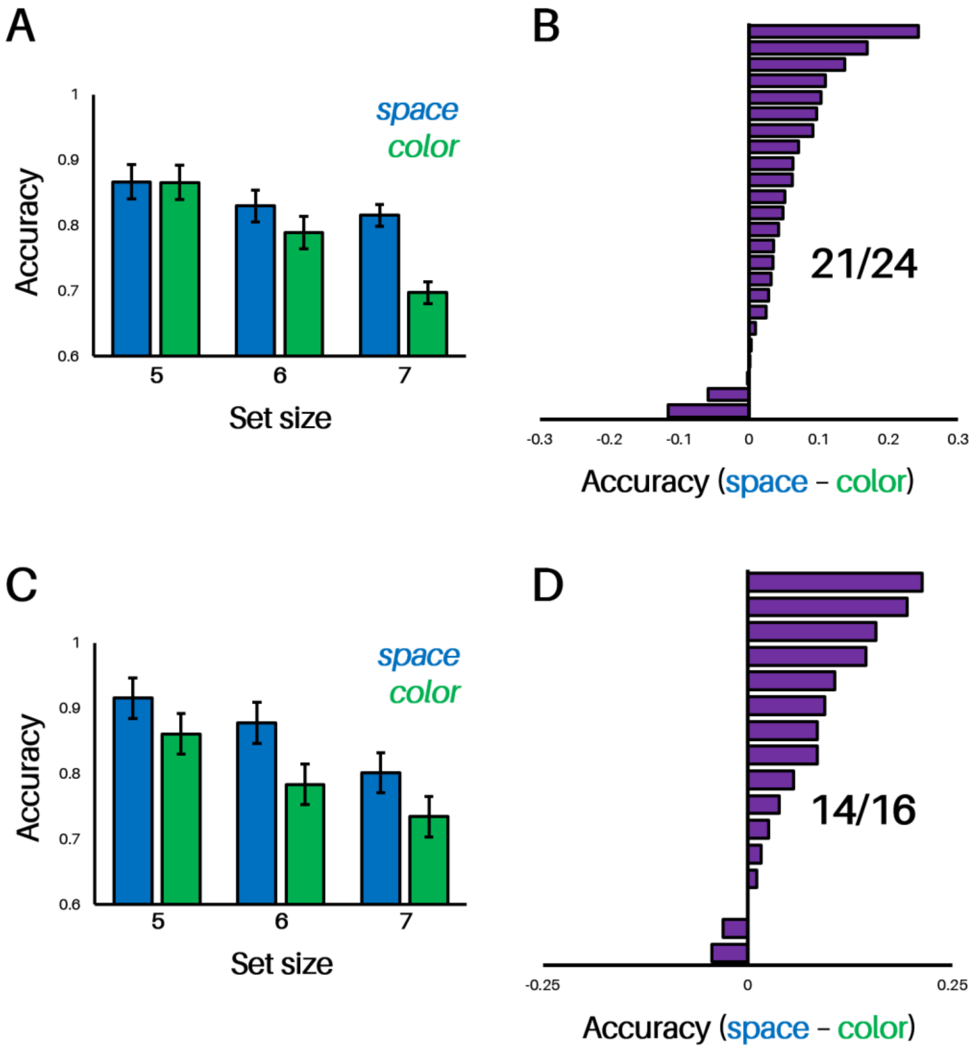



Figure 4.2. Results from Experiment 4.1a (A & B) and 4.1b (C & D). On the left (A & C) average accuracy is broken down set size and by condition. On the right (B & D), difference scores are shown between the space-structured and color-structured condition for each participant. The number of participants showing the predicted effect are shown within each figure. Error bars represent +/- 1 standard error.

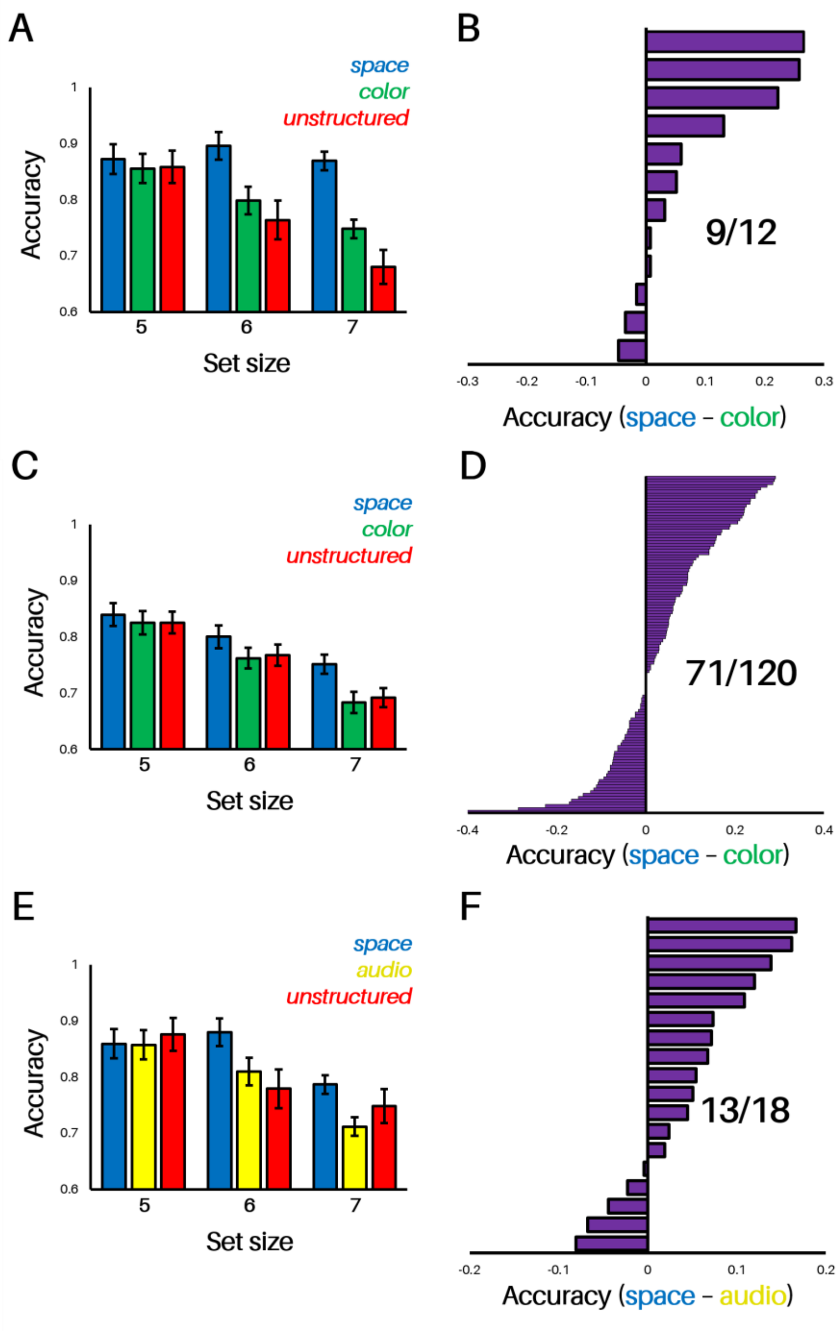



Figure 4.3. Results from Experiment 4.2a (A & B), Experiment 4.2b (C & D), and Experiment 4.3 (E & F). On the left (A & C & E) average accuracy is broken down by set size and by condition. On the right (B & D & F), difference scores are shown between the two most relevant conditions. The number of participants showing the predicted effect are shown within each figure. Error bars represent +/- 1 standard error.

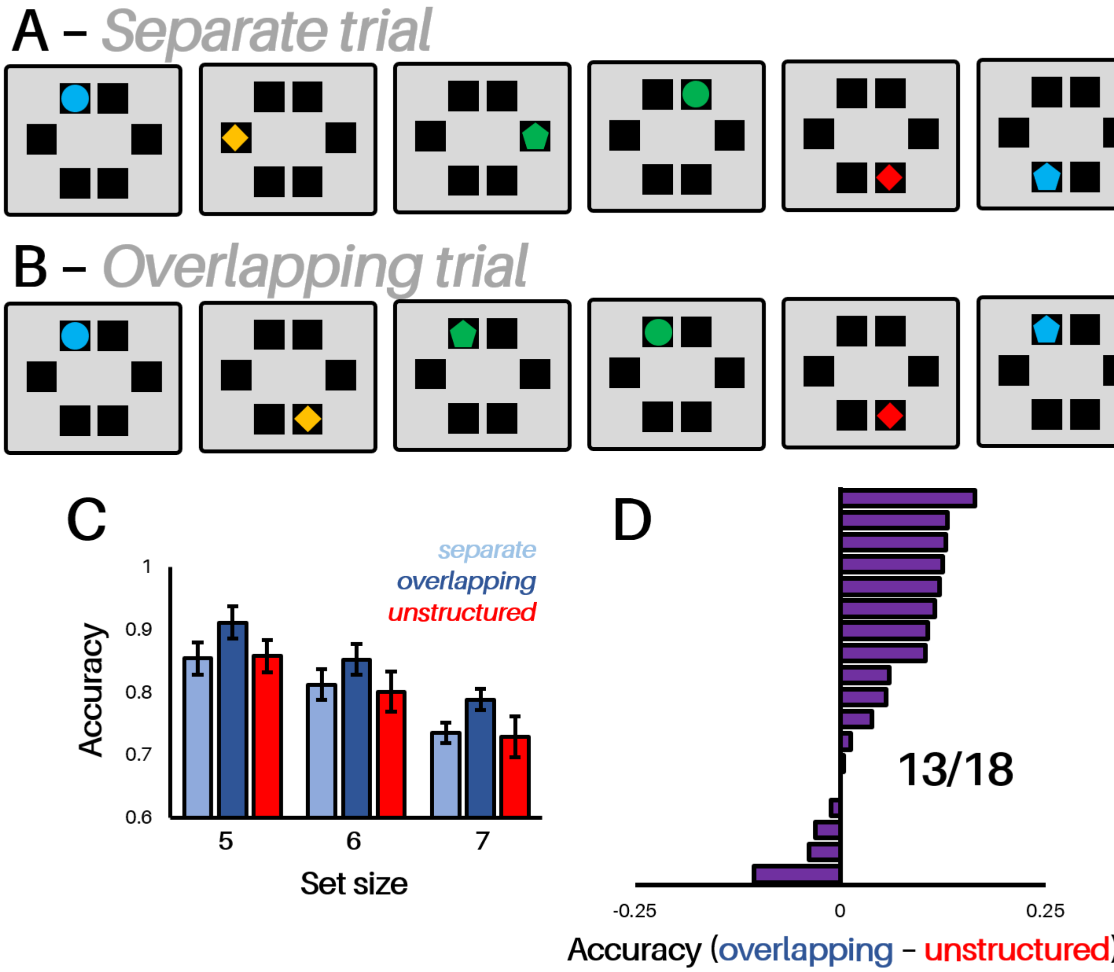



Figure 4.4. Schematic and results from Experiment 4.4. (A) A depiction of a 'separate' trial. In this example, each shape appears in different locations, but no two shapes ever overlap with one another. (B) A depiction of an overlapping trial. In this example, all shapes appear in the same locations each time they appear, but the circle and pentagon appear in the same location. (C) Average accuracy is broken down by set size and by condition. (D) Difference scores are shown between the two most relevant conditions. The number of participants showing the predicted effect are shown within the figure. Error bars represent +/- 1 standard error.

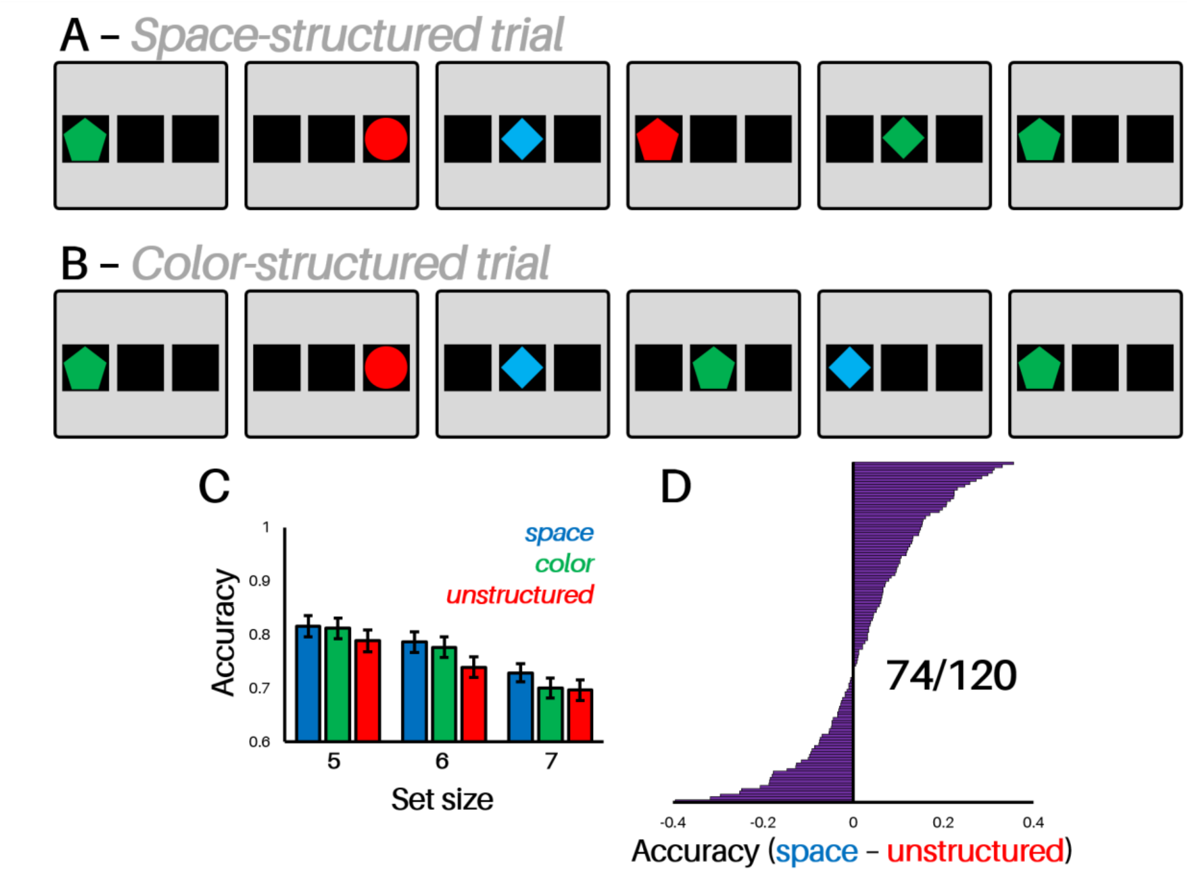



Figure 4.5. Schematic and results from Experiment 4.5. (A) A depiction of a space-structured trial. (B) A depiction of a color-structured trial. Here, unlike previous experiments, there were only three locations, and those three locations appeared in a line. Note that locations are randomized in the color-structured condition and colors in the space-structured condition are randomized, such that it is possible for items to repeat in the same location/color. (C) Average accuracy is broken down by set size and by condition. (D) Difference scores are shown between the space and unstructured conditions. The number of participants showing the predicted effect are shown within the figure. Error bars represent +/- 1 standard error.

# Chapter 5

# Discussion

This chapter contains text and/or materials from the following publications:

Yousif, S. R. (In press). Redundancy and reducibility in the formats of spatial representations. *Perspectives on Psychological Science.*

## 5.1 What does format buy us?

In this thesis, I have emphasized the notion of 'format'. But what do we need this notion for? Perhaps our goal should only be to describe behavior, without worrying about these ephemeral, theoretical constructs.

Discovering 'format', however, is not just a minor curiosity. Fundamentally, as cognitive scientists, our goal is to understand how the activity of neurons in the brain could possibly lead to complex behavior. No matter what models of the mind one favors, there is a simple fact that *at some point* between brain and behavior information must be represented in a specific way. If we want to understand that link, we must understand how that information is being represented. That is why we must think explicitly about format.

Concretely, we also need to understand spatial representation across modalities and domains. Often, those who study the format of visuospatial representations are siloed from those who study the format of motor representations, for example. Sometimes, however, we may want to understand whether the mind solves spatial tasks like these in general ways or not. Do we rely on similar coordinates across modalities, for example? Thinking about 'format' can help us to understand how the pieces of the mind interact; we can test whether similar 'formats' underlie behavior in different modalities.

We must also acknowledge that we are in the midst of a revolution: neuroscience is the future of our field. The answers to many of the questions we most care about at the end of the day are, ultimately, only going to be answered through neuroscience. But this does not mean that we should throw our hands in the air and become neuroscientists. Quite the opposite: I'd argue that the neuroscientists are going to need a lot of help knowing what they ought to look for in the first place.

Some of the greatest advancements in neuroscience in the past two decades have actually been about the very topic of this dissertation: spatial representations. Grid cells and place cells and head directions cells offer a glimpse at how complex behavior can be implemented at the level of neurons. The description of these cells is a triumph of human discovery — quite literally a Nobel-prize-winning achievement. Yet understanding grid cells and place cells alone will not bridge the brain-behavior gap. We know, from much of the work presented and discussed in this thesis, that there are other *forms* of spatial representations in the mind that are not straightforwardly captured by grid and place cells; they may be part of the answer, but they are *not* the answer.

The same is true when we think of other spatial properties like size and shape. Surely we want to understand how these properties are represented in the brain, but it may be vital that we first understand how they are formatted in the mind. If we went off searching for representations of 'true area' only to later find that 'true area' is not represented in the mind at all — well, that would be a bit disappointing. Instead, by thinking about format, by searching for format *first*, we may have a better chance at discovering the truth.

I will briefly comment on a few other matters relevant to this discussion of 'format'.

## 5.2 Other spatial formats

In this thesis, I have discussed the 'formats' of two spatial properties: location and size. However, these are not the only spatial properties for which we may think about 'format'. The notion that the mind represents information in discernible formats — and may utilize multiple, 'redundant' formats — applies as well to spatial properties like shape orientation.

For example, work in vision science has tried to identify the format of shape representations that support object recognition (e.g., Biederman, 1987; Biederman & Bar, 1999, Kanizsa, 1976; Leyton, 1989). Inspired by work in computer vision (e.g., Liu & Gieger, 1999; Shokoufandeh et al., 2005; see also Blum, 1973), it has been suggested that shapes in the human mind are represented as 'shape skeletons' via the medial axis (which describes the set of all points within an object having two or more closest points along the perimeter of that object; see Psotka, 1978). Given the convergence of behavioral and neuroscientific evidence (see also Ayzenberg & Lourenco, 2019; Lowet et al., 2018), shape skeletons — and the medial axis in particular — provide one of the most robust examples of format in human cognition.

Shape skeletons like medial axes are often compared to other models of shape representation, like principal axes (as in Ayzenberg et al., 2019; Firestone & Scholl, 2014); indeed, medial axes often out-perform other possible formats. However, here we might fruitfully apply this notion that representations need not be reduced to a single format. There is, after all, work supporting the role of principal axes (Marr & Nishihara, 1978; Sturz, Boyer, Magnotti, & Bodily, 2017). As with location representations, it may be that the mind utilizes distinct formats in different ways depending on the task. For instance, it may be that the medial axis is used for object recognition, but that the principal axis is used to evaluate rough size (e.g., whether a certain piece of furniture will fit in a certain space).

These same notion of redundant formats can be applied to the study of other spatial properties, like area (see Corbett & Oriet, 2011; Marchant et al., 2013; Raidvee et al., 2020; Solomon et al., 2011; Yousif, Aslin, & Keil, 2020; Yousif & Keil, 2019; Yousif & Keil, 2021a), volume (Bennette, Keil, & Yousif, 2021; Ekman & Junge, 1961; Teghtsoonian, 1965), and orientation (see Appelle, 1972; Li et al., 2003; Girschick et al., 2011; Henderson & Serences, 2021; Sadalla & Montello, 1989; Yousif, Chen, & Scholl, 2020). We can also think about how format translates across reference frames, or how reference frames themselves are formatted (see, e.g., Farah et al., 1990). In other words, there are lessons to be learned in this thesis about spatial representation *in general* that can be applied to other areas of study.

## 5.3 Space as format

The examples in this thesis so far fail to capture the full extent to which we depend on spatial representations. For example, our representations of number (e.g., Aulet et al., 2021; Dehaene et al., 1993) and time (e.g., Núñez & Cooperrider, 2013), and even social relations (e.g., Parkinson & Wheatley, 2013) may be fundamentally spatial. For that matter, virtually all information represented in working memory may be retained in a spatial way (e.g., van Dijck et al., 2014; van Dijck & Fias, 2011; Yousif, Rosenberg, & Keil, 2021). Thus, we might say that the *format* of numerical representations (or social representations, or representations maintained in working memory) is (are) spatial, at least to some extent. And in each of these instances, a question arises about the format of the underlying spatial representations. Do we represent numbers in a Cartesian space? Do we represent social relationships in a polar-esque cognitive graph? Given that non-spatial knowledge may be represented in a spatial way, understanding the format — or *formats* — of spatial representation may help us to understand not only spatial representation, but *all* representation.

This point was recently articulated by Peer and colleagues (2020) who described how non-spatial information could be represented using either Euclidean or graph-like cognitive maps. They write, "...we may represent the people we know in terms of continuous variables such as various abilities that are naturally encoded as a map-like attribute space... or we may represent them in terms of discrete relationships between individuals that are naturally encoded in graph-like formats (e.g., social networks or family trees)" (p. 48). Indeed, some work has shown that hippocampal cells responsible for spatial representation in rodents do encode other dimensions like time (e.g., MacDonald et al., 2011). Perhaps more interestingly, fMRI has revealed Cartesian-like structures for the representation of numerous stimulus dimensions, including some concrete features like visual size and opacity (Theves et al., 2019), as well as more abstract features like popularity and competence (Park et al., 2020). Nevertheless, there is no clear example of any abstract form of knowledge being represented in a graph-like or polar-esque way.

## 5.4 On Euclid's shoulders

I do not know my earliest memory, but I do know my earliest dream. While taking a nap at my preschool, I had a dream that was nothing but numbers — large digits floating in an empty void. It started with the number "1". Then, like a scroll wheel, the number "1" moved 'up' and the number "2" took its place. Then "3", then "4", and so on. When it got to "9", that's when things got tricky. What happens after "9"? (You'll have to forgive me, I was three-, maybe four-years-old at the time.) In this dream, suddenly, something clicked: When the "9" ticks upwards, we start again: "10" is just a "0" with a "1" in front of it; "11" just a "1" plus something new. I realized, in this empty void of my mind, this most basic fact of our number system. It felt exhilarating. I do remember afterwards trying to explain to my mother what I had just discovered, but she was uninterested. I had, after all, previously known that "10" came after "9". Although I had the capacity to come to understand the base-10 system on my own, I lacked the capacity to explain this realization. I must have sounded crazy.

In middle school and high school, I competed in dozens of math tournaments each year. I loved them. Competitive math tournaments in those early years are mostly geometry, or at least most problems can be solved with geometric solutions one way or another. That is where I thrived. The spatial nature of it all just clicked for me. It felt natural to me. Geometric solutions always felt unassailably true — simply elegant, elegantly simple.

In 9$^{th}$ grade I took my first 'formal' geometry class, and I found it a bit disappointing. In our class, we were meant to spend the *whole year* on a geometry curriculum that students of competitive math tournaments would have mostly figured out on their own by that point. I explained to my teacher that the curriculum should be greatly accelerated, but she was unamused.

Then my second year of college, a guest lecturer came to discuss her work on the 'mental number lines' of non-human primates. She described SNARC effects (e.g., Dehaene et al., 1993) to us, and how mental number lines differed across cultures, and how they were applying these paradigms to gorillas and orangutans (Gazes et al., 2017). And in that ordinary guest lecture, in some random class that I *do not even remember*, I felt the same extraordinary delight that I had when dreaming about numbers. This idea that numbers in my mind were spatial — an idea I had always believed — was *real*. Perhaps more importantly: the idea that the numbers in my mind were spatial was *discernable*. It is quite simple, actually. The basic design of a parity judgment task is cognitive science as its very best: straightforward, easy-to-understand, and powerful. (These days, in addition to numbers, I dream of making a contribution as significant as this one.)

This lightbulb moment jumpstarted my academic career. I reached out to the lecturer expressing interest in these the idea of a mental number line, and I found a lab that studied *exactly this*. In the 2.5 years after that, I conducted multiple projects on mental number lines, and, later, on other aspects of spatial representation. Namely, I studied how children use different geometric cues to navigate, and also 'shape skeletons' as the format of human shape representations. The vast majority of my work since then — the work reflected here in this thesis — has been inspired by simply reflecting on the ways in which my own mind perceives and uses space, then trying to turn those vague intuitions into experiments.

When I think about my work is trying to accomplish, I do not think of any psychologist or biologist or cognitive scientist. I think of Euclid. Euclid is considered to be the founder of geometry not because he discovered things that were unknown before, but because he invented a systematic, logical framework to explain them. Euclid perceived the beauty of the external, spatial world, and distilled that beauty down into discernable axioms.

In this era, obviously much more is known about the axioms supporting the external geometric world. But we are only at the very beginning of understanding the axioms supporting our *internal* geometric world. Cognitive science has succeeded in telling us that spatial information is integral to the mind but has fallen short of providing stable axioms and principles — so far. That is what I hope I have begun to do.

This thesis follows humbly in Euclid's footsteps. Hopefully the work here represents even just one *very small step* in helping us to understand the nature of the geometry of our minds.

## 5.5 Conclusion

The 'format' of many mental representations can be surprisingly accessible: whether studying ants in the Tunisian desert or human visual localization in the lab, 'format' often reveals itself in ordinary behavior. Even tiny, almost-imperceptible errors provide valuable insight into how the mind works.

As I emphasized at the beginning of this thesis, *space* is an ideal domain to deepen our understanding of representation in general. Spatial cognition is *foundational* in the sense that it may serve as a building block for other forms of representation; it is *ubiquitous* in the sense that virtually animals rely on spatial representation to some extent, and *tractable* in the sense that research on spatial representation has made steady progress for the last seven decades (which may seem unremarkable, until you consider how few areas of study can say the same). As much as spatial cognition had been central to the cognitive revolution, it will be central to the ongoing neuroscientific revolution. And understanding format in the ways that I have described in this dissertation will be a critical part of the path forward: If there's any hope of meaningfully bridging the gap between brain and behavior this century, we will surely have to know what it is in the brain we are hoping to find.

If we can use insights gleaned in this domain to guide our study of other domains, we can deepen our understanding of not just spatial representation but *all* mental representation.

# References

Abrahamse, E., Van Dijck, J. P., Majerus, S., & Fias, W. (2014). Finding the answer in space: the mental whiteboard hypothesis on serial order in working memory. *Frontiers in Human Neuroscience*, *8*, 932.

Algom, D., & Fitousi, D. (2016). Half a century of research on Garner interference and the separability–integrality distinction. *Psychological Bulletin*, *142*, 1352-1383.

Allen, R. J., Havelka, J., Falcon, T., Evans, S., & Darling, S. (2015). Modality specificity and integration in working memory: Insights from visuospatial bootstrapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 820-830.

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*, 1698-1716.

Anderson, N. H., & Cuneo, D. O. (1978). The height + width rule in children's judgments of quantity. *Journal of Experimental Psychology: General*, *107*, 335.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychological Science*, *25*, 265-270.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number as a primary perceptual attribute: A review. *Perception*, *45*, 5-31.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique effect" in man and animals. *Psychological Bulletin*, *78*, 266-278.

Attneave, F. (1961) In defense of homunculi. In W. A. Rosenblith (Ed.), *Sensory communication*. MIT Press.

Aulet, L. S., Yousif, S. R., & Lourenco, S. F. (2021). Spatial–numerical associations from a novel paradigm support the mental number line account. *Quarterly Journal of Experimental Psychology*, 17470218211008733.

Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 780.

Ayzenberg, V., & Lourenco, S. F. (2019). Skeletal descriptions of shape provide unique perceptual information for object recognition. *Scientific Reports*, *9*, 1-13.

Ayzenberg, V., Chen, Y., Yousif, S. R., & Lourenco, S. F. (2019). Skeletal representations of shape in human vision: Evidence for a pruned medial axis model. *Journal of Vision*, *19*, 1-21.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*, 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of Learning and Motivation* (Vol. 8, pp. 47-89). Academic press.

Barth, H. C. (2008). Judgments of discrete and continuous quantity: An illusory Stroop effect. *Cognition*, *109*, 251-266.

Barth, H., Kanwisher, N. & Spelke, E. S. (2003). The construction of large number representations in adults. *Cognition*, 86, 201–221.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*, 1622-1631.

Bennette, E., Keil, F. C., & Yousif, S. R. (2021). A Ubiquitous Illusion of Volume: Are Impressions of 3D Volume Captured by an "Additive Heuristic"? *Perception*, *50*, 462-469.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, *94*, 115-147.

Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Research*, *39*, 2885-2899.

Blum, H. (1973). Biological shape and visual science (Part I). *Journal of theoretical Biology*, *38*, 205-287.

Brannon, E. M. & Terrace, H. S. (1998). Ordering of the numerosities 1–9 by monkeys. *Science*, 282, 746–749.

Brannon, E. M., Abbott, S., & Lutz, D. J. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition*, *93*, 59-68.

Brannon, E. M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Developmental Science*, *9*, 59-64.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, *47*, 139-159.

Calia, C., Darling, S., Havelka, J., & Allen, R. J. (2019). Visuospatial bootstrapping: Binding useful visuospatial information during verbal working memory encoding does not require set-shifting executive resources. *Quarterly Journal of Experimental Psychology*, *72*, 913-921.

Carbon, C. C. (2016). The folded paper size illusion: Evidence of inability to perceptually integrate more than one geometrical dimension. *i-Perception*, *7*, 1-5.

Chang, L., & Tsao, D. Y. (2017). The code for facial identity in the primate brain. *Cell*, *169*, 1013-1028.

Corbett, J. E., & Oriet, C. (2011). The whole is indeed more than the sum of its parts: Perceptual averaging in the absence of individual item representation. *Acta Psychologica*, *138*, 289-301.

Coren, S., & Girgus, J. S. (1978). Seeing is deceiving: The psychology of visual illusions. Hillsdale, NJ: Erlbaum.

Cowan, N. (1998). Visual and auditory working memory. *Trends in Cognitive Sciences, 2*, 77–78.

Cronin, D., Peacock, C., & Henderson, J. M. (2019). Visual and verbal working memory loads interfere with scene viewing.

Darling, S., & Havelka, J. (2010). Visuospatial bootstrapping: Evidence for binding of verbal and spatial information in working memory. *Quarterly Journal of Experimental Psychology*, *63*, 239-245.

Darling, S., Allen, R. J., & Havelka, J. (2017). Visuospatial bootstrapping: When visuospatial and verbal memory work together. *Current Directions in Psychological Science*, *26*, 3-9.

Darling, S., Allen, R. J., Havelka, J., Campbell, A., & Rattray, E. (2012). Visuospatial bootstrapping: Long-term memory representations are necessary for implicit binding of verbal and visuospatial working memory. *Psychonomic Bulletin & Review*, *19*, 258-263.

Darling, S., Havelka, J., Allen, R. J., Bunyan, E. & Flornes, L. (2020) Visuospatial bootstrapping: Spatialized displays enhance digit and nonword sequence learning. Annals of the New York Academy of Sciences, *1477*, 100-112.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*, 371-396.

Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 149.

Ekman, G., & Junge, K. (1961). Psychophysical relations in visual perception of length, area and volume. *Scandinavian Journal of Psychology*, *2*, 1-10.

Ericson, J. D., & Warren, W. H. (2020). Probing the invariant structure of spatial knowledge: Support for the cognitive graph hypothesis. *Cognition*, *200*, 104276.

Farah, M. J., Brunn, J. L., Wong, A. B., Wallace, M. A., & Carpenter, P. A. (1990). Frames of reference for allocating attention to space: Evidence from the neglect syndrome. *Neuropsychologia*, *28*, 335-347.

Firestone, C., & Scholl, B. J. (2014). "Please tap the shape, anywhere you like" shape skeletons in human vision revealed by an exceedingly simple measure. *Psychological Science*, *25*, 377-386.

Foster, J. J., Bsales, E. M., Jaffe, R. J., & Awh, E. (2017). Alpha-band activity reveals spontaneous representations of spatial position in visual working memory. *Current Biology*, *27*, 3216-3223.

Fougnie, D., Suchow, J., & Alvarez, G. (2012). Variability in the quality of visual working memory. *Nature Communications, 3*, 1229.

Gallistel, C. R. (1990). The organization of learning. Cambridge, MA: MIT Press.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, *1*, 225-241.

Gazes, R. P., Diamond, R. F., Hope, J. M., Caillaud, D., Stoinski, T. S., & Hampton, R. R. (2017). Spatial representation of magnitude in gorillas and orangutans. *Cognition*, *168*, 312-319.

Gigerenzer, G., & Richter, H. R. (1990). Context effects and their interaction with development: Area judgments. *Cognitive Development*, 5, 235-264.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, *14*, 926-932.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, 306, 496–499.

Guida, A., & Campitelli, G. (2019). Explaining the SPoARC and SNARC effects with knowledge structures: An expertise account. *Psychonomic Bulletin & Review*, *26*(2), 434-451.

Guida, A., Leroux, A., Lavielle-Guida, M., & Noël, Y. (2016). A SPoARC in the dark: Spatialization in verbal immediate memory. *Cognitive Science*, *40*, 2108-2121.

Guida, A., Megreya, A. M., Lavielle-Guida, M., Noël, Y., Mathy, F., van Dijck, J. P., & Abrahamse, E. (2018). Spatialization in working memory is related to literacy and reading direction: Culture "literarily" directs our thoughts. *Cognition*, *175*, 96-100.

Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665-668.

Henderson, M. M., & Serences, J. (2021). Biased orientation representations can be explained by experience with non-uniform training set statistics. *bioRxiv*, 2020-07.

Hubbard, T. L. (Ed.). (2018). *Spatial Biases in Perception and Cognition.* Cambridge, England: Cambridge University Press.

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences*, *103*, 19599-19604.

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

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.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175-219.

Kanizsa, G. (1976). Subjective contours. *Scientific American*, *234*, 48-53.

Klauer, K. C., & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology: General*, *133*, 355-381.

Kool, W., Conway, A. R., & Turk-Browne, N. B. (2014). Sequential dynamics in visual short-term memory. *Attention, Perception, & Psychophysics*, *76*, 1885-1901.

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.

Krider, R. E., Raghubir, P., & Krishna, A. (2001). Pizzas: π or square? Psychophysical biases in area comparisons. *Marketing Science*, *20*, 405-425.

Kuipers, B., Tecuci, D. G., & Stankiewicz, B. J. (2003). The skeleton in the cognitive map: A computational and empirical exploration. Environment and Behavior, 35, 81–106.

Langlois, T. A., Jacoby, N., Suchow, J. W., & Griffiths, T. L. (2021). Serial reproduction reveals the geometry of visuospatial representations. *Proceedings of the National Academy of Sciences*, *118*, e2012938118.

Leyton, M. (1989). Inferring causal history froms shape. *Cognitive Science*, *13*, 357-387.

Li, B., Peterson, M. R., & Freeman, R. D. (2003). Oblique effect: a neural basis in the visual cortex. *Journal of Neurophysiology*, *90*, 204-217.

Lidji, P., Kolinsky, R., Lochy, A., & Morais, J. (2007). Spatial associations for musical stimuli: A piano in the head? Journal of Experimental Psychology: Human Perception and Performance, *33*, 1189-1207.

Liu, T. L., & Geiger, D. (1999, September). Approximate tree matching and shape similarity. In *Proceedings of the Seventh IEEE International Conference on Computer Vision* (Vol. 1, pp. 456-462). IEEE.

Lourenco, S. F., & Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Developmental Science*, *20*, 12418-12420.

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science*, *21*, 873-881.

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences*, *109*, 18737-18742.

Lowet, A. S., Firestone, C., & Scholl, B. J. (2018). Seeing structure: Shape skeletons modulate perceived similarity. *Attention, Perception, & Psychophysics*, *80*, 1278-1289.

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, *71*, 737-749.

Maley, C. J. (2021, October 5). The physicality of representation. https://doi.org/10.31234/osf.io/ame6u

Marchant, A. P., Simons, D. J., & de Fockert, J. W. (2013). Ensemble representations: Effects of set size and item heterogeneity on average size perception. *Acta Psychologica*, *142*, 245-250.

Marchant, A. P., Simons, D. J., & de Fockert, J. W. (2013). Ensemble representations: Effects of set size and item heterogeneity on average size perception. *Acta Psychologica*, *142*, 245–250.

Markman, A. B., & Dietrich, E. (2000). In defense of representation. *Cognitive Psychology*, *40*, 138-171.

Marr, D. (1982). *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman, San Francisco.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *200*, 269-294.

McCloskey, M., & Palmer, E. (1996). Visual representation of object location: Insights from localization impairments. *Current Directions in Psychological Science*, *5*, 25-28.

McCloskey, M., Rapp, B., Yantis, S., Rubin, G., Bacon, W. F., Dagnelie, G., ... & Johnson, D. N. (1995). A developmental deficit in localizing objects from vision. *Psychological Science*, *6*, 112-117.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them?. *Psychological Bulletin*, *128*, 278-294.

Morey, C. C. (2018). The case against specialized visual-spatial short-term memory. *Psychological Bulletin*, *144*, 849-883.

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

Moser, E. I., Roudi, Y., Witter, M. P., Kentros, C., Bonhoeffer, T., & Moser, M. B. (2014). Grid cells and cortical representation. Nature Reviews Neuroscience, 15, 466–481.

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

Nachmias, J. (2008). Judging spatial properties of simple figures. *Vision Research*, 48, 1290-1296.

Nachmias, J. (2011). Shape and size discrimination compared. *Vision Research*, 51, 400-407.

Nieder, A. & Miller, E. (2004). A parieto–frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences*, *101*, 7457–7462.

Núñez, R., & Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends in Cognitive Sciences*, *17*, 220-229.

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

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? Psychological Bulletin, *142*, 758-799.

Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology*, *49*, 1103-1112.

Olson, D. R. (2013). *Cognitive development: The child's acquisition of diagonality.* Psychology Press.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology*, *64*, 77-107.

Park, S. A., Miller, D. S., Nili, H., Ranganath, C., & Boorman, E. D. (2020). Map making: constructing, combining, and inferring on abstract cognitive maps. *Neuron*, *107*, 1226-1238.

Parkinson, C., & Wheatley, T. (2013). Old cortex, new contexts: Re-purposing spatial perception for social cognition. *Frontiers in Human Neuroscience*, *7:645, 1-7.*

Peer, M., Brunec, I. K., Newcombe, N. S., & Epstein, R. A. (2020). Structuring Knowledge with Cognitive Maps and Cognitive Graphs. *Trends in Cognitive Sciences*. *25*. 37-54.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8-13.

Pertzov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception, & Psychophysics*, *76*, 1914-1924.

Pica, P., Lemer, C., Izard, V. & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*, 499–503.

Psotka, J. (1978). Perceptual processes that may create stick figures and balance. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 101–111.

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

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*, 157-238.

Race, E., Palombo, D. J., Cadden, M., Burke, K., & Verfaellie, M. (2015). Memory integration in amnesia: Prior knowledge supports verbal short-term memory. *Neuropsychologia*, *70*, 272-280.

Raidvee, A., Toom, M., Averin, K., & Allik, J. (2020). Perception of means, sums, and areas. *Attention, Perception, & Psychophysics*, *82*, 865-876.

Rajsic, J., & Wilson, D. E. (2014). Asymmetrical access to color and location in visual working memory. *Attention, Perception, & Psychophysics*, *76*, 1902-1913.

Rousselle, L., Palmers, E., & Noël, M. P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology*, *87*, 57-84.

Sadalla, E. K., & Montello, D. R. (1989). Remembering changes in direction. *Environment and Behavior*, *21*, 346-363.

Shea, N. (2018). *Representation in cognitive science* (p. 304). Oxford University Press.

Shokoufandeh, A., Macrini, D., Dickinson, S., Siddiqi, K., & Zucker, S. W. (2005). Indexing hierarchical structures using graph spectra. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*, 1125-1140.

Sokolowski, H. M., Fias, W., Ononye, C. B., & Ansari, D. (2017). Are numbers grounded in a general magnitude processing system? A functional neuroimaging meta-analysis. *Neuropsychologia*, *105*, 50-69.

Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *Journal of Vision*, *11*, 1-11.

Sturz, B. R., Boyer, T. W., Magnotti, J. F., & Bodily, K. D. (2017). Do Eye Movements During Shape Discrimination Reveal an Underlying Geometric Structure? *Animal Behavior and Cognition*, *4*, 267-285.

Teghtsoonian, M. (1965). The judgment of size. *The American Journal of Psychology*, *78*, 392-402.

Theeuwes, J., Belopolsky, A., & Olivers, C. N. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica*, *132*, 106-114.

Thelen, E., & Smith, L. B. (1996). *A dynamic systems approach to the development of cognition and action*. MIT press.

Theves, S., Fernandez, G., & Doeller, C. F. (2019). The hippocampus encodes distances in multidimensional feature space. *Current Biology*, *29*, 1226-1231.

Traxler, M., & Gernsbacher, M. A. (Eds.). (2011). *Handbook of psycholinguistics*. Elsevier.

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition*, *34*, 1704-1719.

van Dijck, J. P., & Fias, W. (2011). A working memory account for spatial–numerical associations. *Cognition*, *119*, 114-119.

van Dijck, J.-P., Abrahamse, E. L., Acar, F., Ketels, B., & Fias, W. (2014). A Working Memory Account of the Interaction between Numbers and Spatial Attention. *Quarterly Journal of Experimental Psychology*, *67*, 1500-1513.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483-488.

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.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t-tests. *Perspectives on Psychological Science*, *6*, 291-298.

Woodman, G. F., Vogel, E. K., & Luck, S. J. (2012). Flexibility in visual working memory: Accurate change detection in the face of irrelevant variations in position. *Visual Cognition*, *20*, 1-28.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, 1-11.

Yang, F., & Flombaum, J. (2018). Polar coordinates as the format of spatial representation in visual perception [Abstract]. *Journal of Vision*, *18*, 21.

Yousif, S. R. (2022, January 10). Redundancy and reducibility in the formats of spatial representations. https://doi.org/10.31234/osf.io/g6f8v

Yousif, S. R., Alexandrov, E., Bennette, E., Aslin, R. N., & Keil, F. C. (2022). Do children estimate area using an 'Additive-Area Heuristic'? *Developmental Science*.

Yousif, S. R., Aslin, R. N., & Keil, F. C. (2020). Judgments of spatial extent are fundamentally illusory: 'Additive-area' provides the best explanation. *Cognition*. *205*. 104439.

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. C. (2019). The additive-area heuristic: An efficient but illusory means of visual area approximation. *Psychological Science*, *30*, 495-503.

Yousif, S. R., & Keil, F. C. (2021a). How we see area and why it matters. *Trends in Cognitive Sciences*. *25*, 554-557.

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., & Keil, F. (2021c). Quantity perception: The forest and the trees. https://doi.org/10.31234/osf.io/ydcg7

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.

Yousif, S. R., Rosenberg, M. D., & Keil, F. C. (2021). Using space to remember: Short-term spatial structure spontaneously improves working memory. *Cognition*, *214*, 104748.

Zorzi, M., Priftis, K., & Umiltà, C. (2002). Brain damage: Neglect disrupts the mental number line. *Nature*, *417*, 138-139.