UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Embodying the Future: Modeling Visually Guided Planning as Prospective Mental Simulation

Permalink
https://escholarship.org/uc/item/1wv481cf

Author
Gordon, Jeremy Raboff

Publication Date
2023

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Embodying the Future: Modeling Visually Guided Planning as Prospective Mental Simulation

By

Jeremy R. Gordon

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Information Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John Chuang, Chair
Professor Coye Cheshire
Professor Steven Piantadosi
Dr. Giovanni Pezzulo

Summer 2023

Abstract

Embodying the Future: Modeling Visually Guided Planning as Prospective Mental Simulation

by

Jeremy R. Gordon

Doctor of Philosophy in Information Science

University of California, Berkeley

Professor John Chuang, Chair

What would it feel like to run outside, right now, and attempt a somersault on the first surface you find? Taking seriously an invitation like this to imagine a (perhaps unlikely) future, prompts the activation of evolutionary machinery in the mind and body that took millions of years to emerge. The ability to answer this question depends upon a surprisingly complex model of yourself, the ecology you inhabit, and how you interact within it. Throughout the first several seconds of engaging with this prompt, you have likely tapped memories of a wide spatial vicinity around you, and the types of surfaces likely to be encountered. The season, today's weather, the existence of a lawn and the schedule of a sprinkler system, all brought to bear to produce simulations of the kinds of experience that might result, and the sorts of choices that might be required. Enactive and embodied views of cognition tell us that these simulated experiences are not unitary but parallel and overlapping (A. Clark, 2015), are unbound to typical temporal sequences and durations (Arnold, Iaria, & Ekstrom, 2016), and live not in your head as local neural traces, but leverage your entire body (Thompson, 2010) as you consider the sensory and motor consequences of an activity you may not have performed in quite a while, and certainly never in this particular context. These simulations are constructed under the guidance of past experiences (how fast do you typically run?), counterfactual contingencies (what if the lawn is too crowded?), and knowledge of complex causal relations and social dynamics (how would a passerby respond? what would a colleague think?).

Embodied, prospective simulations of this sort are not a rare occurrence prompted primarily by thought experiments in a thesis. They are ubiquitous in our daily experience. Our cognitive ability to predict how events may unfold in time, how the world will respond to our actions, and how our body will respond to the world, is likely one of the central, and perhaps unique (Bulley, Henry, & Suddendorf, 2016), competencies that allow members of our species to solve problems we've never before encountered, and construct, share, and run experiments on candidate actions and plans before committing to enact them.

This work focuses predominantly on the domain of spatial navigation. As a modeling target, navigation enjoys a long history in computer science, as well as the planning literature, and appears as one of four categories included in Buckner and Carrol's ontology of self-

projection (Buckner & Carroll, 2007). Given its natural low-dimensional spatial state space, navigation is especially well-suited to behavioral study.

Psychologists have gathered a long list of effects outlining key features of the way we learn about and model the world. Tendencies to infer causal relations (Kushnir & Gopnik, 2005), to perceive patterns (even when none exists, e.g., Fyfe, Williams, Mason, and Pickup (2008)), to develop an intuitive understanding of physical systems (Krist, Fieberg, & Wilkening, 1993), to categorize, and hierarchically arrange novel stimuli and task environments (Collins & Frank, 2013), are all suggestive of the processes by which such a model are learned and employed. However, far less is known about the dynamics by which we navigate this internal model, its constraints and topology, or the mechanisms by which simulations traverse large spatial and temporal distances to most effectively instruct the action-oriented needs of the present. In short, we know little about the kinematics of mental simulation. I aim to pursue, then, the following claim: by studying behavior and sensory exploration during navigation, we may begin to characterize the human solution to planning in an uncertain world, and take steps towards a kinematics of prospective mental simulation—the dynamic process by which individuals embody and interact with multiple possible futures.

This work is guided by research questions that I categorize into three areas of inquiry: 1) embodied planning, 2) visual salience and simulation kinematics in navigation planning, and 3) prospection in joint planning.

Embodied planning Questions in this area aim to explore the limits of classical approaches, and suggest avenues and opportunities to develop a more nuanced and naturalistic theory of planning in the real world. I will argue in Chapter 3 that this need to accommodate embodied dynamics—in decision-making as well as less explicitly choice-centered natural activities—requires us, as a scientific community, to develop new models better able to capture the richness of the mental processes that anticipate and drive action in the world. The reviewed work displays a range of approaches to bring embodied intelligence into communication with the classical literature on decision-making and planning. In particular, I'll develop the concept of the affordance landscape, with example applications in various naturalistic (and fundamentally embodied) tasks such as playing team sports, climbing, and crossing a river. These theoretical arguments and methodologies set a foundation for the next research area, in which a mode of embodiment particularly relevant to visual creatures like ourselves is explored.

Visual salience and simulation kinematics in navigation planning This area captures the primary empirical questions of this dissertation, and motivates the use of an agent model designed to iteratively learn the sort of affordance landscapes discussed in Chapter 3. I'll tackle questions about attractors of visual attention, and temporal sampling dynamics, through a variety of behavioral analyses. Key findings in Chapters 4 and 5 include: the predictive relationship between planning-time biometrics and navigation-time decisions, geometric and map-geometry attributes most predictive of attention, as well as various hierarchical aspects of the sequential gaze patterns seen during planning. I also report behavioral differences between top and bottom performers in this task, which raise plausible explanations for how, and through what mechanisms, individuals differ in their spatial planning abilities.

Prospection in joint planning In the third area of inquiry I'll look beyond the individual spatial navigation paradigm to explore the case of multiple collaborators who must, through a variety of techniques, converge on a shared strategy to jointly solve the task at hand. I propose that some of these techniques implicate a planning process driven by the simulation of (shared) counterfactual future trajectories. However, I also acknowledge the significant additional complexity that comes from introducing social interaction, even in relatively simple problem-solving paradigms, and even when verbal communication is limited. I report findings from a two-person collaborative task which suggests a history-dependence of strategy selection when task dynamics changed. Fairness (defined as the balance of effort) was found to be lower in imbalanced task environments, as well as during the more difficult blocks requiring greater partner monitoring. Finally, I find a persistent bias towards spatial separation (which likely helps avoid interaction and enhances decoupling between collaborators), even among dyads using strict color-based strategies, highlighting that non-trivial hybrid approaches which mix conventions can be productively adopted even in non-verbal collaborative settings.

This investigation of embodied prospection builds on a diverse multidisciplinary literature spanning work on mental simulation and episodic foresight in psychology, predictive processing in philosophy of mind, hippocampal preplay and internally generated sequences in neuroscience, embodied and enactive cognitive science, as well as a number of biologically inspired computational models of planning.

Contributions Building on this multidisciplinary foundation, I hope to make three contributions:

First is a conceptual extension made by projecting the body outwards into its likely future, and seeing this projection as a first class representation of the self within which simulations are continuously run. My claim is that this projection is the interface within which future-determining actions are enacted, and is therefore, in an important sense, cognitively inseparable from the individual. Furthermore, the sensorimotor patterns generated in these simulations are not fundamentally different from the inferential (and simulative) process of perception in the present, and therefore are equally constitutive of our moment-to-moment experience.

Second, following Guest and Martin (2021), I propose an instantiation of this theoretical framing within a computational agent given control of a sensory apparatus similar to ours, and tasked with the same types of navigational challenges posed to the human participants in my studies. I argue that models like this one can help to achieve a unified conceptual framing of future-oriented dynamics that links both low-level predictive processing (which may operate over short temporal durations), with the kinds of dramatically extended episodic simulations typical when studying mental time travel. While a diversity of mechanisms are entailed in such a unified framework, I believe seeing these as two ends of a continuum with shared function, and more than likely, algorithmic and process-level correspondence as well.

Third, I share a series of empirical findings relating to visually guided planning and execution within spatial navigation problems. Some results are consistent with the more speculative ideas raised earlier in this section, and others highlight the remaining gaps in our understanding of the tremendously complex processes underlying visual search, prospection, and planning.

By pursuing the larger project within which this work sits, we can work towards a more nuanced understanding of the dynamics by which individuals generate and test expectations for the future. If successful, this project should help deepen our understanding of the way these prospective cognitive processes are embedded in a body, and a complex motor system, that grounds them in the world around us. If so, we may better appreciate the aspects of human intelligence that are truly unique, and defined by a complexity that is still far out of reach of replication in artificial systems. As such, perhaps our computational abilities will be better leveraged towards the design of tools personalized to individual preferences and idiosyncrasies, or even self-acknowledged oversights. Such systems may be able to externalize more legible beliefs, risks, and hypothetical futures, and by availing these to us, help enable more effective communication and collaboration. Finally, by better understanding how individuals project themselves into uncertain futures, and reason about long-term consequences, we might prepare ourselves, as a society, to more effectively face the challenges still ahead.

Acknowledgements

I am so grateful for the last 72 or so months. I have had the opportunity to sit with this time, situated within an incredible institution, deep expertise, and the evolving intellectual interests of other students and mentors. To be able to choose a topic based on nothing other than curiosity, to collaboratively and independently explore multiple bodies of literature, conjure a personal understanding, and then share the result with others who are fascinated by similar ideas — what an immense privilege. I want to acknowledge some of the people, and institutions, who made this time possible and valuable.

I'm grateful for support from the Center for Long-Term Cybersecurity, and NSF's GRFP program[1], which enabled a level of focus on my own research I didn't expect would be possible.

Thank you to Jeff, Subutai, and the team at Numenta, for producing the writings and research that were the most proximal trigger for my rediscovery of interests in neuroscience, the plausibility of modeling the mind, and the decision to apply to graduate programs in the first place. To spend time in Redwood City, around the people whose work has so directly inspired and influenced my own ideas, was a joy.

Thank you to Mar, Alex, Amos, Ed, and Melanie, and others on the Microsoft Research Soundscape team, for giving me reason and structure to ground some of the theoretical work I've been interested in towards the challenging task of designing an experience for real users.

Thank you to Max, Noura, Nick, Richmond, and other former members of the BioSense group, who introduced me to empirical science with a critical lens, and demonstrated what a mutually supportive and curious academic lab can be.

Thank you to Mayank, who continues to instruct me in the purest form of scientific curiosity, and never shies from epistemological marshlands.

Thank you to the ISchool faculty and staff for the unique, open-minded, and deep and authentic commitment to multidisciplinarity that I've come to realize is extraordinarily rare in the academic world. I could, perhaps, have done this work at a psychology department, or neuroscience or computer science, but it would have been narrower, probably, and certainly less fun.

Thank you to Dor and the completely unique Embodied Underground, who immersed me for a semester in a non-stop series of demonstrations of different ways of knowing and communicating about cognition and the mind. Though the language and concepts were unfamiliar and often confusing, they prodded me to question some of the assumptions and intuitions I had come to over-depend on.

Thank you to Coye and Steven for wading through the muddy water of too-rough drafts, and pointing out my oversights with compassion. Thank you to Giovanni for responding to an email from a stranger in 2018, and reflecting back my questions with enthusiasm and suggestions. I never expected that email to become a lengthy thread, let alone a multi-year collaboration (that continues to date). Thank you for sharing your time and energy to advise and guide, and validate and redirect, when necessary — you've become an incredible mentor. Without your support, this dissertation would look very different.

[1]National Science Foundation Graduate Research Fellowship Program Grant No. 2019236659

Thank you to my additional hosts at the National Research Council in Rome, Antonella, Gian Luca, Mattia, who welcomed me despite an unprecedented global context, and educated me about motor control, bounded rationality, Italian politics, and the origin stories of several novel shapes of pasta.

Thank you to John, for sharing your unbounded intellectual curiosity, and nurturing mine with constant interest and patience. I feel so lucky to have had cause to speak with you every other week for the last 6 years. I promise we can reduce our checkin pace now — how about monthly?

Thank you mom, dad, and Allegra. You taught me that questions, introspection, creativity, and earnest pursuit of the unknown are among the best things we can do with our time; what did you expect?

And thanks, finally, to Sandra, who said, more than once, "you should pursue your interests!" and continued to mean it, throughout. This dissertation exists only because I took that (perhaps too much) to heart, and I am so grateful to you for being supportive through all of its twists and turns.

Thank you, thank you, thank you!

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 A present for the future

Both the writing, and the reading, of these words appears to happen within a moment we often refer to as "now." However, as is common in scientific inquiry, the closer we look at an apparently discrete boundary or concept, the harder it is to see its form, and the less confident we become in its use. The blur around our conception of the present is not just a semantic game or a philosophical trick, it is a necessary response to the demands confronted by a constantly changing organism within an equally fluid world. Light incident on this page (or generated by this screen) takes less than two billionths of a second to close the gap to your eye, but the information it contains hits a dramatic bottleneck as it passes into you; the journey from photo-receptors and then ganglion cells in the retina to lateral geniculate nucleus in the thalamus, and onto V1, V2, and higher regions of visual cortex, takes up to a fifth of a second (Humphries, 2021). A grouping of nearby contrasting edges, like, say,

this one,

finds cells particularly excited by its pattern, which, despite the fact you are likely not listening to someone speak, includes populations in speech-selective areas of audio cortex (Perrone-Bertolotti et al., 2012; Skipper, Nusbaum, & Small, 2005). However, by the time such higher cortical activity begins, a burst of spikes in the frontal eye fields has already signaled oculomotor muscles to fire off a ballistic movement—a saccade—that drastically transforms the pattern of light falling on the retina. That is, "present" neural activity in higher levels of visual cortex (e.g. inferior temporal cortex), firing 20-150 times a second, corresponds to moments that have already passed "out there" in the world. Some of these patterns may persist well after landing on a new pattern a little further to the right (or perhaps upwards above the page to investigate a momentary distraction like a bird passing by the window), and will continue as a new packet of neural spikes begin their journey. We know some of this neural activity is in and around the hippocampus, firing patterns that may have been first experienced years ago (perhaps during language learning, or more recently developed associations). Theories of predictive processing suggest that, even before a saccade lands, downward flowing feedback pathways generate anticipatory activations consistent with a variety of likely subsequent patterns which help to overcome unavoidable

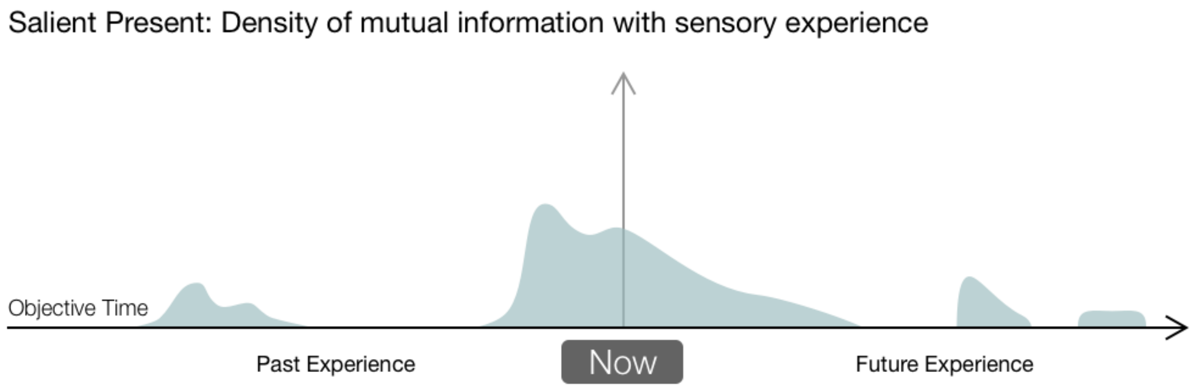


Figure 1.1: Schematic conceptual time-distribution of the salient present. Neural activity at any given moment reflects (contains mutual information with) not only sensory patterns presently available (retinal images, touches, smells, proprioceptive signals), but those recently encountered, those anticipated, and even those from distal events in the past and forecasted to be experienced in the future. The two peaks in the central area of the distribution schematically indicate, respectively: the delay in sensory processing such that present percepts are most correlated with the near past, and prediction signals that attempt to generate anticipated input synchronized to the "true" present (now-casting), as one strategy to overcome this delay.

information processing delays, and guide the control of sensors to disambiguate between competing sensory causes (A. Clark, 2015). These too are *in our experience* of the present, despite not having yet *been experienced*. We know that in a conversation between the deeper hippocampal neurons and higher areas of prefrontal cortex like vmPFC (perhaps in the same region as the anticipatory populations just discussed), neural activations can reflect not just expectations of future words on the page or probable syntactic structures, but entire episodes built from prior experience, but which have not yet occurred—a planned walk before dinner, or a pending email to a colleague.

None of these activities happen in a mathematically precise conception of "now." Or, it might be more helpful to say instead that *all* of this happens now, and that our conception of the present is a continuous amalgamation of neural activation (Buonomano, 2017) (both in the central nervous system and muscle fibers at the peripheries) that is reflective of the *salient present* (see Figure 1.1). This "bump" of activity is particularly aligned with temporally nearby sensory events, yet also likely exhibits correlations with activity spanning years or even decades[1]. In the chapters that follow, I hope to convey and explore, through a combination of empirical and computational methods, a notion of this salient present that has a shape, a dynamic character, and a tight coupling to the sensors and muscles within which it is experienced. It also should be seen as having a future-orientation that is well adapted to preparation for action in an uncertain world. This version of now is molded moment to moment by individuals, but extends out beyond the traditional boundaries of the body, into the world of tools, technologies, and communities in which we interact.

[1] I note, however, that the ubiquity of nonstationarity has been reported across many types of neural signals including electroencephalography (EEG), evoked potential (EP), local field potential (LFP), and others, which poses significant challenges to the study of neural representations over even short time periods (Tong, Li, Zhu, & Thakor, 2007).

1.2 The case for embodied prospection

In the following section, I will argue that adaptive pressures favoring future orientation have been central to our evolutionary history, and as such, we should see prospective abilities as a driving principle underlying cognition. The enormous dimensionality of our modern niche poses challenges for adaptive planning and action, and typical strategies that constrain the space of possible futures via learned experience do not offer a complete solution. I will then propose an alternative solution: that by seeing the internal generative model (which drives anticipated experience) as itself richly embodied, prospection can be rendered tractable under the constraints and enhanced simulation dynamics afforded by embodied intelligence. Next, I will introduce my primary research questions, around the characterization of the kinematics of this embodied generative model. Finally, I will set up the bulk of the work in this dissertation by claiming that the study of behavior and physiology in the dynamic and naturalistic paradigm of prospective spatial navigation is fertile terrain for investigations into these kinematics.

1.2.1 Origins of future-oriented cognition

Humans may be alone in the extent to which we can imagine scenarios we've never experienced and evaluate potential futures before they come to be. This capacity for prospection plays an integral role in situations that both explicitly and implicitly require consideration of the results of our actions, such as planning, and various classes of problem-solving. However, formulating accurate expectations for how the world, and our engagement with it, are likely to unfold, may play a central role in a much wider array of cognitive processes.

To understand why, it is useful to look back at the evolutionary history that shaped us, as well as the other primates, mammals, and other taxonomic branches that share some of these abilities (Cisek, 2019; Raby & Clayton, 2009; Nilsson, 2009). While single-celled organisms are capable of successful adaptive locomotion by using apparently simple and reflexive responses to the immediate sensory environment (e.g. the ascent of temperature, light, or chemical gradients in chemotaxis and klinokinesis) larger organisms confront dynamic niches requiring complex hierarchical motor behaviors that must be prepared and temporally coordinated with a changing environment.

The evolutionary pressure to anticipate likely arose well before advanced motor systems. The physics of even simple perceptive systems, for example the toxin avoidance neural circuit in *C. elegans*, which inhibits forward movement upon detection (Faumont, Lindsay, & Lockery, 2012), has an inherently retrospective nature. All physical processes take time, so at the moment information about the presence of a toxin is available, it may already be too late for an organism to instigate an effective response. As such, even early perception might be seen as motivated by now-casting: the imperative to *infer* the state of the world in the true present, using only evidence from the past. Any evolutionary innovations to perception improving this inferential process, small steps towards synchrony with the world, may therefore have conferred adaptive benefits.

Consider for example the development of whiskers among our nocturnal mammalian ancestors. Whisking, especially when other sensory modalities are unavailable, provides a preview of the local structure of the world; during movement, patterns of activity in somatosensory cortex, produced by active whisking, outline the structure of the local vicinity—candidates for future occupancy. The existence of an obstacle ahead can be brought into awareness (we might say *simulated*) without committing to potentially detrimental effects to the body. Such a view, of course, can be extended to all exteroceptive sensory modalities. For humans, whose visual system is a dominant cause of neocortical activity, our eyes provide similar, though far more spatially extended, capacities for preview (Nilsson, 2009).

While sensory exploration can provide crucial information about the present state of the world, and impart valuable clues to support anticipation of our passage through it, the present perceptual state is not always a good predictor of the next. Though the question of whether perception is discrete or continuous is unresolved (White, 2018), pure sensory inputs are imprisoned to the just-past moment (Buonomano, 2017). To prepare effectively for a cascade of possible future states which may diverge from present sensory evidence (e.g. when extending a hand to touch an unknown object, or peering over a sheer cliff), we turn to another innovation: a mechanism allowing us to learn transitions (temporally ordered associations), which enable the *generation* of expected percepts before confirming data has arrived (A. Clark, 2013). A strong candidate mechanism, known to neuroscientists as prospective coding, is a neural learning rule driven by spike-timing-dependent plasticity (STDP) (Brea, Gaál, Urbanczik, & Senn, 2016). By this Hebbian mechanism, synapses from neuron A to B are strengthened when B fires after A, but within a limited time window. In this way, the activation of A can increase the probability that B will activate, even prior to other inputs to B firing.

There are other candidates for neural implementations supporting prediction and anticipatory activation, but the functional effect, when deployed across larger temporal scales, has been described and theorized about, under a variety of names, for centuries. Notions of *mental simulation* are discussed by Aristotle ("there will always be in the mind of a man who remembers or expects something an image or picture of what he remembers or expects" (Rhetoric 1.1)), through Jamesian phenomenology and gestalt psychology, and most recently have been aided by empirical evidence from cognitive neuroscience, electrophysiology, psychophysics, and vision science. *Mental imagery* allows humans to both reconstruct memories of prior sensory patterns (or episodic sequences of sensory patterns) as well as to stochastically construct novel patterns through recombination, extrapolation, and counterfactual substitution (Kahneman & Tversky, 1981a).

1.2.2 Infinitely forking paths

In an overwhelmingly high dimensional world, however, the complexity of causal dynamics results in an unbounded number of ways the future might unfold. A driver must simulate not only the approximate physics of her own vehicle, but the simultaneous trajectories of multiple actors whose perceptions, choices, and actions interact as her own future manifests. In such a world, which futures should we generate?

The most common answer to this question is experience. A theory increasingly popular in philosophy of mind and computational neuroscience posits that crucial to a number of cognitive processes is an internal *spatiotemporal generative model (SGM)* putatively implemented by the hippocampus and entorhinal system (HE-S) coupled with higher regions in prefrontal cortex, which learns sequences during experience, and arranges them into a hierarchical structure (Çatal, Verbelen, Van de Maele, Dhoedt, & Safron, 2021). Once learned, the SGM furnishes an organism with the ability to produce (a range of) plausible sensorimotor sequences (what Clark calls "fantasies that coincide with reality" (A. Clark, 2015)) conditioned on an initial context or state.

To resolve the dimensionality explosion problem, however, a generative model learned from experience is not sufficient. To see this, we turn to a common computational paradigm for modeling sequential agent-based action in realistic environments. The partially observed Markov decision process (POMDP) is a highly influential formalism used by active inference and reinforcement learning, as well as earlier studies of planning in cognitive science and computer science. In POMDP, we assume the environment and agent evolve according to a Markov decision process (MDP), in which all transitions between a discrete set of states (e.g. an agent in a particular configuration at a particular location), are describable by a probabilistic graphical model. However, in the partially observable paradigm, the agent cannot observe states directly, and must instead infer them from limited observations of the world, coupled with a sensor model enabling state inference. The computational complexity of a problem modeled as a POMDP is a function of both the problem domain, as well as various decisions of the modeler, e.g. the chosen state representation, dynamics, and functional form of the sensor model. However, in even small state spaces, computation of an exact optimal policy is often intractable (Hauskrecht, 2000). Researchers in artificial intelligence and reinforcement learning have developed many solutions to approximate POMDP solutions, often by constructing and maintaining a limited tree of considered states and state-action transitions (Chaslot, Bakkes, Szita, & Spronck, 2008; Silver & Veness, 2010), or using other prioritized samplings methods. However, in the domain of human intelligence, we can leverage a second strategy to focus the space of modeled trajectories, and thereby render planning tractable: we can look to embodied cognition. While past experience may be a useful guide to the future dynamics of our coupled interactions with the world, the body offers a tremendous store of information forged across a long evolutionary history.

1.2.3 An embodied scaffold for the expected future

In this work, I consider the SGM to be embodied in two senses. Embodiment in the first sense is a consequence of the fact that the model both learns from and generates sensorimotor sequences—coupled neural representations that span the sensory and motor systems. Imagine, for example, the neural sequence that plays out during the experience of reaching for a door-knob in the dark (proprioceptive inputs allow us to approximate our arm position), until feeling cold metal against the palm (somatosensory input). Hence, the sequences produced by the SGM are reflective of prior and potential embodied experiences, not purely cognitive ones (Lewis, Purdy, Ahmad, & Hawkins, 2019).

Embodiment in the second sense is a central argument of the present dissertation. Theoretical work from the philosophy of embodied, extended, and enactive cognition suggests that a meaningful understanding of cognition must take into account dynamics well beyond the brain. Specifically, the enactive approach suggests that the mind should be seen as a dynamical system that acts to maintain itself, via homeostasis (Thompson, 2010). When viewed as a temporally extended process, the homeostatic imperative requires agents to identify a small volume within a much larger state space that satisfies its adaptive needs (Ueltzhöffer, 2018). While this volume may span enormous spatial distances, particularly for creatures leveraging internal combustion, the shape of this volume is fluid, spatially fine-grained, and rapidly updated in light of a constantly changing situational context. Hence, an individual's model of its own expected occupancy *enacts* its coupling with a changing environment: it retracts away from threatening future states, and regions with excessive uncertainty, and extends towards stable and rewarding locations. This description of the dynamics of an embodied generative model mirrors a key concept from ecological psychology with links to a number of more recent computational planning techniques: the affordance landscape, which I will explore in more detail in Chapter 3. In the context of driving, we can visualize the contours of this landscape—the opportunities for occupancy in the local environment—as what Gibson called the "field of safe travel" (see Figure 1.2a). Here, obstacles nearing this field preview potential challenges to the physical body, and trigger actions that modify the field in order to preserve its autonomy. Consistent with these ideas, therefore, I propose that the SGM, and the stochastically generated sequences it produces, is itself constitutive of an embodied entity that acts to satisfy the demands of (temporally extended) homeostasis.

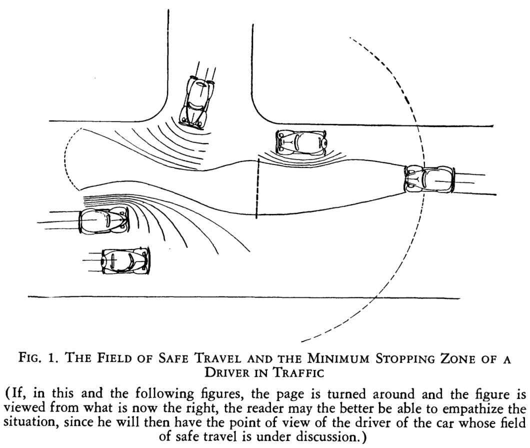
Evolution produced motor systems, bodies, and complex movements, well before a single action potential was fired. In light of this, I aim to press back against a tempting, but ultimately limited, cognitivist framing in which the prospective mind generates a future-aware world for the body to respond to. Rather, the body is seen as a source of adaptive expertise, the nucleus of the SGM, simultaneously enriching and constraining its dynamics as it seeks a robust path into the future[2]. Ultimately, by studying the body as it takes on tasks carefully designed to induce explicit prospective reasoning, we can take a small but important step towards a *kinematics* of embodied mental simulation. Unlike classical kinematics of the body, whose movements are limited by physical processes like force, tension and momentum, there are no established laws governing mental kinematics. How does the generative model expand and contract in response to perception of the environment? And how does the body serve to ground this parallel kinematics?

A recent influential account of mental simulation from György Buzsáki proposes an intriguing potential connection with embodied experience. In this proposal, the same muscular and neural mechanisms that allow us to interact with the world, also allow us, via corollary discharge, to navigate our internal spatiotemporal model of it (Buzsaki, 2019; McCloskey, 2011). In this theory, the activation of motor and premotor populations affords the covert enactment of plausible future episodes, and preparation of adaptive actions before they are needed. As the substrate for these simulations, the body scaffolds the high dimensional space of possible trajectories. A literature on *action simulation* has explored the recruitment of the motor system during these simulative processes (see Section 2.2), and provides encouraging evidence that their features may be illuminated via measurement of the body and behavior. Together, this view suggests that both covert and overt[3] motor activity may be expected to correlate with, and therefore offer measurable insights into, the way individuals navigate an

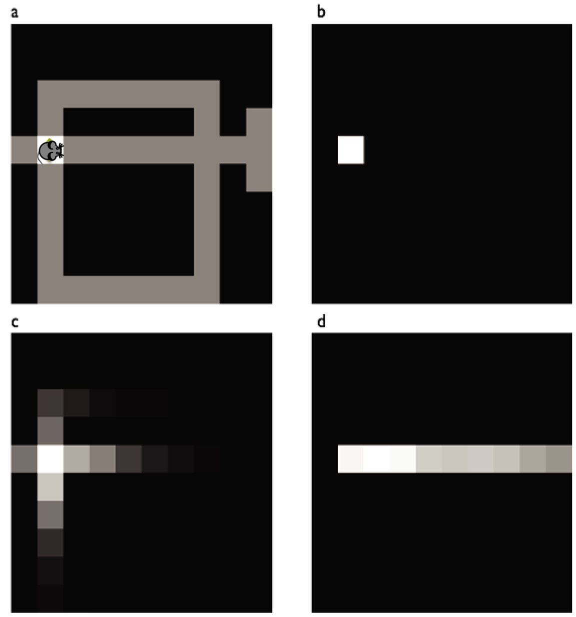
[2]Fascinating findings in ethology, e.g. in studies maze searching in the octopus, suggest that independent peripheral sensory streams can compose a hierarchical process in which the eyes guide gross arm deployments, while each conducts a finer-grained local search via touch (Gutnick, Byrne, Hochner, & Kuba, 2011)

[3]For example, the oculomotor activity underlying eye movements which, in some cases, such as several of the experimental paradigms explored in this dissertation

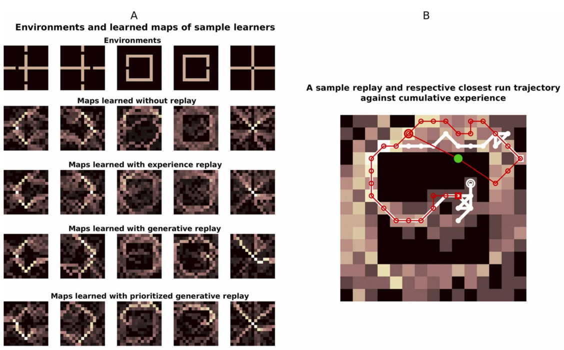
internal world model.



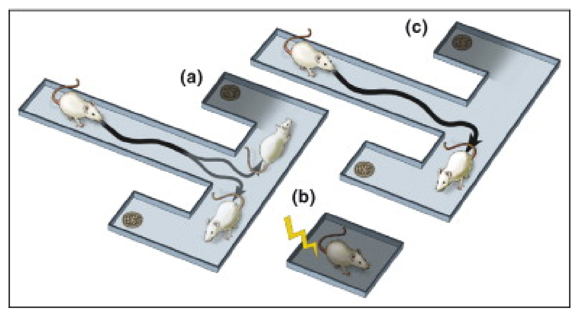
(a) Figure from (J. J. Gibson & Crooks, 1938): a Lewinian analysis of driving a car



(b) From top left, clockwise: environment map, punctate agent state, successor representation, and policy-conditioned successor representation. Figure from (Russek et al., 2017).



(c) Spatial maps learned by the hierarchical generative model proposed by (Stoianov et al., 2020).



(d) A "proto-form" of prospection in a T-maze experiments with mice. Reproduced from (Buckner & Carroll, 2007).

Figure 1.2: Figures reproduced, for conceptual illustration, from relevant other work.

1.3 Domain selection: planning during spatial navigation

To explore the kinematics proposed, it is useful to focus on domains and tasks that most clearly require effective future-oriented cognition. Specifically, tasks that cannot be successfully addressed with a reactive, local policy. Across all three of the primary studies in this work, I have chosen to focus on the domain of spatial navigation. As a modeling target, navigation enjoys a long history in computer science, as well as the planning literature, and appears as one of four categories included in Buckner and Carrol's ontology of self-projection (Buckner & Carroll, 2007). Given its natural low-dimensional spatial state space, navigation is especially well-suited to behavioral study. By selecting this domain, for which it is possible to precisely manipulate dimensionality, stochasticity, and perceptual attributes, I aimed to measure and model the SGM and its kinematics that are the focus of this research.

1.4 Research questions & summary of findings

The principal thesis behind this work is that by studying behavior and sensory exploration during navigation, we may begin to characterize the human solution to planning in an uncertain world, and take steps towards a kinematics of prospective mental simulation—the dynamic process by which individuals embody and interact with multiple possible futures.

Psychologists have gathered a long list of effects outlining key features of the way we learn about and model the world. Tendencies to infer causal relations (Kushnir & Gopnik, 2005), to perceive patterns (even when none exists, e.g., Fyfe et al. (2008)), to develop an intuitive understanding of physical systems (Krist et al., 1993), to categorize, and hierarchically arrange novel stimuli and task environments (Collins & Frank, 2013), are all suggestive of the processes by which such a model are learned and employed.

However, far less is known about the dynamics by which we navigate this internal model, its constraints and topology, or the mechanisms by which simulations traverse large spatial and temporal distances to most effectively instruct the action-oriented needs of the present. In short, we know little about the kinematics of mental simulation.

The work in this dissertation is guided by the following research questions, categorized into three areas of inquiry. For each, I list the motivating questions before previewing the main findings which will be presented in this dissertation.

1.4.1 Embodied planning

The first area of inquiry aims to bring an enactive, and more richly embodied account, to the primarily cognitivist theories that have been dominant in the cognitive science and psychology literature on planning.

Research questions:

1. How can efforts to study, and more carefully model, embodied dynamics help us understand a more naturalistic conception of planning?
2. What is the role of the body in prospective mental simulation during planning?

Summary of findings In Chapter 3, prior to getting to planning, which involves a sequence of decisions, I look first at decision-making in its own right, in order to highlight the limitations of some of the more reductionist models common within the classical decision-making literature. Some of these limitations include a) the inability of the serial decide-then-act paradigm to incorporate feedback from the motor system (decide-while-acting), b) the path dependence inherent to naturalistic decisions in which each choice alters the landscape of future choices, and c) the dynamic nature of continuously changing offers, which raises inaction (e.g. to wait for a more opportune choice set) as an often fruitful strategy. I also align classical and embodied decision settings via the formalization of affordance landscapes (present and future affordances) through their correspondence with economic probabilities and utilities. This chapter reviews a range of scholarship across neuroscience, psychology, and sports analytics, which demonstrate the way new models and a wider scope for experimental paradigms will help to realize the potential of this line of inquiry, and generate findings better able to generalize to the complexity of continuous reasoning and action in a changing world.

The design and findings in Study B are offered as examples for the kind of embodied planning setting Chapter 3 seeks to inspire. In this spatial navigation task, I report findings that arm orientation (which is also used to control direction during navigation) is predictive of initial heading choice, which is likely due in part to memory offloading during planning prior to action (Section 5.3.1). Further findings show that arm movements are increased when planning in maps that contain longer *unambiguous* segments prior to first fork—a finding that may be explained by offloading, as previously mentioned, or motor preplay of definite action sequences.

Additional findings related to the role of the body in the form of visual salience and eye movements, which are signals of particular interest in Study A and B, are covered in the next area of inquiry.

1.4.2 Visual salience and simulation kinematics in planning

In this second area, I focus on open questions within the domain of spatial navigation and planning, and explore the relationship between the dynamics of visual attention during the planning process.

Research questions:

1. How predictive of navigation choices is gaze during planning?
2. What graphical or spatial attributes are most predictive of visual salience (e.g. epistemic value, graphical connectivity, branching points, etc.)?
3. How do patterns of attention correlate with successful spatial planning?
4. Can we identify hierarchical dynamics during visually-guided planning, such as preferential attention at key decision points, temporal patterns in explore-exploit trade-offs, etc?
5. How do planning dynamics vary with respect to task complexity and the extent of uncertainty?
6. How well can an agent model, endowed with a sequential Monte Carlo-based forward planner, predict visual salience and temporal gaze kinematics?

Summary of findings These questions stem from an interest in the spatial characteristics of an underlying plan, under the assumption that both gaze and navigation trajectories are causally related to (and therefore correlated with) this planning process. Findings in both Study A and Study B help to support this key assumption. Visual search patterns prior to first move (in Study A), and at planning time (in Study B) were both predictive of initial navigation direction, and in the latter, also predictive of later choices at fork points in maps exhibiting significant participant navigation variability. I also find that planning duration was correlated with path efficiency on successful trials, but in the free-form plan-while-navigating paradigm of Study A, success required a balance of time spent planning versus navigation to effectively manage trial time constraints (see Section 4.4.1 for discussion).

A combination superficial spatial factors and deeper graphical connectivity attributes are found to predict visual salience. In Study B, key tile types (goal caches and muddy water), as well as critical tiles—those to which path connectivity and path-to-goal are most sensitive to changes—predict visual attention during planning. Similarly, both empirical and agent results in Study A show a preference for attention on high ambiguity locations—those for which traversal is most uncertain—as well as forks with multiple similar options. I interpret this finding as a resource investment in disambiguating between conflicting futures (especially when variance in expected return is high, as is the case with a possibly inaccessible jump) (Section 4.5).

A number of relationships are identified relating spatial salience and planning performance. In Study A, I observe a positive relation between attention distance and score, particularly in higher difficulty maps, suggesting the importance of forward planning in more challenging navigation settings (Section 4.4.1). Similarly, in Study B, I find increased correspondence between planning gaze and paths computed via larger planning depths, especially within higher uncertainty maps like those with one or more forks (Section 5.3.4).

Temporal analyses explored typical sequences of exploration, both in terms of spatial features and patterns in eye movement dynamics such as saccade distances and fixation duration. In Study A, I report a decreasing trend in (egocentric) attention distance through each trial (Section 4.4.1), which is consistent with the predominantly decreasing trend seen in gaze distance from origin found in Study B (Section 5.3.2). Additionally, various hierarchical dynamics are observed in planning gaze, such as a broad to narrow shift as plans are established, and improved correspondence between planning gaze and navigated route, particularly towards the end of planning among top performers (Section 5.4.1).

The agent models developed in this work, ADP (specified in Section 4.3), and then further developed as MC3 (see Section A.3.3), combine a gaze controller with a forward simulation mechanism capable of exploring sequential trajectories in parallel. In Study A, ADP exhibits similarities in the specific map features attracting visual salience, as well as strong correlations with human performance in map-level success rate and duration (Section 4.4.2). Results from Study B show a similar efficacy in visual salience prediction compared with static feature-map approaches, and in some cases demonstrate improved performance (see Figure A.5). Additionally, the proposed agent model brings the added benefit of generating spatiotemporal trajectories rather than static heatmaps, which exhibit some temporal characteristics corresponding with human participants' gaze kinematics (Section A.3.4).

1.4.3 Prospection in joint planning

In this final area of inquiry, I extend my inquiry focused on individual behavior during planning to pairs of individuals planning collaboratively towards joint goals. As discussed in relation to collaborative sensemaking in Section 2.6.2, the addition of another actor invites a wide variety of complex social dynamics that are beyond the scope of this work. However, there is sufficient evidence that mental simulation abilities are frequently leveraged for the understanding and anticipation of the behaviors of others (e.g. mirror neurons thought to provide vicarious learning during observation of others (Gallagher, 2007)), and therefore likely to play an important role in interactive tasks and joint problem solving. If so, understanding joint planning, and the process of strategy selection, convention building, and work splitting that comes with it, requires an understanding of the future-oriented representations that are developed during this process, and the efforts to share and synchronize these continuously during an extended collaboration.

Research questions:

1. How do collaborators coordinate joint strategies in a novel task, when verbal communication is unavailable?
2. Can an improved understanding of embodied planning help us understand how individuals reason about the consequences of actions and interactions with others?
3. How might this understanding aid collaboration towards shared goals?

Summary of findings The paradigm used in Study C (Chapter 6) involved a weaker embodiment than Studies A (which enabled continuous visuospatial exploration, despite the online setup) and B (which used an immersive fully embodied task), but still allowed me to explore questions related to embodied decision-making through the necessary usage of nonverbal communication strategies. These include participants' use of their own movement (and monitoring of the movement of others) to convey intentions, plans, or even requests, in a behavior sometimes referred to as *sensorimotor communication* (Pezzulo et al., 2018).

From Study C, I report findings suggesting a history-dependence of strategy selection when tasks dynamics change. This implies an asymmetry in either the expected cost of establishing a joint convention cost, the expected cost of its implementation, or both, dependent on which task context was presented first. The balance of work (i.e. fairness) was found to be lower in imbalanced environments, as well as during the more difficult blocks requiring greater partner monitoring. Finally, I find a persistent bias towards spatial separation (likely used to avoid interaction and enhance decoupling between collaborators), even among dyads using strict color-based strategies, highlighting that non-trivial hybrid approaches which mix conventions can be productively adopted even in non-verbal collaborative settings.

1.5 Contribution and importance

The scope of questions explored here are broad, multidisciplinary, and in some cases speculative to a degree that poses challenges to present-day empirical inquiry. Despite this, I aim to make a material contribution to the long-term project alluded to in the questions above: to develop a computational model predictive of measurable correlates of human prospective reasoning and planning, and to validate and improve this model via the guidance of human behavioral experiments.

The questions outlined above are of significant interest to cognitive science and psychology, as well as artificial intelligence efforts taking inspiration from human cognition[4], and the designers of systems to support human planners. Though they pose considerable challenges to empirical study, modern threads in embodied and enactive cognition suggest that measurement of the body can shed light on these internal processes. According to a recent argument from Yael Niv, behavioral measures may offer insights into not only *what* the brain does, but *how*, while also improving in a variety of ways upon traditional neuroscientific methods (Niv, 2020).

This investigation of embodied prospection builds on a diverse multidisciplinary literature spanning work on mental simulation and episodic foresight in psychology, predictive processing in philosophy of mind, hippocampal replay and preplay in neuroscience, embodied and enactive cognitive science, as well as a number of biologically inspired computational models of planning. Chapter 2 reviews these foundations.

With an improved understanding of the dynamics by which individuals generate expectations for the future, and plan for the future, we might one day design supporting tools personalized to individual preferences and idiosyncrasies, and perhaps oversights. Such systems may be able to communicate legible beliefs, risks, and hypothesized futures, and help enable more effective collaboration. Additionally, by better understanding how individuals reason about the future, we might confront our own limitations in preparing for long-term out-of-distribution phenomena such as the climate crisis and growing automation risks.

1.6 Outline

This literature review, arguments, and presentation of novel results (both human behavioral and computational) that compose this dissertation are organized into the following chapters.

In Chapter 2, I review several bodies of literature which inspired and provided a foundation for the theoretical interests of this dissertation. These include contributions from philosophy of mind, ecological psychology, neuroscience, and the cognitive sciences.

In Chapter 3, I present an array of arguments for the study of more naturalistic, embodied decision-making behaviors as an antidote to the reductionism that has both enabled and limited the ability of the sciences of the mind to understand and satisfactorily explain behavior outside of the laboratory.

In Chapter 4, I review a first foray into my empirical study of prospective planning (Study A). This work was designed to capture richer, more embodied data, during a visual-exploration and pathfinding task, despite the constraints to in-person research created by the pandemic. I also present an initial computational model of visual search during planning, called "Active Dynamical Prospection" (ADP), which became the foundation for subsequent iterations of the model.

In Chapter 5, I present Study B, an immersive virtual-reality based experiment extending the lines of inquiry and modeling begun in Study A, but using methods that enabled the capture of robust eye-tracking data during planning and navigation. The data collected allow me to explore in finer detail the outlines of the representation developed by participants

[4]In a recent talk Stuart Russell noted that a key missing piece of the modern artificial intelligence project is to develop models capable of long-range thinking at multiple levels of abstraction (Russell, 2023)

(what I have previously introduced as the SGM) as they develop an action plan robust to environmental uncertainty. In the Appendix for Study B (Section A.3), I describe an updated agent model, MC3, enabling generation of stochastic gaze trajectories, that improves prediction of participant gaze behavior during planning.

In Chapter 6, I explore the joint setting, in which two collaborators coordinate to take on a shared task, and confront tradeoffs of efficiency, risk, and fairness. This study highlights the importance of aligning future-oriented representations, as well as decoupling each collaborators information needs, in order to effectively coordinate in tandem.

In Chapter 7, I take a step back to consider a number of ways the predominantly theoretical work explored in this dissertation, as well as some of the methods developed, might be productively applied to real-world contexts, as well as important ethical considerations in doing so.

Finally, in Chapter 8, I summarize the work conducted, revisit the theory and argument for a more richly embodied perspective on prospective simulation, and highlight some of the many remaining open questions.

Chapter 2

Related work & theory

In this chapter I review theoretical and empirical contributions made across psychology, neuroscience, philosophy of mind, cognitive science and computer science, to illustrate the conceptual foundations upon which this work sits. My main focus here is on theories of prospection and planning, as well as work exploring the influence of the body on these and related processes. Towards the end, I include a brief methodological review of the use of eye tracking, one of my central empirical methods, within the cognitive sciences.

2.1 Prospection as mental simulation

Perhaps due to a mismatch in the ease of introspective access (indeed, according to Baumeister and Vohs (2016), we spend around a third of our lives anticipating the future) and the challenges of empirical measurement, the topics of mental simulation and planning have been approached by a range of scholars across many disciplines, and decades, and referred to under a range of labels.

2.1.1 Mental simulation

Among the names for related concepts to mental simulation are: constructive episodic simulation, self-projection, prefrontal synthesis, mental time travel, and episodic (p)replay. Though theorized about by William James, and philosophers and psychologists before him, simulation as a central heuristic strategy to evaluate the likelihood of future outcomes was brought to wider attention by Kahneman and Tversky (1981a). A key aspect of their proposal is that in evaluating the likelihood of a set of potential future events, we imagine each happening, assigning higher probability to those which are simulated more easily. The authors highlight connections with stochastic sampling methods as well: "...mental simulation yields a measure of the propensity of one's model of the situation to generate various outcomes, much as the propensities of a statistical model can be assessed by Monte Carlo techniques" (p. 2).

A number of classical results in psychophysics highlight notable early efforts to quantify the usage and dynamics of mental simulation. In one of the most famous of these, Shepard and Metzler (1971) showed that when participants are asked to rotate (and then match the result) of a visual schematic of a complex 3-dimensional object, response time is linearly proportional to the degrees of rotation. A similar relationship is observed for 2D images.

In an attempt to organize a diversity of related theories and results, Buckner and Carroll (2007) proposes an ontology of mental simulation processes, suggesting that they underlie four categories of cognitive abilities: episodic memory, prospection, theory of mind (conceiving the viewpoint of others) and navigation. Further, they argue that each is not distinct, but rather functionally related, and perhaps involving the same core neural processes. Prospection, the second of these categories, and the main interest of this work, is reviewed below.

2.1.2 Prospective simulation

Prospection refers to simulation of plausible future events, whether immediate or distant. This future-oriented subset of mental simulation has been claimed by some scholars to be so central to human cognition as to warrant the creation of a new field of psychology dedicated to its study (Seligman, Railton, Baumeister, & Sripada, 2013). Shortly after this call, Szpunar, Spreng, and Schacter (2014) proposed a taxonomy of prospective cognition with four modes—simulation, prediction, intention, and planning—with instances of each mode varying on an additional dimension of episodic to semantic. The last two decades have seen significant empirical work, primarily in humans, but also in primates, birds, and other animals (see Section 2.1.3), exploring behavior plausibly linked with the generation of likely future episodes.

A recent study in the psychophysics paradigm illustrates how mental simulations that are explicitly prospective in nature can bring about predictions for future outcomes; in this case, whether a ball bouncing in a box would pass through a target region (Hamrick, Smith, Griffiths, & Vul, 2015). Here, the authors leverage prior work proposing a binary decision-making model related to drift diffusion—sequential probability ratio test (SPRT)—in which an agent chooses the maximum a posteriori (MAP) hypothesis following a series of Bernoulli samples resulting in a parameterized decision threshold. By combining SPRT with a trajectory simulation model, they are able to predict response times for the bouncing ball task. Results validate the model's prediction that more simulations must be run in situations where evidence for a binary outcome is most balanced (and therefore ambiguous). Notably, in order to generate response time predictions using this model, an implicit assumption is made that simulations run sequentially, and mirroring Shepard and Metzler (1971), have duration consistent with the underlying physical process (here, movement rather than rotation). Models I present in Chapter 4 and Section A.3 diverge from this assumption in favor of a parallel approach[1].

In the context of behavioral economics and decision theory, intertemporal choice tasks are commonly used to evaluate peoples' (and animals), abilities to reason about and compare offers across varying delays. Some authors have suggested that prospection may allow us to overcome the challenges of temporal discounting (Boyer, 2008; Daniel, Sawyer, Dong, Bickel, & Epstein, 2016), in which payoffs in the future are out-competed by nearer term,

[1]The unitary vs probabalistic divide is controversial in cognitive science and philosophy of mind. For an argument against the common perception of simulation and imagery as "univocal," with a predictive processing lens, see A. Clark (2018). Similarly, in the domain of navigation, see arguments for hippocampal spatial representations existing along a manifold, rather than single unitary maps (Arnold et al., 2016)

but perhaps lower value, immediate gains. Others have extended this argument, and claimed that we as humans enjoy a unique capability: to pursue goals for days, months, and years without explicit reward (Bulley et al., 2016). In a model of intertemporal choice proposed by Bulley et al., the simulation of future possibilities and episodes feeds back two primary outputs into the decision-making process: the value of a future episode, and its likelihood. They note, however, that the interaction and weighting between these two information types is not well understood (p. 38).

A number of studies in developmental psychology and neuroscience have documented the time course by which future oriented thinking develops in humans during childhood and adolescence. For example, in an fMRI intertemporal choice study, Banich et al. (2013) found greater differentiation with age, when comparing patterns of brain activity for immediate versus future choices. Three regions showed this increasing differentiation: areas involved in current behavioral control, computing affective value to choice options, and imagining future outcomes. Similarly, in a delay discounting task, younger adolescents (under 16) were less likely to opt for a larger reward in the future as compared with counterparts aged between 17 and 30 (Steinberg et al., 2009). Hartley and Somerville (2015) proposes a possible adaptive benefit of this delayed development of prospective thinking, by which adolescents exhibit increased exploratory behavior (versus exploitation) supported by greater tolerance for uncertainty and more future discounting.

2.1.3 Prospection in other animals

Studies of the previously mentioned intertemporal choice paradigm are perhaps the most common class of attempts to evaluate prospective abilities in other species. However, many other compelling studies have documented behaviors suggestive of future-oriented representations. That we should find evidence for this cognitive ability in other animals is unsurprising, especially in light of prior arguments in Section 1.2.1 about the fundamental adaptive benefits (even to much simpler organisms) of prediction as control. Indeed, numerous examples exist across a wide range of species, e.g.: chimpanzees observed pick up stones to use as nut-crackers in other stone-lacking sites (Boesch & Boesch, 1984), and scrub jays learning to cache in private after having the experience of stealing from other jays' caches (Emery & Clayton, 2001).

However, the literature in animals highlights the challenges raised by studying internal cognitive processes without the benefit of (verbal) self-report (Clayton, Bussey, & Dickinson, 2003). There is no mirror test[2] for future thinking. Indeed, it would be tempting to set the threshold for classifying a cognitive process as "prospective" based on its ability to consistently generate actions resulting in adaptively positive future outcomes. As Pezzulo, Friston and others have argued, however (Pezzulo, Rigoli, & Friston, 2015), the informational imperative to learn a useful (temporal) model of an organism's reciprocal interactions with the world is a multi-scale process which can also be seen to transcend the life of any single organism via evolutionary pressures. These too, can be meaningfully seen as prospective, in the sense that the adaptations of the body and mind are themselves predictions of the sort

[2] By which I refer to what has become a standard assessment for self-awareness in both (young) humans and a wide range of animals (Gallup Jr, 1970), but which despite this widespread interest, has not avoided controversy around what in fact is being measured in this exercise (Kakrada & Colombo, 2022).

of niche a member of any particular species is likely to inhabit throughout its life, and the range of capabilities it is likely to require.

To highlight this ambiguity, we can take the example of the decision of when to flee as a predator approaches. A classical model from Ydenberg and Dill suggests that distance is selected based on the comparison of two cost curves: the cost of staying compared with the cost of fleeing (which have an inverse relation with present distance, and therefore intersect at the location of likely flight) (Ydenberg & Dill, 1986). To the extent that organism behavior is well approximated by such a model, we must still ask whether this finding implies a simulative capability (in which organisms generate the likely sensory consequences of flight at multiple distances as they are approached), or rather, reflexive mechanisms adapted over time to optimize a fitness function that is itself necessarily future oriented (since it operates over the entire lifetime of an individual, or perhaps communities of individuals). In the latter case, any number of (non-simulative) cognitive processes might function as heuristics (e.g. a set of context-aware thresholds of safe proximity) which trigger flight when exceeded. Despite the fascinating nature of these questions, I avoid the thorniest challenges of defining and evaluating these boundaries by choosing to focus on paradigms requiring the simulation-based, deliberative form of prospective reasoning which (while also of course a product of evolutionary processes) enables vicarious experimentation and feedback on much shorter timescales.

One cognitive ability especially frequently claimed to be the sole domain of humans is the ability to reason about hypotheticals, that is, to generate and consider *alternate* ways that the world—whether in specific episodes, or through manipulation of semantic properties—might be (R. M. Byrne & Girotto, 2012). This ability is, of course, closely linked to simulation in general, and prospection in particular (since the future, by definition, is never fully realized), but is particularly relevant to the present work, and reviewed independently in the following section.

2.1.4 Counterfactual reasoning

Inherent to many mental simulation processes is a particular type of simulation in which much of the modeled world is held fixed, while a single attribute or event outcome is changed. According to Judea Pearl, these precisely designed interventions can be seen as experiments that provide a central mechanism underlying causal inference. As he elegantly, though anthropocentrically writes: "every other creature can see what is. Our gift, which may sometimes be a curse, is that we can see what might have been" (Pearl & Mackenzie, 2018). Pearl proposes a ladder of causation mirroring the development of causal understanding. After observation (identifying sensory patterns), and intervention (discovering the effects of physical actions), is the highest rung, underpinned by imagination, which involves the conception and evaluation of counterfactuals. This third level, enabled by what Pearl calls the *do*-calculus, allows us to hold the world constant while "wiggling" a variable of interest to identify its causal influence.

In "The Rational Imagination: How People Create Alternatives to Reality," Byrne explores questions about the tendencies and idiosyncrasies of these human counterfactual abilities (R. M. J. Byrne, 2007). Citing evidence from a variety of empirical psychological studies, her primary argument is that imaginative counterfactual thought is based on the same principles underlying rational thought. She considers the dynamics that lead to peoples' consideration of some counterfactual alternatives rather than others, and the biases inherent to this process. Her findings indicate that people tend to consider a small number (due to working memory constraints) of likely alternatives (events that might actually happen). She also discusses a number of heuristics and biases at play, such as the tendency for people to favor recency in causal judgments (temporal anchoring). When considering expressions in the *subjunctive mood* (e.g. "if he had driven the car then he would have fastened the seat belt") listeners are cued to consider hypotheticals and uncertainties, and generate two scenarios at once (Quelhas & Byrne, 2003). Byrne also offers the concept of "unformed thoughts" as markers of ambiguity still remaining which can be deferred and resolved at a later time (Johnson-Laird & Byrne, 1991).

Another domain in which models of counterfactual simulation have been successfully developed is that of intuitive physics. Prior work has shown that people spontaneously run counterfactual simulations when asked to make judgments about causality (Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum, 2017). Several recent studies use the block towers to evaluate prospective and causal evaluation of the physical structure of a scene (L. Zhou, Smith, Tenenbaum, & Gerstenberg, 2023; Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2021), and develop a model based on a noisy physics simulator (aware of masses, friction, etc), called the Counterfactual Simulation Model (CSM). L. Zhou et al. (2023) propose three main steps to the generation of causal judgements: 1) representation with a mental model, 2) imagination of a counterfactual change acted upon that model, and 3) simulation of consequences. In the block tower support task, participants view an image of a particular tower of blocks, and are asked how many blocks will fall if a target block is removed. During perception of this stimulus, several types of uncertainty are in play: perceptual uncertainty about the exact location of each object, and dynamic uncertainty about how the scene will (prospectively) unfold, i.e. what will happen next when the target block is moved. L. Zhou et al. (2023)'s experiments highlight a close relationship between the level of responsibility the target block is assigned, and the number of blocks that would fall if the target object were moved. Interestingly, when the movement of several other blocks mediated the eventual disruption of other blocks in the scene, a lower responsibility was assigned to the target, indicating a divergence from this basic relation, and interpreted as a diffusion of responsibility (due to the role of other causal factors).

Together, these studies of human counterfactual reasoning help to shed light on the numerous ways people contend with an uncertain world, and attempt to develop useful representations of its causal dynamics. This work makes a strong case that a significant value of investing energetic resources in an internal dynamics world model is the ability to run simulations against this model: both to generate more precise future predictions, and to highlight divergence from realized outcomes (which may imply the need to adjust the model).

2.1.5 Planning

An extensive literature exists around planning across cognitive science, psychology, neuroscience, and artificial intelligence. As pointed out by Taylor, Pham, Rivkin, and Armor (1998), mental simulations are well-suited to support planning for a number of reasons, including the ability of simulation to recapitulate causal structure reflective of the true task, and to generate information that may be critical to identifying a solution, but which might otherwise go unnoticed.

A number of results in cognitive and reinforcement learning highlight a fundamental value for entropy-maximization that likely extends to the planning domain. Recent work by Rens et al. showed that people prefer options that open up more options (Rens et al., 2023). Similar entropy-seeking behavior (in which the expansion of future action-state occupancy was the sole motive) was shown to produce a range of complex behaviors without an explicit external reward.

Often, planning problems are posed in line with classical problem solving, in which the environment is fully observed with known dynamics and formalized as a Markov decision process (MDP). In this case, solution entails identifying a sequence of actions resulting in a goal condition (Newell, Simon, et al., 1972). This deterministic, fully observable, and durationless abstraction is well-suited to common problems used by neuropsychological tests assessing planning abilities. In the Tower of Hanoi, for example, participants are required to move disks of differing size between pegs, one at a time (and subject to additional constraints), in order to reach a goal state. However, in problems more typical of the real world, a number of complexities render the classical MDP formalism insufficient. These include continuous rather than discrete state spaces (since no two configurations of the world are identical), partial observability and sensor noise (resulting in uncertainty over the present state or configuration), temporally extended actions (requiring the scheduling of interdependent, potentially overlapping actions and effects), as well as significant uncertainty about the probabilistic dynamics of the world, and therefore the consequences of action. Much of the literature on classical approaches to automatic planning are out of scope for this work, however the most relevant biologically-aligned computational models are reviewed in Section 2.4.

2.2 Embodied action & simulation

The discussion of prospective simulation and planning, thus far, reveals a bias in much of the literature towards a cognitivist-leaning perspective that sees these processes as in many ways disembodied, or at least, plausibly implementable by an independent brain. And yet the acknowledgment of a deep interdependence between simulated experiences and their sensorimotor substrate is more than a century old. William James makes this connection in relation to memory: “For where should a past feeling be embodied, if not in the same organs as the feeling when present? It is only in this way that its identity can be preserved; a feeling differently embodied would be a different feeling” James, Burkhardt, Bowers, and Skrupskelis, p. 180.

Consistent with James’ position, it is now well established that visual imagery activates the oculomotor system (M. J. Spivey & Geng, 2001), and motor imagery produces measurable premotor and muscular outputs (Guillot et al., 2007); indeed, the latter phenomenon forms the basis for many brain computer interfaces when used as prosthetics. These findings are related to the wider study of action simulation, in which the brain systems involved in overtly executing goal-directed actions are also recruited during perception and imagination of these actions (Gallese, 2000; Pezzulo, Candidi, Dindo, & Barca, 2013; Lara, Gaona, Escobar, Pardo, & Hermosillo-Valadez, 2021). A range of adaptive benefits have been proposed, across fields such as motor control (Lepora & Pezzulo, 2015), motor learning, affective forecasting (Benoit, Szpunar, & Schacter, 2014), as well as empathy and social interaction (Gallese, 2005).

The potential to inform and nuance classical decision-making with theory and empirical work from embodied cognition extends well beyond action simulation. Efforts to identify shortcomings in classical decision-making paradigms, and develop new theories for more realistic accounts of decisions outside the laboratory have recently gained momentum. In embodied decision-making, agents are situated within complex, noisy, and uncertain environments in which, importantly, they must control both sensors and other motor outputs while simultaneously planning future actions in an online fashion. One critical role of the body in this decision-making process has been explored under the name *affective forecasting*. In one proposal, existing theory of mind systems may get re-used to anticipate the affective state of the 'future self' (Ersner-Hershfield, Wimmer, & Knutson, 2009). Bringing to mind a valenced stimulus can produce a cascade of physiological processes that together form an emotional response (Damasio et al., 2000). This embodied response may then be leveraged during reasoning and decision-making, as argued by Seligman et al.: "Affect is the brain's common currency for value, and conscious, subjective affect would permit the possible futures to be brought into the open for explicit comparison with each other" (Seligman et al., 2013). Lepora and Pezzulo's Embodied Choice framework demonstrates the limitations of considering decision-making a strictly serial process, and instead recasts it as dynamic and inherently active (Lepora & Pezzulo, 2015). In a visual decision task, the authors tracked the correlates of active consideration via movements of a mouse towards a target in a reaching task. Their findings suggest that two components: action preparation, and commitment effects, are critical to accurately model movement trajectories in ecologically valid paradigms where both speed and accuracy are required. In later work, Cos, Pezzulo, and Cisek (2021) demonstrated that perturbations to the arm during a similar reaching task can prompt changes of mind, further reinforcing the theory that deliberation continues dynamically during action execution. A more thorough review of the role of the body in cognition and decision-making is covered in Chapter 3.

2.2.1 Spatial navigation

Due in part to the density and variety of electrophysiological and optogenetic evidence from the study of rodents during maze navigation, as well as the implicit nature of planning required by this domain, spatial navigation has long been a focus of study among empiricists and computational modelers.

According to Montello (2005), navigation can be decomposed into two components: 1) locomotion, in which the body is coordinated to its local surrounds, and 2) wayfinding, in which a goal-directed agent plans actions aided by memory of both the local and distal environment. Studies on active navigation have investigated the relationship between sensory exploration and pathfinding.

In a psychophysics experiment, Arnold et al. (2016) showed that humans adaptively compress simulations of potential routes during prospective route planning. In a recent empirical study in primates Lakshminarasimhan et al. (2020) showed that eye movements could be used to infer latent beliefs such as the location of a hidden goal during virtual spatial navigation. The authors also found that controlling fixations had detrimental effects on performance, indicating the involvement of free visual search in effective navigation.

Edward Tolman's studies of rats during navigation provided some of the first empirical evidence and computational modeling of spatial decision-making in animals. One finding of particular interest is the observation of *vicarious trial-and-error (VTE)*, in which rats consistently move their heads back and forth at decision points in a maze to view alternative paths. In particular, the rate of VTE (percent of trials in which the rat looks one or more times before making an overt choice) was higher in trials requiring more challenging perceptual discrimination to identify the reward location (Tolman, 1938, p. 359). Though not possible at the time, later work showed that hippocampal place cells with receptive fields downstream of the decision point fire in synchrony with these visual scans, lending further evidence linking this behavior with prospective deliberation—the active previewing of multiple candidate futures.

Human and computational studies of navigation have identified a number of heuristics people use to reduce the burden imposed by exhaustive search (Bonet & Geffner, 2013; Russell & Norvig, 1995). One example is the use of a pruning strategy during mental search—when individuals identify a tree node that appears unpromising, the entire branch of the tree may be discarded (Huys et al., 2012). Another heuristic is to sample only a few promising routes, or truncated versions of a large number of routes (Keramati, Smittenaar, Dolan, & Dayan, 2016). It is possible to simplify the computational burden of planning by using a hierarchical approach, that is, by splitting a given problem into manageable, independent subproblems (Solway et al., 2014; Balaguer, Spiers, Hassabis, & Summerfield, 2016), or by switching repeatedly between planning and execution; for example, plan until a certain subgoal, then revise and elaborate on the initial plan during execution (Bonet & Geffner, 2013).

2.3 Neural underpinnings

2.3.1 Default mode network

Spontaneous simulation of future episodes, as well as recall of past episodes, is believed to involve a large and coherent network (Seligman et al., 2013) composed of the hippocampus, ventromedial prefrontal cortex (vmPFC), and dorsolateral prefrontal cortex (dlPFC). These areas were originally named the *default mode network (DMN)* as they consistently appeared in fMRI studies during activity measured between blocks; that is, when participants were not responding to any task in particular.

In one fMRI study investigating prospection, participants were asked to list 200 people and places and then imagine randomly chosen pairings (Benoit et al., 2014). They then rated the expected pleasantness of each episode. Analysis showed increased activation in vmPFC with more familiar people and places, as well as a correlation with ratings of expected pleasantness. The authors conclude that vmPFC codes for expected affective response to future scenarios, even novel ones. The finding that activation was not a simple sum of activation during each individual component, is taken as evidence for the integration and simulation of novel scenes in vmPFC.

2.3.2 Hippocampal formation as a spatiotemporal generative model

The discovery of place cells (encoding specific spatial locations) in hippocampus CA1, grid cells in medial entorhinal cortex (MEC), and a range of other putative "cell types" which selectively activate in response to specific spatio-relational configurations (e.g. head direction, proximity to a boundary, or orientation with respect to an obstacle), have inspired a reconceptualization of hippocampal formation function centering spatial memory and reasoning (Moser, Kropff, & Moser, 2008).

This theoretical framework has since evolved based on additional evidence in rodents showing a sequential reactivation of such cells during wakeful rest and sleep that is consistent with the temporal ordering of past trajectories. Thus, a theory of the hippocampal formation's involvement in the production of *internally generated sequences (IGS)* has been proposed, with direct links to prospection and planning. Summarized by Pezzulo, van der Meer, Lansink, and Pennartz (2014, p. 647): "internally generated sequences may be productively considered a component of goal-directed decision systems, implementing a sampling-based inference engine that optimizes goal acquisition at multiple timescales of on-line choice, action control, and learning." Recent proposals have gone further, suggesting that prior conceptions of the hippocampus as an episodic memory buffer may be too limited. Rather than storing and replaying past experiences with maximum feasible fidelity, the hippocampal system may constitute a spatiotemporal generative model (SGM) capable of *generative replay*: the sampling of fictive trajectories that may generalize true historical experience, counterfactual alternatives, and plausible futures (Stoianov et al., 2020). Evidence of prospective coding (the activation of place cells not yet visited, e.g. during vicarious trial-and-error) may support this framing (A. Johnson & Redish, 2007). Chersi, Donnarumma, and Pezzulo (2013) presents a computational model of a hippocampal-striatal circuit supporting simulated and overt action during maze navigation. In their model, place cells in the hippocampus respond to detection of a unique maze location, and connectivity between this subpopulation and dopaminergic neurons in ventral striatum encodes expected value. Covert simulation in an artificial motor cortex enables the agent to transform sensory input via motor imagery, generating a learned prediction of consequent sensory states.

A potential neural mechanism underlying sequence generation are sharp wave ripples (SWRs), an oscillatory pattern measured in hippocampus with EEG, which, according to Buzsáki, are "...the most synchronous population event in the mammalian brain" (Buzsaki, 2019). Buzsáki argues that memory, imagination, and planning may all be neurobiologically unified. "Overall, the selection of possible forward paths during sharp wave ripples can be conceived as an internalized vicarious trial-and-error process that flexibly 'imagines' real or fictive alternatives to select an optimal path or construct novel inferences without the need for movement-based exploration."

One potentially valuable feature of these forward simulations, which has garnered significant empirical evidence across animals and humans, is that simulation avoids not only the risk but also the latency of true action in the world. That is, episodic simulations may take on their own subjective timescale, which plausibly enables the anticipation of longer term outcomes. This subjective time could theoretically be achieved by a flat accelerated simulation, but evidence exists that temporal abstraction—organizing remembered and anticipated episodes by extracting discrete events—is a more likely mechanism to enable a faster simulative ability. The identification of boundaries between remembered sub-episodes is a well-studied topic known as event segmentation (Zacks, 2020). A number of computational models have been proposed to endow artificial agents with the ability to make longer-timeframe predictions by ignoring low-information intermediate events. In Zakharov, Crosby, and Fountas (2020), an active inference agent computes the salience of each event it experiences using prediction error (Bayesian surprise) to identify key boundaries across which to learn transitions. Baldassano et al. (2017) apply a Hidden Markov Model to human fMRI data during perception of naturalistic narrative stimuli. Monitoring transitions in the model's latent variables found correspondence in patterns of activity across cortex, which showed relative stability during consistent events, followed by rapid transitions at event boundaries.

Finally, evidence for hippocampal formation involvement in coding entities more abstract than spatial locations and trajectories (e.g. bird shapes varying across multiple dimensions in Constantinescu, O'Reilly, and Behrens (2016), and abstract relational knowledge in Garvert, Dolan, and Behrens (2017)), suggests that this generative model may be capable of supporting a wide array of reasoning tasks beyond spatial navigation. In line with this more abstract relational capability of the hippocampal formation, Behrens and colleagues have proposed a model called the Tolman-Eichenbaum Machine, which extends Tolman's theory of cognitive maps into a system capable of representing more abstract relational structure (in entorhinal cortex), as well as to generalize this relational knowledge to novel contexts, including non-spatial ones (J. C. Whittington et al., 2020; J. Whittington, Muller, Mark, Barry, & Behrens, 2018). This theory suggests that regularities in structure (e.g. similar connected graphical representations applied to social relationships, logical comparisons, and spatial environments) enable generalization by decoupling this structure from "sensory objects". Recent work by the authors has suggested that this decoupling may provide the foundation for constructive reasoning and imagination of not yet experienced events through the mechanism of hippocampal replay. Specifically, by generating state spaces via composition of already learned structural primitives, the hippocampus may enable previewing realistic sequences of novel states (Bakermans, Warren, Whittington, & Behrens, 2023).

2.4 Biologically inspired models of planning

2.4.1 Model-based reinforcement learning

The reinforcement learning (RL) paradigm is one of an embodied agent situated in an unpredictable world. During the early development of RL, researchers used methods such as Dynamic Programming and focused on systems where agents had access to a perfect simulator of the world (R. Sutton & Barto, 1998b), but contemporary developments acknowledge that real systems never have such a guide, and must instead approximate it. Model-based reinforcement learning is a subclass of algorithms that learn to approximate the transition dynamics of the environment in the form of a world model, and exploit this model to select goal-directed actions. These approaches are contrasted with model-free strategies such as Q-learning, which learn only a 'habitual' policy: a distribution over the best actions given the present state, and therefore need not predict the effects of actions explicitly.

A unifying concept central to both RL approaches is the learning of a value function which represents the approximate goodness of any particular state an agent might find itself in. Value functions are inherently forward looking, as they seek to collapse the complexity of the expected future into a unidimensional score (an approximate utility) simplifying action selection in any given state. This "forward stance" is at the very heart of the RL paradigm. Below, I review a few examples of algorithms most parallel (whether via mimicry or convergence) to the presented conceptions of human mental simulation.

A wide-ranging research program introduced by Jürgen Schmidhuber in 2015 lays out a high level vision for the future of model-based RL (Schmidhuber, 2015). He proposes a large world model M (usually implemented as a recurrent neural network) capable of predicting future states given present input. This model is then paired with a controller C that is able to flexibly "think with M" by passing in sequences of inputs, running queries, and reading outputs, synchronously with world interaction. The ideas presented show many parallels with mental-simulation-based reasoning within a generative world model, including a discussion of online queries, in which C interacts with M during, and in support of, ongoing overt action.

An example following from these ideas was demonstrated in David Ha's "World Models," in which a model-based system (composed of a Variational Autoencoder-based (VAE) vision module, and an RNN-based memory), is paired with an evolution-optimized controller to perform well on a number of common game environments (Ha & Schmidhuber, 2018). Interestingly, in more complex environments, the authors find that training the agent within its own learned dynamics model (learning via hallucination), was efficient and especially effective when a temperature parameter was used to increase uncertainty in the learned model. These results encourage a deeper consideration of the dynamics of forward mental simulation, and the potential benefits of learning transition dynamics that differ from those experienced, e.g. with higher uncertainty.

In "Imagination-Augmented Agents for Deep Reinforcement Learning," an algorithm is proposed that uses a learned transition model of the world in addition to a directly model-free approach (Weber et al., 2017). This algorithm equips agents with the ability to learn which approach to use (simulated or reflexive) conditioned on a particular context. The authors propose an "imagination core," composed of the learned environment model and an encoder (implemented by a recurrent network known as long short-term memory (LSTM)) which encodes and aggregates multiple imagined trajectories into an "imagination code," which in turn is used as context for action selection in the policy network.

Monte Carlo methods are frequently used in RL to sample trajectories during value estimation, and therefore to support the planning of future actions. One example, famous for its success in planning through huge decision trees in complex games like Go, is Monte-Carlo Tree Search, an algorithm that combines full Monte Carlo rollouts (simulations of potential trajectories generated by sequential sampling), with value function learning in a smaller tree of likely trajectories beginning from the present time-step (Chaslot et al., 2008). Building on MCTS, Silver and Veness (2010) proposed Partially Observable Monte Carlo Planning (POMCP) to make value estimation tractable in high dimensional state spaces.

In this work, particle filtering is used to efficiently approximate belief state updates when access to the true generative process is not available.

2.4.2 Successor representation

Another elegant computational model proposed as a biologically plausible implementation of spatial transition dynamics is the successor representation (SR) which was developed by Peter Dayan to improve temporal-difference (TD) learning techniques in reinforcement learning (Dayan, 1993). The SR uses a simple store of discounted expected occupancy of a destination state s', when originating from a discrete state s, often formalized as the matrix $M(s, s')$. The SR may offer a useful middle-ground between high flexibility model-based methods, and computationally efficient model-free methods (Gershman, 2018). By providing an agent with access to a weighted distribution of likely future states, this representation, which is also easily learned using a TD update rule, renders value estimates computationally trivial as the product of the reward map and present state SR (see the bottom two figures in Figure 1.2b).

The successor representation has also been applied to the classic Dyna algorithm (R. S. Sutton, 1991), which uses Monte Carlo simulations from a remembered "prior experience buffer" to avoid catastrophic forgetting, a persistent problem in the domain of continual learning. In SR-Dyna, offline replays are used to update a state-action version of the SR matrix $H(sa, s'a')$, and the value function and policy are induced in the usual way (Russek et al., 2017).

2.4.3 Planning as (active) inference

A conception of planning as probabilistic inference has been particularly influential in computer science and computational neuroscience. Citing weaknesses in typical approaches to goal-directed planning in RL, Botvinick and Toussaint introduced planning as inference (PAI), dependent on an internal generative model encoding the joint distribution of states, actions, and rewards (Botvinick & Toussaint, 2012). In PAI, the traditional expected reward maximization process is inverted, and planning is seen as a process of inverse inference that *follows* conditioning on the expectation of rewarding states. The agent acts to minimize the difference between two probability distributions: the expected state-action occupancy given the current policy, and the same distribution conditioned on the assumption of future reward. Botvinick notes, however, that PAI must further engage with evidence from classical cognitive research showing that human planning is defined by a range of simplifying strategies and heuristics, subject to the constraints of working memory, when solving complex planning problems (Botvinick & Toussaint, 2012, p. 488).

Within the same theoretical lineage as PAI, is a framework called *active inference*, in which agents resolve prediction errors (divergences between the present model and sensory evidence) in two ways: 1) by adjusting the model towards perceived evidence, and 2) by orienting available sensors in such a way as to conform sensory data with expectations, and therefore to disambiguate competing perceptual hypotheses (Seth, 2014; Friston, FitzGerald, Rigoli, Schwartenbeck, & Pezzulo, 2016). When applied to the setting of visual inference, percepts can be seen as hypotheses (e.g. *the dog is hungry*, or *the man is planning to buy* *some flowers*), and the ocular motor outputs that produce saccades (eye movements) are seen as experiments aiming to confirm or deny prior beliefs (Friston, Adams, Perrinet, & Breakspear, 2012).

2.4.4 Limitations in existing models

The models reviewed above offer a range of perspectives on normative answers to the questions of efficient prospection and planning. However, none of these models yet offer complete accounts capable of making predictions about the kinematics of simulative processes theorized to underwrite planning. While a conception of mental simulation—either to maintain learnings from past experiences, or approximate the value of a given state—is often prominently featured, few models, if any, center the body in a mechanistic account of counterfactual reasoning or planning. As such, there is an opportunity to extend existing models, or propose new ones, that can explain a number of theoretically predicted and empirically observed behaviors. These include: the interaction between simulated future action and ongoing activity in the present; a description of attentional dynamics, and how simulations are selected to iteratively tune model dynamics; the conditions under which the motor system is (measurably) recruited into simulative processes; and finally the specific ways in which embodied simulation may support the detection of opportunities to improve the model and plan.

2.5 Eye tracking in cognitive science

Both theoretical interest in eye movements, and eye tracking as a scientific methodology, enjoy an extremely lengthy history, as well as a rich modern resurgence in the present. Even early philosophers wrote about the ability of the eyes to somehow internally capture aspects of the outside world, as well as the informational potency of the eye movements that make this possible. Apicius, an ancient Roman who wrote one of the earliest cookbooks, alludes to the way vision allows us to preview experiences even before they arrive, for example, to evaluate how a new food might taste (Apicius, 2016). And from William James, we see allusions to the volitional and interactive nature of visual sensing, an early preview of the enactivist rebellion against notions of passive information processing in the brain: "Millions of items of the outward order are present to my senses which never properly enter into my experience. Why? Because they have no interest for me. My experience is what I agree to attend to" (James et al., 1890).

However, moving beyond these intuitions into models capable of predicting or interpreting quantitative eye tracking data has been a much harder and much slower to develop project. The origin of the modern quantitative study of eye movement is attributable largely to Alfred Yarbus, who devised the first system capable of capturing precise eye movements[3](Yarbus, 1967). Yarbus' studies and writings make the case for an understanding of eye gaze that

[3]Yarbus' system involved beaming light off of an angled mirror attached to the cornea by a suction cup, and capturing the movement of this beam on photographic paper. According to his own account, participants noted intense pain after 2-5 minutes of this procedure, a reminder of the benefits of modern technologies, and a better-regulated research ethics review process, we benefit from today.

is interactive, contextual, and often goal-directed. Further, he noted that the scan paths captured may be used to make inferences about internal states of mind (e.g. the task at hand, epistemic states, etc).

Such a finding is consistent with the driving thesis of embodied cognitive science, and 4E cognition more broadly. Embodied and extended cognition ask us to see the mind as a larger system that encompasses the brain, body, tools, and even social interactions, and highlights the way bodies leverage motor systems to offer a form of embodied intelligence, and goal-oriented behavior that predates complex central nervous systems. In the 4E paradigm, organisms are fundamentally seen as control systems pursuing allostasis—by which the needs of the body are anticipated and acted upon before they arise—by interpreting internal (interoceptive) and external (exteroceptive) sensory evidence (Sennesh et al., 2022). This idea, that cognition can only be reasoned about with an understanding of what it is for—movement and interaction (James et al., 1890), in service of an allostatic imperative—highlights the necessity of studying signals captured from the body, and indicative of this ongoing predictive process, in cognitive science.

Recently, a new wave of eye tracking research has benefited from a combination of reducing costs of hardware—to the point that eye trackers are beginning to be found in consumer-facing products such as VR headsets and laptops—as well as improved models and analytic techniques that have significantly augmented our ability to interpret eye movements and their relation with internal mental processes. Next, I will review several recent studies exploring the use of eye tracking to infer various kinds of internal states of mind, and then, more central to the present work, its uses in detecting and understanding planning processes.

2.5.1 Inferring internal states via gaze

One recent approach is to fit a generative model to observable gaze data, and then correlate temporal patterns seen in the model's hidden states to structure in the task at hand. A recent example of this in a human visual search study used a Dynamic Bayesian Network (DBN) as a task-independent representation of cognitive state, and played the role of a bias for the production of saccades (Joseph MacInnes, Hunt, Clarke, & Dodd, 2018). The DBN learned conditional probabilities governing saccade attributes like velocity and duration, as well as transitions between cognitive states. This model was then leveraged to generate synthetic saccades across multiple task types, and compared with statistics from human participants. Notably, this model affords additional validation by analysis of temporal dynamics such as return angle, which reflected patterns in the human data despite not being directly modeled.

Shimojima et al. studied inferential processes as participants reasoned about diagrams expressing a variety of conceptual relations such as "A is lighter than B" (Shimojima & Katagiri, 2013). When audio information was provided which augments the relational information provided by the diagram, eye movements show a covert or simulated "drawing" process that aids inference in locations in the diagram consistent with the information already available. Response time delays and error rates further indicated the influence of the spatial constraints of the diagram (e.g. tasks were more challenging when the position of newly inferred concepts could not be definitely located). This work reinforces a notion of visually guided simulation that is fundamentally linked with structure in the world, as opposed to purely abstract and cognitive. Such a finding helps to validate the assumed relation between gaze on external information sources and apparent correspondence with spatial aspects of an internal representation, which supports some of the design decisions of Study B.

One recent eye tracking study highlights enactive claims that action is part and parcel of cognition. Krajbich, Armel, and Rangel (2010) augment the classic drift diffusion model in a two-alternative forced choice (2AFC) task by additionally considering the currently fixated object, and then adjusting the Gaussian Process dynamics with a bias towards the attended option. By doing so, their model is able to make predictions about the relationship between (temporal) visual attention patterns and choice results, e.g. that the last fixation is predictive of choice, and that longer fixations on option A can only be overcome by longer fixations on B (to ultimately choose B).

Recurrence Quantification Analysis (RQA)[4] is a method developed for the analysis of dynamical systems, and applied to a wide range of biosignals including EEG, electrocardiography (ECG), and blood pressure. A number of studies in psychology and cognitive science have also begun to leverage RQA for a deeper temporal analysis of eye movements. In one example, Gurtner et al. studied mental imagery processes, in which participants shown visual images, and then after removal of the stimulus, imagine the image just seen (Gurtner, Bischof, & Mast, 2019). Gaze patterns during visual processing of the image, and the simulation (via mental imagery) in the second stage, are compared. Their analyses find higher determinism (DET) during imagery indicating imagery-based gaze sequences are "re-enactments of prior sequences," as well as longer median fixation time, and fixation targets closer to the median (more clustered), during mental imagery. One interpretation for the increased determinism (indicating gaze returned more often and sooner to previously inspected areas) during imagery is the need to support reactivation and maintenance of the visual memory, which is not needed during perception of the image given environmental availability.

Eye tracking is also of interest to education scholars to better understand the ways students and those with varying levels of expertise interpret visual imagery, especially information displays common across a range of disciplines. In one recent study of graph interpretation, eye tracking data was captured from non-science undergraduates, science undergraduates, and science educators as they viewed a variety of figures displaying data, alongside a question requiring its interpretation (Harsh et al., 2019). More experienced participants were found to spend more time viewing the graph data earlier in the interpretation period, as compared to less experienced who focused more attention on the question and answers and other cues related to the task at hand. Interestingly, this work also elicited an "interpretation plan" from participants prior to graph viewing, and compared the planned sequence of AOIs (e.g. variables, title, question) to the AOI sequences recorded during eye tracking. Lower expertise participants had similar plans, but diverged from these plans more often according to eye tracking data as compared to more experienced groups. These results overall indicate greater visual information search skills are developed with expertise through increasing alignment between planned search strategies and those enacted.

[4]RQA takes in a, typically discretized, timeseries of states $X(t)$ and first generates a recurrence plot, constructed by marking each cell $RP(i, j)$ black when $X(i) = X(j)$. This 2D plot is then analyzed to generate a variety of metrics which correlate with various temporal dynamics of the timeseries, including recurrence rate (RR): the average density of recurrent states, determinism (DET): percent of states forming diagonal lines, or temporally extended repeated sequences, and several others.

2.5.2 Inferring plans via gaze

I now turn my attention to studies with a focus on prospective or planning-based processes and their correlations with eye movements. While it is clear from the literature reviewed above that eye tracking data contains rich signals enabling an understanding of various internal processes such as inference, conceptual reasoning, and mental and motor imagery, the domain of planning is unique in the inherently sequential nature of multi-step action reasoning, and the handling of uncertainty which may prompt multiple (possibly conflicting) contingencies to be considered. I ask, then, how good of a proxy does gaze provide for the detection and interpretation of a plan?

In one study related to this question, Grison et al. look at gaze behavior as individuals reason about transit routes between a given origin and destination (Grison, Gyselinck, Burkhardt, & Wiener, 2017). Analyses divided the trial duration into two phases: search for origin and destination, and path planning. Features such as the ratio of time spent looking at peripheral versus central stations, and the ratio of time spent looking at origin and destination stations, within each phase, were extracted and analyzed. This study found that saccades during search were longer than during planning, and that saccade amplitudes decreased in more complex tasks (a finding in line with prior work such as Pomplun, Reingold, and Shen (2001)). Pupillometry revealed larger dilation during the more cognitively demanding planning phase, which required consideration of possible transfers between lines and other connectivity considerations going beyond the initial search for key start and end landmarks.

In another study, the use of eye tracking data to predict future navigation choices and implicitly communicate intention was explored in the context of human robot interaction (HRI). Chadalavada et al. performed an AOI analysis of gaze data during pedestrian navigation around a moving robot encountered in the path (Chadalavada, Andreasson, Schindler, Palm, & Lilienthal, 2018). A trivial intention predictor based on binary AOIs was shown to be accurate in 72% of trials.

Using a more complex wayfinding paradigm, and one upon which some aspects of the map geometry for Study C were based, Zhu et al. study visual search patterns prior to navigation through a virtual landscape (Zhu, Lakshminarasimhan, Arfaei, & Angelaki, 2022). Findings show dynamics consistent with rapid sequential 'preplay,' as well as a balance between fixations on reward locations and critical transitions (walls or openings) in the environment. Another study of visual search in humans has shown, further, that eye movements themselves exhibit behavior consistent with multi-step planning (in this case, planning of a gaze sequence matched to the shape of a given search space), rather than greedy selection of informationally rich or otherwise salient locations (Hoppe & Rothkopf, 2019).

Together, this literature demonstrates a variety of techniques for leveraging eye tracking data to better understand human planning and decision-making, especially as it relates to spatial navigation. However, many open questions remain about the temporal characteristics of eye movements during planning in complex naturalistic tasks, and no models capable of producing plausible fixation sequences during planning, have been developed. Filling some of these gaps is a central project of the remainder of this dissertation.

In Chapter 3, I will present a series of arguments for the importance of studying embodied dynamics in more naturalistic settings. These arguments build upon much of the literature reviewed above, and go beyond planning, by looking at the realm of decision-making at large, and highlighting some of the ways richer generative models might help us build more robust theories of cognition and action across a variety of disciplines.

2.6 Joint reasoning and strategy selection

Scaling up theories of cognitive processes to even a second participant (let alone larger social groups), introduces enormous complexity that we cannot assume is captured by models designed to understand individual behavior. Though the main focus of this dissertation is on the dynamics of a single mind reasoning through spatial (and therefore not, primarily, social) problems, work in Study C (Chapter 6), and some of the speculative applications proposed in Chapter 7, break this constraint and do take an interest in how multiple individuals might develop, share, and act upon corresponding beliefs about the future.

2.6.1 Collective decision-making

Prior work on collective decision-making and action coordination has shown that features of the information channels available between collaborators (and their environment) is critical to any cooperative pursuit. In particular, the kind and frequency of feedback provided about the performance of others influences a dyad's ability to effectively divide task demands. As one example, in a cooperative visuospatial tracking task, dyads received different kinds of information about the other's performance following each trial (Wahn, Kingstone, & König, 2017). Dyads receiving both performance information (partner efficacy), as well as target selection information (partner's task choices), performed better than dyads receiving only one type of feedback. Teams can also learn to decompose work even in the absence of objective feedback (e.g., the performance of other team members), through verbal or non-verbal social interaction (Bahrami, 2012; Bahrami et al., 2010; Pezzulo, Roche, & Saint-Bauzel, 2021).

An additional example of this is a study of motor coordination in the absence of online perceptual feedback (Vesper, van der Wel, Knoblich, & Sebanz, 2013). Dyads who could not see or hear one another were tasked with performing jumps, given independent target distances, that landed at the same time. Participants were able to coordinate their jumps successfully, adjusting their motion onset timing and trajectories to land at a similar time, even when target distances were different. Of particular interest to the present work, this study also showed that participants converged upon an asymmetric (and therefore perhaps "fairer") convention to solve the task, where the jumper with the shorter distance adjusted their action more than their partner with the larger distance (and therefore the harder task). Motor simulation is putatively the driving mechanism behind this result, as the shorter jumper must evaluate the time dynamics of their partner's task even as they plan their own movement.

2.6.2 Collaborative sensemaking

An alternate literature on sensemaking, which applies social psychological theory to organizational behavior, raises useful concepts that specifically aims to apply to the establishment of collective understanding (within organizations and other multi-party settings) (Weick, 1995).

Helpfully, sensemaking fits somewhat naturally into the paradigms of cognitive science closest to the core foundation of this work (e.g. the embodied homeostatic imperative discussed in Section 1.2.3, and the *4E* approach discussed in Section 2.5 above), due to alignment in several of its core principles. First, its recognition of the central role of inference based on the extraction of contextual cues, is consistent with the attention-controlled approach to prediction error minimization found in active inference (e.g. individuals' ability to bring their model in line with sensory evidence by adjusting their sampling to model-consistent observations). Second, that individuals actively participate in a continuous and recurrent exchange with others, and that this exchange constitutes *enaction* of a shared understanding, resonates deeply with the embodied and, indeed, enactive accounts of cognition (Weick, Sutcliffe, & Obstfeld, 2005). For example, a correspondence can be seen between Weick's "sensemaking as activity that talks events into existence," and Thompson: "a cognitive being's world is not a pre-specified, external realm, represented internally by its brain, but a relational domain enacted or brought forth by that being's autonomous agency and mode of coupling with the environment" (Thompson, 2005).

Weick claims sensemaking is triggered by "disruptive ambiguity," helps to formulate a plausible story of the situation at end, and ultimately "answers the question 'now what?' [which emerges] from presumptions about the future, articulation concurrent with action, and projects that become increasingly clear as they unfold" (Weick et al., p. 413).

2.7 Conclusion

In this section, I reviewed a selection of theoretical frameworks and empirical findings which provide a rich and diverse foundation for the main efforts presented in this dissertation. In the next chapter, I will present arguments towards a deeper commitment to understanding the role of embodiment in planning and decision making, and further reference aligned projects in related disciplines including ethology, robotics, and ecological psychology.

Chapter 3

Affordance landscapes as embodied constraints in planning and decision-making[1]

3.1 Introduction

The ability to make effective value-based decisions is crucial for the survival of living organisms. Despite its popularity in psychology and neuroscience (Gold & Shadlen, 2007), value-based decision-making is often studied in restricted laboratory settings, using simple tasks that are inspired by economic theory, such as binary choices between lotteries (e.g., do you prefer 50 dollars with 20% probability or 20 dollars with 50% probability) or intertemporal offers (e.g., do you prefer 20 dollars now or 50 dollars in a month?). In this "classical" setting, there are a limited number of choices to be selected (often two) that are prespecified by the experimenter and presented simultaneously. Furthermore, the relevant choice dimensions (e.g., rewards and probabilities) are usually fixed throughout the experiment and unambiguous. In human studies, choice dimensions are usually labeled with symbols and easy to identify. Similarly, in animal studies, choice dimensions become stereotyped and unambiguous as an effect of long learning periods. Finally, the action component is often trivialized, not just because the action itself is simple (e.g., a button press) but also because there is no effect of the action on subsequent perception—i.e., there is no action-perception loop. The experiments conducted in this classical setting have crystallized a serial view of decision-making, which identifies three distinct sequential stages (Fodor, 1983; Pylyshyn, 1984): the perception of the attributes of the predefined alternatives, the decision between the alternatives, and the reporting of the decision by action (i.e., decide-then-act).

While using the classical setting permits methodological rigor, this comes at the expense of an excessive focus on just one kind of choice—economic choices that can be most easily studied in the lab—while disregarding the fact that there are other kinds of choices that are equally (or even more) frequent in our lives and important from an evolutionary perspective. The classical setting is relatively well suited to certain types of choices, such as choosing

[1] This chapter is based on work published in Neuroscience and Biobehavioral Reviews under the title "The road towards understanding embodied decisions"(Gordon, Maselli, et al., 2021)

between a fixed number of dishes from a restaurant menu. However, these are just a small subset of the choices that we make in our everyday lives. For example, on the way to the subway station one makes countless choices about actions, such as whether to cross before or after an oncoming vehicle at an intersection, how to effectively pass by a group of other pedestrians on the sidewalk, which seat to take, etc. In many of these situations, the choices themselves as well as their costs and benefits are in constant flux, and they are difficult to model using classical concepts from economic choice settings.

Furthermore, classical economic choices are not the kinds of scenarios that drove the evolution of the brain's neural mechanisms. Instead, throughout its long history the brain adapted to deal with very different choice situations, such as deciding whether to approach or avoid an object, how to navigate to feeding sites, and how to move among obstacles (some of which might themselves be in motion). These kinds of "embodied decisions" dominated animal behavior long before humans existed, and are accomplished by neural mechanisms that have been conserved for hundreds of millions of years (Rodríguez et al., 2002; Saitoh, Ménard, & Grillner, 2007; Striedter & Northcutt, 2019). Furthermore, the innovations of neural circuitry introduced in mammals and more recently, primates, were all made within the constraints of that ancestral context (Cisek, 2019; Passingham & Wise, 2012). Even now, in our daily lives, humans continue to perform many similar embodied decisions every day, such as when we play a sport, drive on a busy road, prepare a meal or play hide-and-seek with children, and it has often been suggested that our cognitive abilities are constructed upon a scaffolding provided by the sensorimotor strategies of such embodied behavior (Hendriks-Jansen, 1996; Piaget, 1952). Thus, one could argue that understanding embodied decisions is of primary importance for understanding many aspects of human cognition and behavior.

3.1.1 Embodied decisions

Embodied decision settings are very different from classical settings for a number of reasons (Cisek & Pastor-Bernier, 2014) (Figure 3.1). First, the number of offers that enter the deliberation is not predefined and can vary on a moment-by-moment basis. For example, it is likely that a lioness in front of hundreds of gazelles does not consider each as a separate offer, but instead clusters them into "patches" of potential food sources. Second, choice offers and their dimensions are rarely discretized or labeled with symbols and units (e.g., "option 1", "option 2", "dollars", "percentages"). Rather, the decision-maker has to identify the options perceptually (under uncertainty) and to select the relevant choice dimensions; these choice dimensions often include geometric dimensions and affordances, such as the relative lion-gazelle distance, which are disregarded in classical settings.

Finally, and most importantly, perception, decision and action dynamics are intertwined. The action component is not simply a way to report a choice but an essential component to secure the reward (i.e., the lion has to actually chase the gazelle). Furthermore, action dynamics change the perceptual landscape (e.g., one or more gazelles can go out of sight when the animal starts moving) and the decision landscape (e.g., some gazelles can go out of reach and hence cease to be valid offers, or conversely become more available and extend the offer menu). Action and decision dynamics cannot be fully separated in time and are instead "continuous" (Yoo, Hayden, & Pearson, 2021). The decision-maker can start acting before completing the decision, to buy time (Barca & Pezzulo, 2012; Pezzulo & Ognibene, 2011) or

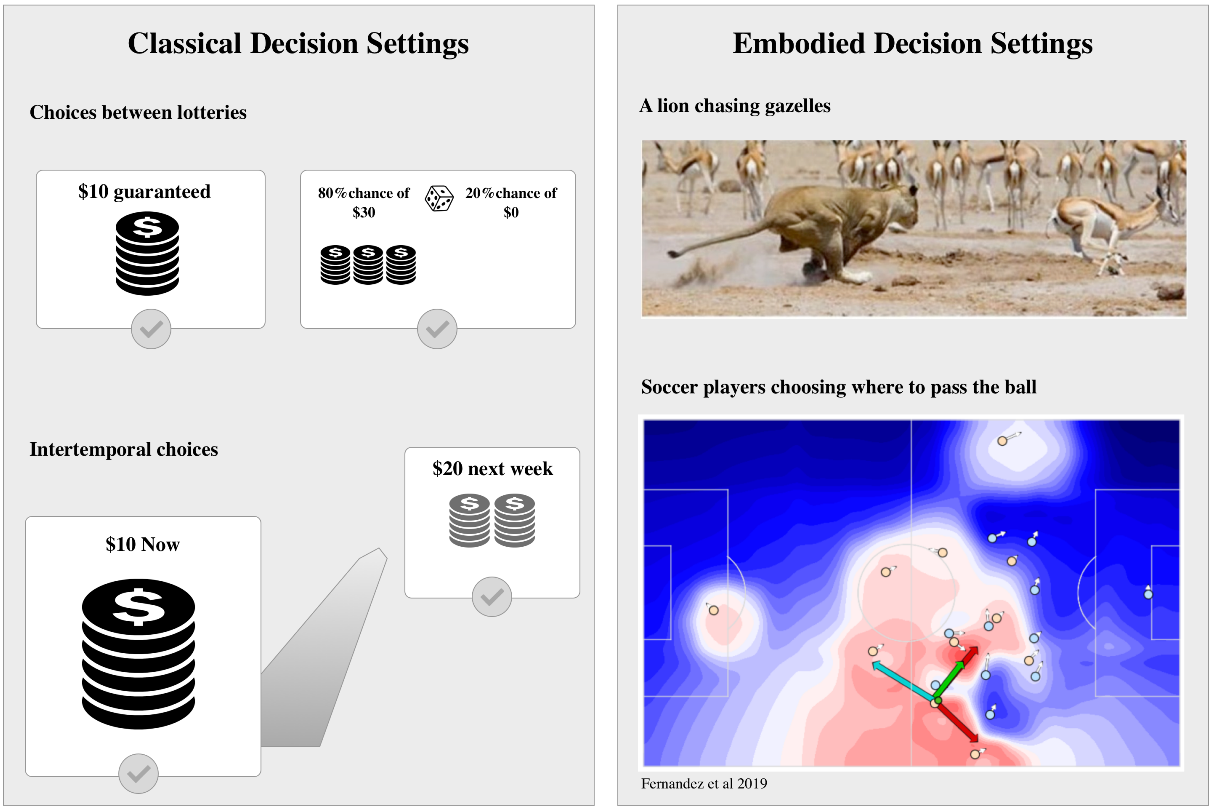


Figure 3.1: Examples of classical decision settings (left) versus embodied decision settings (right). Bottom right figure shows expected pass value (EPV) surface in soccer, reproduced from (Fernández et al., 2019). Warm (red) colors represent higher EPV whereas cool (blue) colors represent lower EPV. Yellow circles represent the location of the player holding the ball (green circle) and his teammates. Blue circles represent the locations of the opponents. The colored arrows identify potential passes having the highest EPV on the surface (green arrow), the highest utility (red arrow) or the lower turnover expected value (cyan arrow), which (roughly) corresponds to the lowest probability of a counterattack in this setup. This visualization illustrates that a continuous decision process can be described in terms of classical concepts and analyzed with established techniques. For example, this formalism permits assessing whether soccer players chose "risky" (i.e., higher utility, lower probability) or "safe" (i.e., higher probability, lower utility) passes.

to exploit an option that would otherwise disappear; for example, a lion can start running toward a group of gazelles before deciding which one to chase (i.e., act-while-deciding), otherwise the gazelles may simply run away. Furthermore, the decision-maker can change their mind while acting, due to reconsidering a previous decision (Resulaj, Kiani, Wolpert, & Shadlen, 2009), gathering novel evidence, or the appearance of novel opportunities.

These examples illustrate that there are crucial differences between the classical laboratory settings in which decisions are usually studied and the embodied settings in which real life decisions are deployed. While it is intuitively clear that studying decisions in restricted laboratory settings is methodologically simpler than studying them "in the wild", it worth asking whether certain aspects of lab settings, such as the focus on stable, predefined alternatives and the trivialization of action, have contributed to a distorted (or at least incomplete) view of how we solve decision tasks. For example, one may ask to what extent the serial view of decision-making (decide-then-act) applies beyond simple binary decisions in the lab, or whether capturing the richness of real life decisions (act while deciding and decide while acting) requires a different class of (embodied) models, which acknowledge the fact that decision and action dynamics deploy in parallel and influence each other bidirectionally (Lepora & Pezzulo, 2015).

For example, a classic model of decision-making posits that one deliberates by accumulating sensory evidence until it reaches a threshold, and initiates movements at that time (Gold & Shadlen, 2007; Ratcliff, 1978). But how does that model generalize to situations, common during real-time behavior, in which one must already be acting (and thus "above threshold") while still deliberating? Must the very concept of a threshold be abandoned when considering embodied settings, and if so, then where does that leave current models? Recent studies are starting to examine how humans make decisions during ongoing actions (Grießbach, Incagli, Herbort, & Cañal-Bruland, 2021; Michalski, Green, & Cisek, 2020) but models of the underlying mechanisms will need to go beyond traditional ideas of initiation thresholds, possibly toward ideas of task-specific subspaces in high-dimensional neural populations (M. T. Kaufman, Churchland, Ryu, & Shenoy, 2014).

I argue that to broaden our understanding of decision-making, it is necessary to go beyond restricted laboratory settings and design novel experimental paradigms bringing embodied dynamics tasks under controlled conditions. To this end, it is necessary to remove some conceptual and methodological barriers that make embodied decisions challenging to address.

3.2 Towards a deeper understanding of embodied decisions

This chapter has three goals. The first goal is to clarify the key characteristics of embodied decisions and the novel questions they raise. The second goal is to identify a novel methodology for the study of embodied decisions, by distilling key insights from recent studies in sports analytics, experimental psychology and other fields. The third goal is to discuss to what extent the study of embodied decisions requires novel theories, or will change our understanding of what decisions are.

Specifically, the remainder of this chapter addresses the following three points:

(1) Novel questions. What are the key differences between classical and embodied decision settings? What are the novel experimental questions that only arise when studying embodied settings and are instead ignored in classical settings? What can we learn from embodied settings that we cannot learn in the "classical" way?

(2) Novel methodologies. How can we experimentally address the above empirical questions? Given that studying embodied decisions poses additional challenges compared to restricted laboratory settings, is it possible to develop a methodology that does not sacrifice rigor? Can we identify success cases? Can we borrow methodologies from other fields that address similar problems?

(3) Novel theories. Would a widespread use of embodied settings change the way we understand decision-making? Is it possible that by studying decision-making in the classical way we have mischaracterized its mechanisms? Will the study of embodied decisions require novel conceptual frameworks and how would they differ from standard decision models? What will be the impact of novel studies of embodied decision making for cognitive science, neuroscience, robotics and other fields?

3.2.1 What novel questions can we ask through the study of embodied decisions?

In the introduction, I argued that embodied decision settings include a number of dimensions that are missing from classical settings. Here, I reconsider these unique dimensions of embodied settings and highlight that they prompt novel research questions that are impossible or difficult to study in classical settings. Below I focus on three novel questions that, though not exhaustive, exemplify the heuristic potential of embodied settings. To illustrate the novel questions, I will use the example of a soccer player who decides where (or to which teammate) to pass the ball.

Question 1. How are offers and their attributes identified? In classical settings, the offers are clearly identifiable by the decision-maker and their number is fixed and known in advance. The situation is different for the soccer player, because (despite an awareness that she has 10 teammates) she might not know exactly where they are and would hardly deem all of them good choices. This situation prompts the novel question of how many (and which) choice alternatives she considers in the first place.

A closely related question is what are the dimensions that the soccer player considers during the decision. The choice can clearly benefit from considering multiple possible dimensions, such as the distance from the player, whether the positioning of the receiving player is appropriate or advantageous to score a goal, etc. I will argue in the next section that it is advantageous to group these dimensions into two classes—probabilities and utilities—which permits us to establish a mapping with the two typical dimensions of expected value in economics.

Crucially, however, the soccer player is not provided a priori with a list of the dimensions to consider and of the values of these dimensions, but has to figure them out as part of the decision-making task. This prompts a number of additional questions: which dimensions are considered? To what extent are these dimensions context-sensitive and time-varying? What are the contextual factors (e.g., situated aspects of the choice) and individual differences (e.g. personality or cognitive factors) that influence the selection of relevant dimensions?

Another set of questions regards the way the values of these dimensions are estimated under uncertainty and how this influences dimension selection (e.g., it would be ineffective to select a choice dimension for which one has no access to reliable values). In classical studies of value-based decision-making, the perceptual phase is often trivial. Under the rubric of perceptual decision-making or attention selection, perceptual dynamics are studied in isolation from value-based computations. In contrast, in the domain of embodied decisions, the different (e.g., perceptual, attentional, value-based) aspects of the decision process are tightly intertwined. While studying each in isolation may be fruitful as a first approximation, embodied settings may offer a unique window into how they interact.

Finally, the classical setup suggests that the identification of offers and evaluation of their attributes are largely sequential processes. Conversely, embodied decision settings allow us to study their interactions. For example, as you identify the offers, you may discover the attributes that differentiate them, which then helps to eliminate some offers from consideration (e.g. some players may be just too far to be considered, or too close to an opponent), leaving a new set of offers with a new set of attributes to consider (e.g. some players may be moving too fast). These examples suggest that the consideration of some attributes comes before and influences the identification of offers, which might compose the first stage of a decision process that filters out irrelevant choice alternatives. These are all problems that hardly arise in classical settings.

Question 2. How is the deliberation between choice offers performed? Classical settings start from the premise that choice offers are predefined and presented in parallel, and the offer-attribute mapping is fixed throughout the decision. These assumptions are directly incorporated into decision-making models, like drift-diffusion models (Ratcliff, Smith, Brown, & McKoon, 2016), accumulator models (Usher & McClelland, 2001), or attractor-based models (Wang, 2008). For example, the drift diffusion model requires fixing two thresholds (one for each offer) and the decision variable (that reflects the value of the first offer minus the value of the second offer) for the duration of computation. The attractor-based model requires fixing the synaptic connections between populations that encode attribute values and populations that encode offer values. Without these fixed mappings, the mathematical guarantees of these models break down. However, as has been discussed above, in embodied decisions neither offers nor their mapping with attributes are predefined or fixed, nor can we assume they are presented simultaneously. It is unclear whether the most popular decision-making models can deal with these complexities in embodied situations, both mathematically and conceptually (e.g., embodied choices would require a rapid reconfiguration of the neural architecture supporting attractor-based decisions, but it is unclear to what extent this is plausible).

One may therefore argue that we need a different conceptualization of the deliberative process that leads to embodied decisions. At minimum, we should consider that the mathematical guarantees of models that deal with fixed data streams are not appropriate for embodied decisions, where the choice conditions change over time and hence the most recent data streams are more relevant (Cisek, Puskas, & El-Murr, 2009). A more drastic change of perspective comes from studies of foraging (Charnov, 1976) (which have strong decide-while-act components), where the standard assumption is that the deliberation is between "select the current offer" (exploit) vs. "search for other offers" (explore), not "decide between two fixed choices" as in classical decision settings. A series of studies suggest that birds (and other animals) use by default strategies for dealing with sequential choices and these lead them to make irrational decisions when faced with simultaneous choices (Kacelnik, Vasconcelos, Monteiro, & Aw, 2011). However, it remains an open question whether this or other perspectives offer a broader and more appropriate conceptualization of situated decision-making in real life situations.

Question 3. How do perception, decision and action processes influence each other? In embodied settings, perception, decision and action processes are deeply intertwined. Leaving aside perceptual processes (which I briefly discussed above), the interplay of decision and action—and the "continuity" of the decision (Yoo et al., 2021)—is evidenced by the fact that decision-makers can start acting before completing a decision (act-while-deciding) and can change their mind along the way, perhaps to consider new offers that were not initially present (decide-while-acting). These situations create novel opportunities that are rarely (if ever) addressed in classical settings.

As an example of a novel question, it is unclear at what point of the decision a commitment emerges for one of the alternatives and under which conditions a person can change their mind (Cos et al., 2021). Additionally, what is the range of novel strategies that embodied settings afford? For example, a common observation in studies that ask people to answer by clicking one of two response buttons with a computer mouse is that participants sometimes move rapidly in between the two buttons, perhaps as a strategy to "buy time" and avoid committing until necessary (Pezzulo & Ognibene, 2011). These strategies are not normally available in classical settings but are very important in real life, where decision commitment dynamics are less constrained and postponing is often an option.

Finally, while decisions are normally treated as independent from the context and from one another, sequential dynamics are ever-present in embodied settings. As each decision and action influences subsequent decisions, actions and perception dynamics, the "dynamical affordance landscape" of decisions (Pezzulo & Cisek, 2016) cannot be easily disregarded. For example, if the soccer player has the ball and goes to the left, he will open some affordances (a left attack) but also close others (an attack to the right).

3.2.2 What novel methodologies do we need to address these new questions?

Embodied decision settings prompt several novel research questions. Some studies are already meeting these challenges. For example, studies that track continuous (eye, hand or mouse) movement dynamics during the choice reveal how decision and action processes influence each other; for example, how decision uncertainty provokes adjustments of movement trajectories during a continuous decision, and how motor costs influence the initial decision and the subsequent changes of mind (Barca & Pezzulo, 2015; Burk, Ingram, Franklin, Shadlen, & Wolpert, 2014; Cos et al., 2021; Marcos, Cos, Girard, & Verschure, 2015; Resulaj et al., 2009; M. Spivey, 2007; M. J. Spivey & Dale, 2006; M. Spivey, Grosjean, & Knoblich, 2005). Other studies address how we make decisions while we are already engaged in performing an action, and reveal the importance of geometric and situated aspects, such as how the relative distances between currently pursued and other available targets, influence the choice of changing target (Michalski et al., 2020).

Studies of embodied decision-making sometimes target interactive situations too, such as joint decisions (Pezzulo, Roche, & Saint-Bauzel, 2021) and joint actions, either collaborative, as in grasping something together, or competitive, as in martial arts (Candidi, Curioni, Donnarumma, Sacheli, & Pezzulo, 2015; Pezzulo, Donnarumma, & Dindo, 2013; Pezzulo & Dindo, 2011; Sebanz, Bekkering, & Knoblich, 2006; Yamamoto et al., 2013). These studies reveal how the decision and action processes of two (or more) co-actors are interdependent and how they can become aligned over time, especially when the two can engage in a reciprocal sensorimotor (nonverbal) communication (Pezzulo et al., 2018) (the task explored in Chapter 6 addresses this kind of situation).

Another example is a promising trend in systems neuroscience toward studies of more naturalistic settings. This includes neural recordings in animals that are freely moving (Chestek et al., 2009; Schwarz et al., 2014; Sodagar, Wise, & Najafi, 2007) or navigating through controlled environments (Etienne et al., 2014; Krumin, Lee, Harris, & Carandini, 2018), as well as studies in humans using portable magnetoencephalography (MEG) or electroencephalography (EEG) equipment (Boto et al., 2018; Djebbara, Fich, & Gramann, 2021; Djebbara, Fich, Petrini, & Gramann, 2019; Topalovic et al., 2020).

3.2.3 Towards a methodology that connects classical and embodied domains of decision-making

Despite these promising developments, most aspects of embodied decisions remain unaddressed—in part because embodied decisions are more challenging to study compared to classical settings. A key question is whether it is possible to develop a sound methodology for the study of embodied decisions in their full complexity, without sacrificing rigor. One aspect that makes the study of embodied decisions challenging is that most of the choice dimensions reflect spatial and geometric aspects of the situation, such as "passability" affordances for a soccer player. Unlike choice dimensions usually considered in classical settings, such as dollars and seconds, affordances are more difficult to formalize, but they are exactly the dimensions that our situated brains evolved for.

The methodology proposed here involves mapping the choice dimensions of embodied problems into the two factors—probabilities and utilities—that are often used in classical problems to define expected value. Establishing a formal correspondence between classical and embodied settings will permit us to address classical questions about (for example) utility maximization, risk-avoidance and sunk costs in embodied settings.

Example 1: The case of soccer

Fortunately, there are some success cases that exemplify this methodology, especially in sports analytics, as in the case of statistical studies of basketball (Cervone, D'Amour, Bornn,

& Goldsberry, 2016) and soccer (Fernández et al., 2019). For example, the bottom right image in Figure 3.1 shows an example of a soccer player (yellow circle) who has to decide to whom he wants to pass the ball (small green circle)—or more precisely, where to pass it, since the passage can be to any location of the field, not just to his teammates' current positions. The figure shows the "expected pass value (EPV)" surface for the soccer player (zones in red and blue have high and low EPVs, respectively). The notion of EPV is exactly the same notion of expected value used in economic studies and it is calculated by combining the usual two factors of probabilities and utilities. In this setting, the probability that a pass will succeed is itself modeled as a function of several subfactors, such as the distance from teammates and whether the passage zone is "under control" by a teammate or opponent. The utility of a pass is modeled as spanning both positive and negative dimensions, in consideration of the fact that passing to a certain zone can change the chances of scoring a goal (positive dimension) or receiving a counterattack (negative dimension). The positive and negative dimensions of utility are thus calculated separately by considering several subfactors, such as where the target of the pass is located within a precomputed "zone value surface" that spans the whole playing field (and is higher in the opponent's field), and how many times a successful pass to the target zone resulted in a goal or a counterattack in the model's training data. The values of each of the subfactors that contribute to probabilities or utilities are in part derived analytically (e.g., by calculating the zones that each player "controls") and in part learned from data (e.g., how many times a goal was scored from each zone of the playing field, in a large database of recent matches).

While the ways probabilities and utilities are calculated is context dependent (e.g., depends on specific game models that differ for soccer, basketball and other games), the methodology is general and it permits establishing a strong formal connection between embodied and classical settings. In other words, regardless of the difficulties of establishing how probabilities and utilities should be modeled for each embodied setting, and how to estimate their values, once these two factors have been estimated they can be combined to form an expected value surface, in the same way they are combined in classical settings and economics. This is evident if one considers that in Figure 3.1, it is possible to identify the choices (passes) that provide the highest rewards, defined here as moving the ball into an area from which it is possible to shoot at the goal (red arrows), those that provide the best expected value (green arrow) and those that provide the lowest risk of a counterattack (cyan arrow). This makes it possible to ask (for example) whether the soccer player selected the pass that maximizes utility or whether he is risk-seeking or risk-averse.

Interestingly, despite the EPV being continuous, in most cases it is possible to enumerate a discrete (and small) number of choice offers, by focusing on the "peaks" of EPV on the field. This exemplifies a possible approach to map continuous aspects of embodied decisions (e.g., physical space) to discrete choice dimensions, which are most commonly used in standard decision-making models. However, while standard decision-making models specify a priori a discrete set of options, it is possible that these emerge from continuous surfaces. This phenomenon can be formalized using various computational models that show decision surfaces developing peaks and troughs (Amari, 1977; Cisek, 2006; Furman & Wang, 2008; Sandamirskaya, Zibner, Schneegans, & Schöner, 2013; Schneegans & Schöner, 2008) and transitioning from representing multiple options to choosing a single winner (Grossberg, 1973; Standage, You, Wang, & Dorris, 2011).

Example 2: The case of crossing the river

I illustrate how to use the same methodology used in the soccer study to derive "expected value surfaces" to deal with other embodied decision settings; for example, the case of a person who has to cross a river by jumping between stones having different sizes and placed at different distances, see Figure 3.2.

As in the soccer example, this situation can be addressed by deriving various "surfaces" that consider spatial and geometric elements of the task, such as child-stone distance and stone sizes (as examples of probability surfaces) and distance between stones and the end of the river (as an example of a utility surface).

Before analyzing the task of crossing the river, it is useful to explicitly mention the guiding principle used to distinguish probability and utility factors. In this analysis, the former (probability) factor relates to "present affordances" and the latter (utility) factor relates to "future affordances", respectively (J. J. Gibson, 1977; Pezzulo & Cisek, 2016). Probabilities (and present affordances) only depend on the current situation (e.g., current position of the jumper, distance from the stones and their size), not the intended destination. As a result, computing the probability (or "present affordance") surface amounts to asking: "which stones are most 'jumpable' given my current situation"? This question can be asked irrespective of the intended destination. On the contrary, the notion of utility depends on the future, up until the goal destination (in my example, the end of the river). Hence, computing the utility (or "future affordance") surface amounts to asking: "Will this jump increase the chances of reaching the goal destination (and by how much)? What affordances will I create (or destroy) by completing a certain jump?"

A concrete example may help. Let's consider the child preparing to jump shown in Figure 3.2. He has the choice between three options: the small stone to the right, the small stone to the left and the big stone to the center. The right stone is closer and hence has the greatest present affordance, but it has lesser utility, as it avails less progress towards the goal. The center stone has the greatest utility as it is closer to the destination, but (despite its size) it may have a lower present affordance, given its distance from the jumper. The jumper actually selects the left stone, which seems a good compromise despite having neither the greatest present affordance nor the greatest utility. Yet, its utility is noteworthy as it "opens up" novel affordances: it permits reaching the next, bigger stone with little effort, and eventually progressing until the goal.

Figure 3.2b-f shows a schematic view of an affordance landscape summarizing the child's decision (Figure 3.2a), with black circles representing the stones, and the white triangle representing the child's present location and orientation. In this example, I formalize the "probability" (or "immediate affordance") component in terms of two subcomponents. The first subcomponent of probability, A_d, takes into consideration the child-stone distance and prioritizes shorter jumps. It is calculated as:

$$A_d(x, y) = -\sqrt{((x - x_{child})^2 + (y - y_{child})^2)} \tag{3.1}$$

Note that the units in which x and y can be expressed relative to the participant's size, to account for the fact that distance depends on bodily parameters (Warren, 1984).

The second subcomponent of probability takes into consideration "landability" and assigns positive values to locations with stones and zero value to locations without stones (as

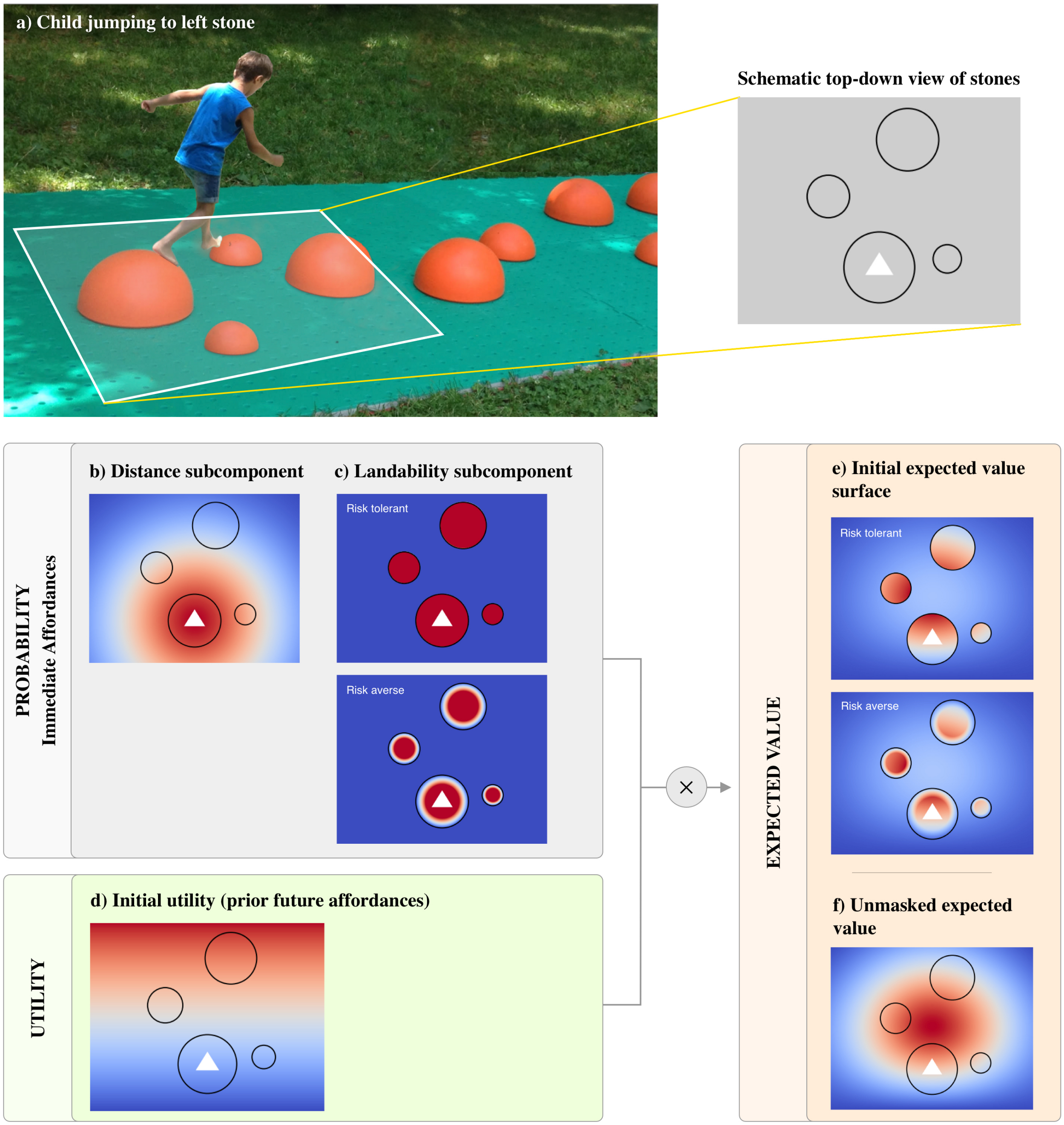


Figure 3.2: "Cross the river" task setup. This figure illustrates the two components of the affordance landscape for the situation shown in the photo in a), where a child must decide between three candidate stones for their next jump. (b-f) show a schematic view of an affordance landscape summarizing the child's decision, with black circles representing the stones, and the white triangle at the child's present location. The figure shows both probability surfaces and utility surfaces and how they are integrated to form an expected value surface for jumping. See the main text for explanation.

it is impossible to land on water). For illustrative purposes, two landability affordances are shown, which cover the whole stone or only its center, respectively (see Figure 3.2c). The former may be perceived by very accurate (or risk-seeking) jumpers, whereas the latter may be perceived by less accurate (or risk-averse) jumper, who may not perceive the outer edge of the stone being landable, knowing that their jumps have some variance and hence aiming at the outer edge poses a risk of falling (Trommershäuser, Maloney, & Landy, 2008).

Furthermore, the “utility” component is formalized as the forward progress towards the end of the river (see Figure 3.2d). This is approximated as:

$$A_{u,0}(x, y) = y - y_{child} \tag{3.2}$$

This naïve method does not correspond to the optimal utility surface (see below) but it is sufficient in this simple example to provide some directionality.

By combining the two (probability and utility) components it is possible to calculate an expected value (EV) surface, see Figure 3.2e. This can be calculated simply as the product of the scaled probability and utility surfaces discussed above:

$$EV_0 = (A_d \cdot A_l) \cdot A_{u,0} \tag{3.3}$$

Min-max scaling is used, where $X_{scaled} = (X - X_{min})/(X_{max} - X_{min})$.

Note that the EV surface is shown in two versions, each integrating one of the two landability surfaces shown in Figure 3.2c. It can be informative, as well, to visualize the EV surface with subsets of its subcomponents. For example, Figure 3.2f shows only the distance subcomponent of probability and utility (Figures 3.2b & 3.2d), which expresses the preferred location of stones irrespective of their true position, which may provide a useful guide during visual search—or if the task is to place stones optimally.

While useful, the EV surface is myopic as it only considers the immediate utility of each jump derived from a naïve notion of “direction to goal”, rather than long-term utility. Given that this is a sequential decision problem that involves multiple jumps, long-term utilities should consider not just the utility of the next jump but also of the successive ones that may become available or unavailable as a function of the next jump; for example, using a look-ahead planning algorithm such as REINFORCE (R. Sutton & Barto, 1998a).

Look-ahead planning permits addressing cases in which the myopic strategy is maladaptive and leads to dead ends. One such case is illustrated in Figure 3.3. This is a variant of the situation shown in Figure 3.2, in which the final stone has been moved from the left to the right. A myopic strategy would still assign the highest EV to the left stone, even if this is a dead end (Figure 3.3a). However, using a one-step look-ahead planning algorithm such as REINFORCE, will permit predicting how the value surface will change after jumping to the left or right, respectively. In turn, considering this future utility instead of the immediate utility would change the EV surface at the start position – and assign the highest utility to the right stone (Figure 3.3b). This simple example illustrates the benefits of planning in situations where naïve notions of utility (such as direction to goal) are misleading.

From an empirical perspective, people may be quite effective in avoiding dead ends while crossing the river (in person or in a videogame-like setting), but at the expense of investing some (cognitive) cost – which may reflect the process of planning ahead, while overriding current affordances. Interestingly, such cognitive costs can be modeled as the computational costs required to update the expected value surface, from the "immediate" utility of Figure 3.3a to the "future" utility of Figure 3.3b-c. By using REINFORCE or a similar Monte Carlo algorithm, it is possible to generate predictions for this cost as the number of rollout steps required to disambiguate between competing immediate affordances.

In general, the greater the differences between the "immediate" utility surface of Figure 3.3a and the "future" utility surface of Figure 3.3b, the greater the computational cost to compute the latter (Ortega & Braun, 2013; Todorov, 2009; Zenon, Solopchuk, & Pezzulo, 2019). If these computational costs correspond to mental effort, they could plausibly become apparent when comparing the RT distributions of peoples' choices before a jump, when the "immediate" and "future" utility surfaces are more or less similar.

Note that calculating the most accurate expected value surface may have a high computational cost. If people optimize a trade-off between the accuracy of the expected value surface and the computational costs of obtaining it, they could select adaptively the amount of resources to invest (e.g., number of rollouts; compare the two updated expected utility surfaces of Figure 3.3b-c which are obtained using 2 rollouts (Figure 3.3b) or 10 rollouts (Figure 3.3c) of the REINFORCE algorithm). In tasks that involve multiple choices, people might also select the strategy to plan ahead (and invest cognitive resources) only during the most critical decisions; for example, when they need to choose between two sequences of stones in opposite directions, but not in simpler situations, such as the one shown in Figure 3.2, where the benefits of planning might not be worth its cost. By looking at subjects' choices during the task, it would be possible to infer the expected value surface that they actually computed, which could provide an indication their (cognitive) effort investment – and whether and how often they look ahead during embodied choices.

Example 3: The case of climbing

I now consider one final application of "expected value surfaces": the case of a rock climber who has to plan the best way to climb a wall – and more specifically, the best sequence of holds to traverse with their hands and feet, until reaching the "top" hold that marks the end of a climbing problem. Figure 3.4a shows a climber solving an example problem on a bouldering wall (called MoonBoard). The MoonBoard wall implies a standard configuration of all possible holds, and a problem is defined as a subset of these holds that are allowed. The goal is to start with each hand touching two predefined holds, then executing a sequence of movements until reaching a predefined "top" hold with both hands. Figure 3.4a shows a climber on an example bouldering problem on the MoonBoard, whereas Figures 5b-g show various probability and utility surfaces for the same problem (the holds that belong to the problem are marked with circles Figures 5b-g; and the current positions of the climber's hands and foot are marked with short arcs around the holds; the green arc around marks the position of the climber's left hand). As in the case of crossing the river, the value surfaces combine the two factors of "probabilities" (or current affordance, e.g., which holds are reachable given the current situation?) and "utilities" (or future affordance, e.g., which holds are more useful to open up future affordances that permit progress towards the top of the wall?).

However, there is an important difference in the way "probabilities" were calculated in the two cases of crossing the river and climbing. In the case of crossing the river, they were

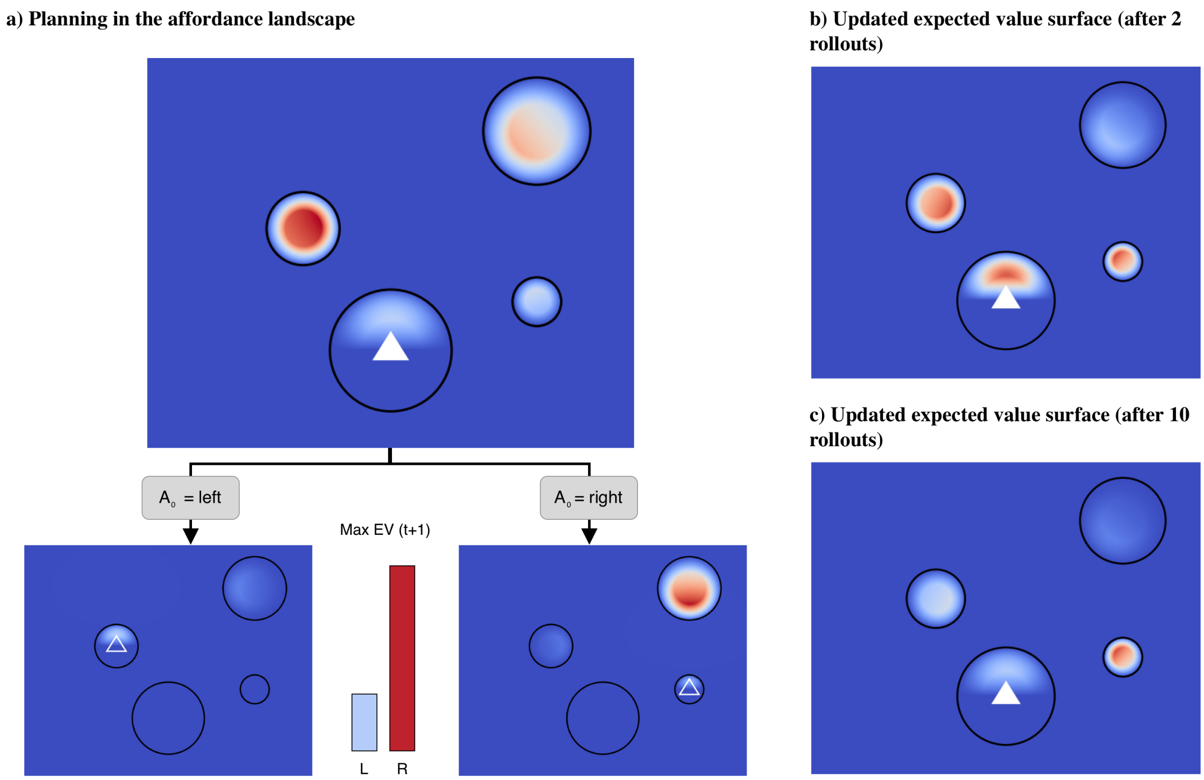


Figure 3.3: An example of planning while crossing the river. (a) Expected value surface without planning, and below, the expected result of each available action (left, right). (b-c) Expected value surfaces updated based on 1-step planning. The cost of computing these updated expected value surface is a function of various parameters specific to both the planner (e.g. learning rate, planning depth) and problem context (e.g. number and differentiation of offers available, branching factor, etc). Here, two updated utility surfaces are shown, calculated using 2 rollouts (b) or 10 rollouts (c) of the REINFORCE algorithm with the same parameterization (R. Sutton & Barto, 1998a). Please note that these are just examples of updated utility surfaces, as the updating depends on the specific parameterization of the algorithm (e.g., its learning rate).

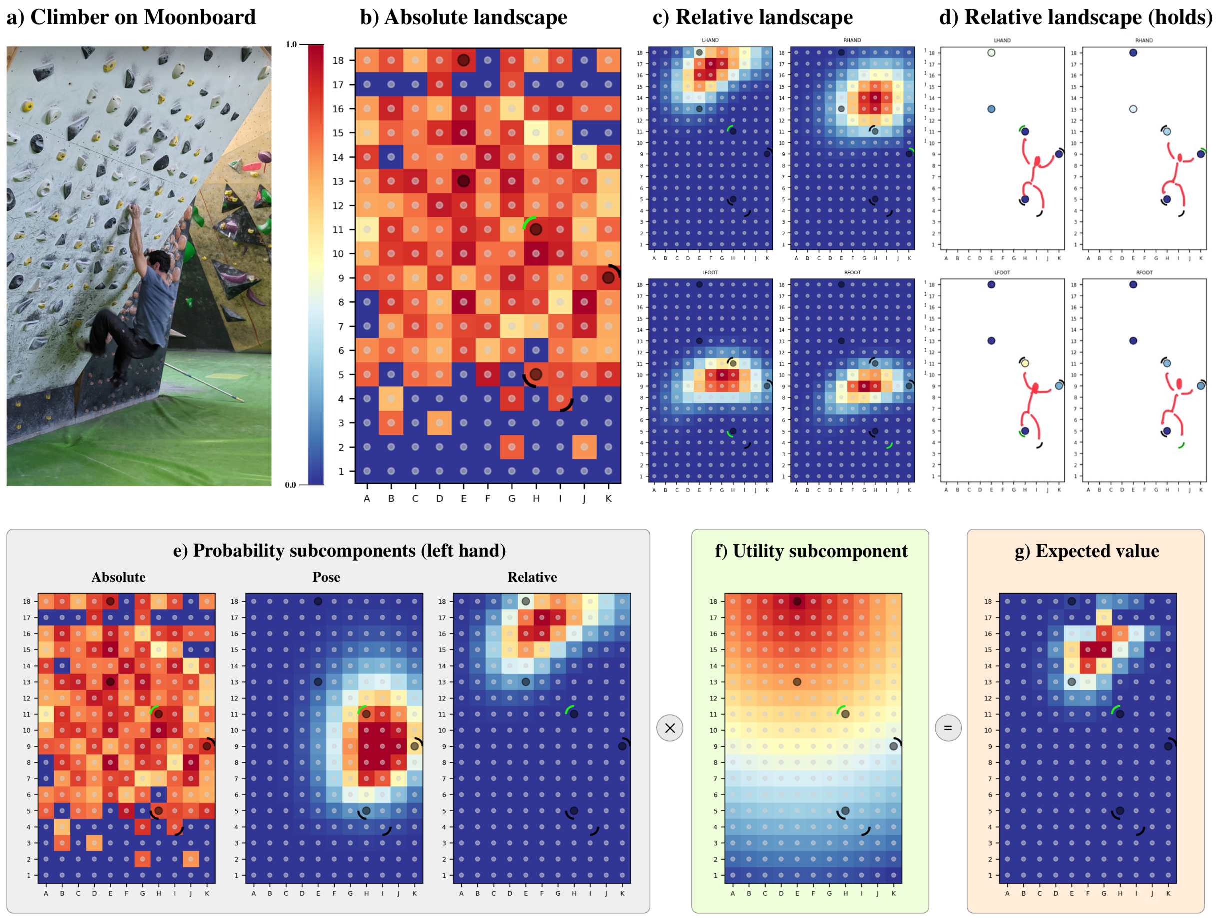


Figure 3.4: The present affordance landscape of a bouldering wall (the MoonBoard). (a) A climber attempts an example problem on the MoonBoard. (b-d) Example subcomponents of the probability surface: absolute landscape, relative landscape (unfiltered) and relative landscape (filtered to show only the available holds for this problem, with a sketch of a climber superimposed to show the posture). The black circles show the positions of the climbing holds in the example problem. The curved lines show the current positions of the hands and feet of the climber in Figure a, with the green curved line highlighting the limb whose landscape is shown. (e) The three subcomponents that are combined to form the probability surface of the climber's left hand, which is the next to move; these are the absolute landscape (which is the same as b), the pose subcomponent, and the relative landscape of the left hand (which is the same as the top-left panel of figure c). (f) This is a naïve utility component: a distance gradient starting from the final hold of this MoonBoard problem. (g) The final Expected Value for the left hand, which is obtained by combining probability subcomponents of Figure e and the utility subcomponent of Figure f. See the main text for explanation.

calculated analytically, by only considering geometric aspects of the situation: agent-stone distance (for the distance subcomponent) and stone sizes (for the landability subcomponent). However, this analytical approach is much more challenging in the case of climbing, as a climber's current affordances plausibly depend on a much larger set of factors (e.g., climbing hold types, posture, distance) that are challenging to model. To generate the probability surfaces in this case, a small set of subcomponents were first identified, after which a data-driven approach was used to learn their associated probabilities from a large dataset of climbing data (similar to the approach used in the soccer example described above (Fernández et al., 2019)). One implication of this method is that the probability surfaces considered here reflect the average climbing patterns of multiple climbers.

Figure 3.4 show the subcomponents of the probability surface considered in this example. Figure 3.4b shows the first subcomponent, called "absolute landscape", which depends only on the specific climbing hold (i.e., its type, shape, size) in any given position, not on the climber's current position, and only depends on the kind and shape of the climbing hold. Here, the holds in the positions marked in red are "good" (or "easy to grasp and hold") while those in blue are "bad" (or "difficult to grasp and hold"). The absolute landscape subcomponent is calculated empirically, using data freely available from the MoonBoard app (`https://www.moonboard.com/moonboard-app`), under the assumption that "good" (bad) climbing holds occur more often in easy (difficult) climbing routes.

Figure 3.4c shows a second subcomponent of probability: the relative landscape. This is the likelihood of moving each limb to a particular region of the board and hence capturing aspects of limb-hold distance. Note that this subcomponent is specific for the current position of the climber; in the figure, the positions of the hands and foot are indicated by the small curved lines (with green indicating the limb whose subcomponent is shown, i.e., the left hand). Furthermore, the relative landscape subcomponent is limb-specific, which means that there is a separate subcomponent for each limb: the two top panels of Figure 3.4c are for the two arms, while the two bottom panels are for the two feet. The relative landscapes are calculated empirically from sequences chosen by climbers (not based on geometric considerations as in the case of crossing the river), conditioned on the starting pose of each limb. Figure 3.4d shows the same landscape, but "filtered" to show only holds that exist in the given problem.

Figure 3.4e shows the aggregate probability surface for the climber's left hand. This is the combination of the two probability subcomponents introduced earlier (the absolute landscape of Figure 3.4b and the relative landscape of Figure 3.4c, here limited to the left arm, i.e., the top-left panel) plus a third subcomponent: the pose. The pose subcomponent favors more common spatial configurations of the hands and feet, and discourages those that are less common. For example, low probability is assigned to moves resulting in configurations that break simple biomechanical constraints such as hands positioned farther apart than a climber's arm-span would allow (e.g. in the example above, the pose constraint limits movement of the left hand away from the other limbs, which can be seen in the second plot in 3.4e). Like the relative subcomponent, the pose subcomponent is computed empirically by fitting a set of bivariate Gaussian distributions over the distributions of the Euclidean distance of limb's pairs as observed in the solutions dataset.

Figure 3.4f shows a naïve utility surface for this climbing problem. This is a simple gradient that originates from the last climbing hold of the problem, similar to the utility surface for the case of cross the river. Figure 3.4g shows that the probability surfaces (Figure 3.4e) and the naïve utility surface (Figure 3.4f) can be combined, to derive a final expected value (EV) surface for the left hand of the climber (Figure 3.4g). Obviously, the same can be done for the other three limbs.

As in the cases of soccer and crossing the river, deriving expected value surfaces permits modeling the embodied choices of a climber who faces the same problem as shown in Figure 3.4a. Figure 3.5 illustrates a simulated climber that uses the expected value surfaces described so far to solve the problem of Figure 3.4a. Note that at each step during the simulated climb, the model considers *four* value surfaces, one for the movements of each arm and foot—which means that at each moment in time, the climber has a choice not just between where to move a limb but also which limb to move. Figure 3.5a-b illustrates the resulting climbing behavior of a "greedy" planner that always selects the climbing hold whose expected value is the highest across the four surfaces (the alternation of hands and feet is an emergent effect of this strategy). This figure shows that some climbing holds (those at the top of the wall) have very low expected value at the beginning but acquire value as the climber progresses to the top of the wall. This exemplifies the concept of an evolving "landscape" of affordances as a function of the changing position and pose of the climber; and the fact that by moving the climber creates some affordances (and destroys others) (Pezzulo & Cisek, 2016).

Figure 3.5c-d shows an empirical solution of the same problem, i.e., the actual sequence of movements selected by an experienced climber, with an expected utility landscape (generated from the same model) superimposed above. In this particular example, some of the expert's movements were well-predicted by the greedy planning strategy, but not all. This is not surprising, given that the climbing scenario poses significant modeling challenges. Furthermore, it is worth reminding that the simulated climber of Figure 3.5a-b uses a naïve utility surface, without look-ahead, which can potentially lead to dead ends. A better utility surface can be calculated with model-based planning, e.g., by doing rollouts from the current position up until the goal to be reached (i.e., the top hold); see the previous discussion of planning while crossing the river.

Interestingly, while the four value surfaces have been considered as independent so far, they become interdependent during planning; for example, by moving the left arm I can change the value surface of the other limbs. This is interesting from a planning perspective, because it makes it possible to model the fact that in climbing, foot movements are often instrumental to create future affordances for the hands (e.g., render a distal hold reachable). Given that the goal is reaching the top with the hands, feet have ancillary roles: they are fundamental to change the hand value surfaces.

As prior discussion exemplified, modeling climbing is significantly more challenging than modeling the case of crossing the river, as there are many more dimensions that contribute to the notion of "affordance" in this setup; and these are difficult to treat analytically. For this, a data-driven methodology was used, analogous to what done in other sports like basketball (Cervone et al., 2016) and soccer (Fernández et al., 2019). This methodology can be readily adapted to model other domains of embodied choice that resist an analytic treatment. However, it is worth noting that the data driven methodology used here blurs the distinction between probabilities and utilities that have been assumed so far, by endowing probability surfaces with some element of utility. This is evident from the fact that the

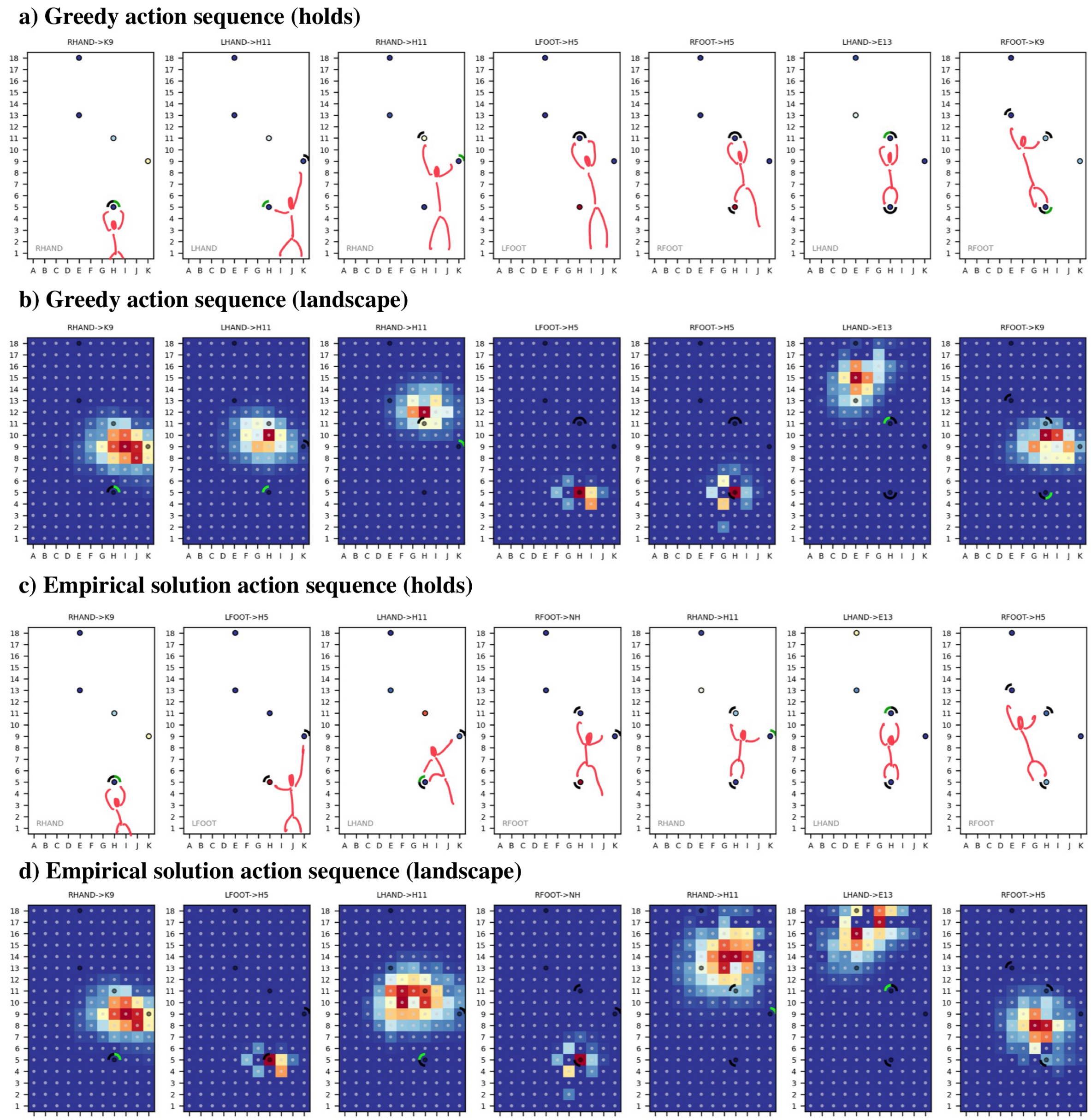


Figure 3.5: Partial action sequences for a selected problem showing the complete underlying affordance landscapes (a & c), and the value of each hold (b & d). The first two plots show a greedy policy in which the maximum hold value (across all four limbs) is selected and the corresponding expected value map shown. The second two plots show a partial sequence from the true solution as annotated by an expert climber.

maxima within the relative landscape are above the current positions of the limbs (this is different from crossing the river, where the distance subcomponent was independent of direction). The fact that the relative probability surface is imbued with some utility results from the fact that the relative landscape was computed from empirical data—from climbers whose movements all solve the goal-directed task of reaching the top of the wall.

3.3 Borrowing methodologies from other fields

Apart from sports analytics, many other disciplines have addressed embodied decision problems and developed methodologies to deal with them.

Robotics One field where embodied decisions have been an object of interest is robotics. For example, when a robot has to make a decision between which objects to grasp, it has to consider situated and embodied aspects of the task, such as biomechanical constraints and the presence of obstacles. Recent work illustrates that deriving good movement policies benefits from learning internal (latent) codes that represent body-scaled "distance" metrics. Crucially, latent distance is body-scaled and sensitive to the robot arm biomechanics and to the presence of obstacles, rather than reflecting Euclidean distance. This line of research shows that it is possible to derive body-scaled perceptions akin to affordance (reachability) surfaces that I considered above for the tasks of crossing the river or climbing, using methods from deep learning and robotics. See also (Roberts, Koditschek, & Miracchi, 2020; Zech et al., 2017) for recent reviews of other models of affordances in robotics.

Reinforcement Learning The Reinforcement Learning (RL) paradigm, which typically presupposes a situated agent interacting with a changing environment, is sufficiently general to take on some of the embodied dynamics of interest. Model-free techniques, such as Q-learning, center on the learning of a value function to score state-action candidates. This value function, however, is typically global and amounts to a reflexive stimulus-action mapping that is not easily extended to the kind of forward planning required by humans in complex environments. One attempt to bridge model-free and more flexible model-based approaches in RL, which exhibits some similarity with the relative landscape used in the climbing example, is the successor representation (SR) (Dayan, 1993; Gershman, 2018; Momennejad et al., 2017). The SR uses a simple store of discounted expected occupancy mapping each origin state to a distribution over destination states. This technique avails a temporally extended local representation from which expected value can be efficiently computed without the use of Monte Carlo sampling. The SR can be combined with online Dyna-like replay to accommodate changes in environment and reward structure that pose challenges for continual learning (Momennejad et al., 2017). Other work in RL has begun to integrate the concept of affordances more explicitly. In one recent effort affordances are defined as action intents in a Markov Decision Process setting in order to reduce the space of state-action pairs evaluated (Khetarpal, Ahmed, Comanici, Abel, & Precup, 2020), though empirical results are so far limited to simple grid-world settings. In another example, a neural network is trained to model contextual affordances in order to predict the consequences of an action using a particular object (identified a priori) and thereby avoid arrival in a failure condition from which the goal cannot be reached (Cruz, Parisi, & Wermter, 2018).

Ecological Psychology In ecological psychology, Gibson's work on affordances (Gibson, 1977) has been extended into an approach known as ecological dynamics, which applies methods from dynamical systems theory to understand behavioral interactions between organisms and their environment. Consistent with the conception of decision-making presented in this work, the ecological dynamics literature acknowledges that decisions are temporally extended and not easily separated from their behavioral expression (Beer, 2003; Nolfi, 2009; Nolfi & Floreano, 2001). Araújo et al. further suggest that because decisions are ultimately expressed as actions, an ecological analysis of human movement offers a "grounded" method to understand decision-making (Araújo, Davids, & Hristovski, 2006). In one demonstration of a such an analysis, a study of human locomotion and obstacle avoidance analyzed path data as participants chose either an "inside", or less direct "outside", path to a goal location (Warren & Fajen, 2004). The authors modeled behavior as a differential equation where acceleration was a function of both goal distance and the angle between the goal and an intervening obstacle. Their analysis finds "bifurcation points" in the space of initial conditions (e.g. when the obstacle-goal angle exceeds 4 degrees) from which participants consistently shift to the inside path. According to this model, the selection of a path can be seen as an emergent behavior that arises from the dynamics of steering interacting with a particular environmental structure (Araújo et al., 2006). Furthermore, paths could be explained by these on-line dynamics without reference to explicit planning within an internal world model. Another insight that comes from ecological theory of development is Eleanor Gibson's notion of prospectivity (E. J. Gibson, 1997). Prospectivity is one of four fundamental aspects of human behavior and it describes anticipatory, future-oriented behaviors such as an infant who reaches out for an object in anticipation that it will move within catching distance. This is an example of what we might call a future affordance, as the movement of the hand creates a new affordance (catching) that did not exist in the original configuration. This concept, that planning can be seen as the search for useful transformations to the available affordance landscape (Pezzulo, Barca, Bocconi, & Borghi, 2010; Pezzulo & Cisek, 2016; Rietveld & Kiverstein, 2014), is a central theme for the "affordance calculus" proposed.

Motor control While the two domains of decision-making and motor control have long remained separated, a more recent trend sees motor control as a form of decision-making — or the problem of maximizing the utility of movement outcomes under various sources of uncertainty (Shadmehr, Huang, & Ahmed, 2016; Wolpert & Landy, 2012). This analogy licenses the cross-fertilization of ideas and formal tools (e.g., statistical theory) across the two domains; and the development of novel computational models that combine objective functions used in economics, such as the maximization of expected utility, and in motor control, such as the minimization of motor costs (Ganesh & Burdet, 2013; Lepora & Pezzulo, 2015; Wispinski, Gallivan, & Chapman, 2018). These formulations become relevant when considering complex redundant tasks for which the same outcome can be achieved with different actions, as in a throwing task for which it is possible to hit a target (a fixed locations) with virtually infinite combinations of the release velocity of the projectile. This redundancy poses the issue of selecting one among a set of actions, or sequences of actions, that guarantee the same level of performance. Interesting insights on how humans select actions in redundant tasks have been provided by computational (R. G. Cohen & Sternad, 2009; Müller & Sternad, 2004) and analytical approaches (Cusumano & Dingwell, 2013; Scholz & Schöner, 1999; Tommasino, Maselli, Campolo, Lacquaniti, & d' Avella, 2021) based on the characterization of the action-to-performance mapping geometry. Such mappings can be seen as landscapes describing performance in a continuous space of actions, and the way in which executed actions are distributed in relation to the landscape geometry (e.g. gradient and Hessians) can reflect idiosyncratic differences in task execution strategies or learning patterns (Sternad, 2018; Tommasino et al., 2021). A similar approach could be beneficial for modeling embodied decision-making. For example, in the case of soccer, action selection could be further informed by considering the EPV (expected pass value) surface gradients, rather than just the peaks, so to take into account the risk associated with the intrinsic variability of motor execution (i.e. motor noise).

Also, when the overall goal requires the execution of subtasks, while for a given subtask the solution manifold may allow for distant actions, only a part of the manifold could be compatible with concurrent or subsequent tasks. In this case, planning could be formalized as reciprocal constraints on subtasks manifolds. It has been shown empirically that in sequential actions, each action can be executed differently as a function of the preceding and the subsequent actions in the sequence. This coarticulation effect has been observed across various domains, such as speech, fingerspelling, and reaching and grasping actions (Jerde, Soechting, & Flanders, 2003; Rosenbaum, Chapman, Weigelt, Weiss, & van der Wel, 2012). For example, people grasp (and place their fingers on) a bottle differently depending on the next intended movement (Sartori, Straulino, & Castiello, 2011). This and other similar results can be accounted for by a computational model where the solution manifolds of subsequent actions reciprocally constrain each other (Donnarumma, Dindo, & Pezzulo, 2017).

In general, the close integration between decision-making and motor control is supported by a large body of neurophysiological data showing the neural correlates of decisions throughout cortical and subcortical regions that are clearly implicated in sensorimotor control (for review, see (Cisek & Kalaska, 2010; Gold & Shadlen, 2007)). For example, decisions about where to move the eyes take place in oculomotor circuits (Basso & Wurtz, 1998; Ding & Gold, 2010; Platt & Glimcher, 1999; M. N. Shadlen & Newsome, 2001), while decisions about reaching or grasping occur in arm or hand areas (Baumann, Fluet, & Scherberger, 2009; Cisek & Kalaska, 2005). In particular, when decisions are made about different targets for reaching, neural activity in dorsal premotor cortex reflects the changing probability (Thura & Cisek, 2014) and relative utility of the choices (Pastor-Bernier & Cisek, 2011), the competition between targets is modulated by the geometry of their placement (Pastor-Bernier & Cisek, 2011), and the very same cells continue to reflect online changes-of-mind (Pastor-Bernier, Tremblay, & Cisek, 2012). In short, embodied decisions appear to unfold as a continuous competition between neural correlates of potential actions (affordances) biased by all factors relevant to the choice (Cisek, 2007; M. Shadlen, Kiani, Hanks, & Churchland, 2008) and modulated by an urgency signal that helps to control the trade-off between the speed and accuracy of decisions (Cisek et al., 2009); see also (Standage et al., 2011; Thura & Cisek, 2017).

3.4 Novel theories in the study of embodied decisions

In the introduction, I reviewed the fundamental differences between classical and embodied decision settings. There is a range of types of decisions—for example, choosing which class to take is not the same as choosing where to sit in the classroom—and these may require different mechanisms, serial or parallel, dynamic or static, etc. While the study of classical settings has advanced our knowledge in many ways, it is important to ask whether the knowledge gathered by studying classical choices, and the models developed to explain them, generalize to other situations, or whether they produced a distorted (or at least incomplete) view of how we solve decision tasks.

The literature on value-based decision-making is currently fragmented into two main classes of models. The first (classical) class of models directly stems from economic theory and target economic-like laboratory tasks – and indeed, the study of the neural underpinnings of decision-making is often called "neuroeconomics" (Glimcher & Rustichini, 2004). In this approach, decision-making is described as the selection between a menu of prespecified offers whose values are putatively coded by the orbitofrontal cortex (Padoa-Schioppa, 2011). In this class of model, decision is therefore a centralized (prefrontal) process that only assigns action and perception systems ancillary roles.

The second (action-based) class of models instead assumes that decision-making is a more distributed process within which the motor system plays an important role. When choices map directly to specific actions, as is the case in many of the choices for which our brains plausibly evolved, the selection can directly involve motor or premotor cortices — hence consisting in a competition between action affordances (Cisek, 2007). The action-based view does not assume that the prefrontal structures highlighted by the "classical" view are irrelevant; but rather that the decision emerges from a "distributed consensus" between several brain areas, each contributing different inputs to the decision (e.g., subjective values versus motor costs) (Cisek, 2012). Indeed, mapping at least some aspects of the choice into the action system is more advantageous than a centralized model for animals that have to decide and act in real time. Since it is implausible that humans have developed a completely different decision architecture to deal with economic decision in the lab, it is possible that the ancestral action-based architecture is also in play when choice-action mappings are more arbitrary. We risk, perhaps, only observing parts of this system when we contrive lab-based choice settings, which often remove the situated aspects that motivated its evolution.

Elaborations of action-based models, called "embodied models," highlight that the distributed consensus architecture does not complete the decision before action initiation but rather continues deliberating afterwards, as the action unfolds in time. This is important for at least two reasons. First, as the deliberation continues, it is possible to change one's mind or revise an initial plan along the way if novel opportunities arise. Second, action dynamics change the dynamical landscape of choice (e.g., the motor costs required to pursuing the selected plans). To the extent that these changing factors can be incorporated in the deliberative process, action deployment feeds back on the decision process, hence breaking the serial assumption of classical models (Lepora & Pezzulo, 2015).

While these models seem more appropriate than classical models to deal with embodied decisions, current research has just scratched the surface of the possible ways the brain may implement decisions “in the wild” (and perhaps also in the lab). In the same way we are in need of novel methodologies to study embodied decisions, we also need a new theoretical framework to study them. It is important to acknowledge that future models of decision-making may radically change some of the core assumptions of current models. For example, embodied choice models are at odds with the classical serial view of decision-making. Decision models used in foraging theory (Charnov, 1976; Hayden & Moreno-Bote, 2018) are at odds with the standard definition of a decision as a competition between two (or more) offers and propose instead that the competition is between “exploit the current offer” and “explore alternatives”. While these decision models were originally developed to deal with offers that are not presented simultaneously (as is common in foraging studies), they can deal equally well with offers presented simultaneously (as common in lab studies) if one assumes that the offers are attended to serially (Hayden & Moreno-Bote, 2018).

Another fundamental limitation of classical decision setups (and corresponding models) is that they keep perceptual aspects of the task extremely simplified (to avoid confounds). However, the situation is completely different in embodied settings, which often require making decisions based on perceived affordances (J. Gibson, 1979) and body-scaled perceptions (Proffitt, 2006). Similarly, the field of decision-making often assumes that we have ready-to-use representations of utilities (Padoa-Schioppa & Assad, 2006), but instead I showed that utility surfaces for embodied decisions need to be computed in context dependent ways—for example, to reflect future affordances—and in real time. I showed examples of how to quantify present affordances (or probabilities) and future affordances (or utilities) to study embodied decisions in ways that are commensurate to the study of classical neuroeconomic tasks. However, it remains to be assessed whether these formalizations really reflect the ways our brains process affordances or are only convenient ways to study them (Pezzulo & Cisek, 2016).

In sum, the field of embodied decisions offers opportunities not just for conducting novel experiments but also for developing novel theories of decision-making and other cognitive processes. Future models of decision-making may not even answer the questions currently being pursued but instead lead to new types of questions; for example, if one sees the brain as a control system, the most important question is not “how does the brain represent knowledge about the world and subjective values”, but instead becomes “how does the brain learn to control increasingly complex interaction with the world”. In turn, the development of these theories could lead to a novel fundamental understanding of the brain as an organ that evolved for interaction, not for contemplation.

An additional challenge for future theories of embodied decisions is offering a more comprehensive description of “embodied decision-makers,” which considers for example the fact that they feel emotions, remember past experiences, and monitor their ongoing performance. In the literature review, I largely omitted the important influences of arousal, motivation and emotion on decision-making processes and the roles of our neurocognitive capacities for executive functioning, attention and working memory in the solution of embodied decision tasks. While a comprehensive analysis of the above processes is beyond the scope of this chapter, advancing the field of embodied decision-making will require their systematic analysis and integration. In turn, a systematic study of embodied decision settings might potentially help refine our taxonomies of cognitive, emotional, motivational and executive processes, which we often take for granted, but which might not correspond to separate processes or neuronal substrates (Barrett & Finlay, 2018; Buzsaki, 2019; Cisek, 2019; Passingham & Wise, 2012). One example that was discussed multiple times already is the deep involvement of sensorimotor brain areas in decision-making processes, which is at odds with classical taxonomies (Cisek & Kalaska, 2010; Gold & Shadlen, 2007; Song & Nakayama, 2009); but there are other putative segregations (e.g., between planning and attention processes) that might be less compelling if one considers embodied choice situations.

3.4.1 Potential impacts

The novel insights gathered from embodied decision experiments could have a significant impact on several fields, such as cognitive science, neuroeconomics, cognitive robotics and other disciplines interested in decision-making and sensorimotor control – as well as their deficits. As discussed, most decision-making studies in cognitive science and neuroeconomics focus on simple choices between fixed menus. Embodied decision studies can shed light on other kinds of decisions—and possibly decision-making circuits in the brain—that are more deeply integrated with sensorimotor processes than traditionally considered. As noted in the introduction, brain evolution has for millions of years been driven by the challenges of embodied decisions during closed-loop interaction with a dynamic environment (Ashby, 1952; Cisek, 1999; J. Gibson, 1979; Maturana & Varela, 1980; Pezzulo & Cisek, 2016; Powers, 1973), and the neural systems meeting those challenges have been highly conserved (Cisek, 2019; Striedter & Northcutt, 2019). It is likely that even the abstract abilities of modern humans are built atop that sensorimotor architecture (Hendriks-Jansen, 1996; Pezzulo & Cisek, 2016; Piaget, 1952; Pezzulo, Parr, & Friston, 2021) and would be difficult to understand outside of its context (Buzsaki, 2019; Cisek, 2019). Furthermore, clarifying the neuronal and computational processes that living organisms use to make embodied choices would be extremely important for sports psychology and analytics as well as for cognitive robotics. All these fields deal natively with decisions deployed in embodied settings and have developed useful methodologies to study them (see the discussion above of analysis methodologies) that have a great potential for cross-fertilization. From a technological perspective, the realization of effective robots able to operate in unconstrained environments (e.g., homes or hospitals) critically depends on their ability to address embodied decisions, such as the choice between available affordances (including social affordances). Unlike current AI systems that are largely non-embodied, these robots will need to more deeply integrate decision and sensorimotor processes, and this is why the novel paradigms of embodied decision studies (e.g., decide-while-acting as opposed to decide-then-act) could be most beneficial. Finally, the novel insights gathered via embodied decision studies can help shed light on deficits that involve a combination of motor and decision processes. For example, it has been suggested (Mink, 1996), that the inability of Parkinsonian patients to take a step (apparently, a purely motor deficit) might be due to an inability to resolve the competition between actions (hence, a decision deficit) or to endow them with sufficient vigor or urgency. The development of novel theories of how we make embodied decisions in the first place could help us better understand how these processes can break down, as apparent in Parkinson's or other diseases.

3.5 Conclusions

Every day we make countless embodied decisions, when we drive in a busy road, prepare a meal or play hide-and-seek with children. In this paper, I discussed how embodied decisions differ from classical economic decisions and which novel questions they raise that are poorly investigated. I then reviewed recent studies in fields like sports analytics and experimental psychology that start addressing these questions and distilled from them a general methodology to study embodied decisions more rigorously. I provided examples of how to cast key dimensions of embodied choices (namely, present and future affordances) into the attributes of classical economic decisions (namely, probabilities and utilities) - hence helping to align the two decision settings. My hope is that this contribution will raise interest towards the emerging field of embodied decisions and pave the way for the realization of novel experiments, novel formal tools and novel theories that capture more appropriately its most distinctive features.

In the next chapter, I present results from the first of three experimental contributions. In Study A, I also introduce a computational agent capable of leveraging probability and utility landscapes in line with the affordance calculus described in this chapter. Using a Monte Carlo sampling approach similar to that described in several of the examples in Section 3.2.3, the agent iteratively updates this representation during continuous exploration and navigation of an uncertain environment.

Chapter 4

Simulating mental simulation: active dynamical prospection (Study A)[1]

4.1 Introduction

Unlike single-celled organisms that can achieve adaptive success by responding reflexively to immediate changes in the local environment, humans, and likely a range of other animals (Gerrans & Sander, 2014), hold the capacity (and sometimes responsibility) of prospection: to foresee multiple plausible futures, and harvest information from them to guide action in the present. In an uncertain and constantly changing world, this is no small task.

How do we respond, for example, when confronted with a common challenge: we know where we want to go (perhaps we can even see our destination already), but we are not yet sure the best way to get there, or even if we can. This is the problem posed to agents during spatial navigation and pathfinding, and its solution may give us insights that extend into the more abstract domain of planning in general. Indeed, a wide array of human activities (from those that occur within seconds, to those that span years) can be framed as a filtering problem: how can we identify a subset of the infinite possible trajectories through the future that are most likely to take us somewhere we'd like to go? And further, how can we explore and evaluate these candidate paths before experiencing them directly?

In this chapter, I aim to analyze and model pathfinding behavior in a task paradigm that is more continuous and dynamic than those historically chosen by the planning literature. In the designed task, participants (and agents) must coordinate both visual exploration and navigation within a partially observable environment in which the dynamics of movement result in ongoing uncertainty about the true passability of potential paths.

This contribution has three primary components: 1) an analysis of behavioral data from a novel pathfinding task conducted as an online experiment, 2) a proposal to model mental simulation during navigation as particle filtering, and 3) an instantiation of this proposal in an agent capable of solving the task in ways that share attributes with human performance.

By developing a computational model of active perception, simulation and movement

[1]This chapter is based on work presented at ALIFE 2021 under the title "Active Dynamical Prospection: Modeling Mental Simulation as Particle Filtering for Sensorimotor Control during Pathfinding"(Gordon & Chuang, 2021)

during the pathfinding task, and comparing results with human behavioral data, I hope to shed light on the following questions:

- How are simulations of potential future actions coordinated during pathfinding and navigation?
- Which path characteristics attract attention and forward simulation?
- What are the distributional and temporal dynamics of attention, and how do they relate with pathfinding performance?
- Can a common computational mechanism successfully drive the coordination of both visual attention and navigation?

4.1.1 Background & related models

Situated planning

An extensive literature exists around planning across cognitive science, psychology, neuroscience, and artificial intelligence. Often, planning problems are posed in line with classical problem solving, in which the environment is fully observed with known dynamics (e.g. formalized as a Markov Decision Process), and solution entails identifying a sequence of actions resulting in a goal condition (Newell et al., 1972). In reinforcement learning, Monte Carlo methods are frequently used to sample trajectories during value estimation, and therefore to support the planning of future actions. (Silver & Veness, 2010) proposed Partially Observable Monte Carlo Planning (POMCP) to make value estimation tractable in high dimensional state spaces. In this chapter, particle filtering is used to efficiently approximate belief state updates when access to the true generative process is not available.

In embodied planning, agents are situated within complex, noisy, and uncertain environments in which, importantly, they must control both sensors and other motor outputs while simultaneously planning future actions in an online fashion. Though common for some time in robotics, efforts to develop theories of realistic embodied planning have recently gained momentum, propelled by multidisciplinary contributions from dynamical systems, ecological psychology, and reinforcement learning.

To select just a few examples, (Cos et al., 2021) demonstrated that perturbations to the arm during a reaching task can prompt changes of mind, indicating that deliberation continues dynamically during action execution. (Pezzulo, Donnarumma, Maisto, & Stoianov, 2019) proposed a connection between specific neural dynamics (sharp-wave ripples and theta sequences) as mechanisms to support planning in two regimes: at decision time, and in the background to optimize a behavioral controller. In a foraging paradigm, (Yoon, Geary, Ahmed, & Shadmehr, 2018) developed a model of normative utility based on the marginal value theorem, and applied it to a visual information harvesting experiment in which fixation duration (time spent at a patch) and saccade speed (movement vigor between patches) were measured. Their findings suggest a shared principle of control may underlie both aspects of foraging behavior.

Navigation, simulation & prospection

According to Montello (2005), navigation can be decomposed into two components: 1) locomotion, in which the body is coordinated to its local surrounds, and 2) wayfinding, in which a goal-directed agent plans actions aided by memory of both the local and distal environment. Though a range of neuroscientific mechanisms have been proposed to support both components (e.g. cells in the hippocampal formation encoding position, orientation, and head direction, among others), the dynamics by which internal models of the environment are queried offline (via simulation) and integrated with present sensory information (e.g. the observation of landmarks), is not well understood.

Mental simulation, often also referred to as replay or preplay, is the generation of internal sequences reflecting previous or possible engagements with the world. In a psychophysics experiment, Arnold et al. (2016) showed that humans adaptively compress simulations of potential routes during prospective route planning. (Chersi et al., 2013) developed a computational model of simulated and overt action during maze navigation, involving the hippocampus and striatum to support recall and cache action values respectively.

Especially important to the present work, is active inference, which has been proposed as a suitable formal theory to support the project of embodied perception, action, and planning (Friston, FitzGerald, Rigoli, Schwartenbeck, & Pezzulo, 2017). In active inference, agents act to reduce prediction error (free energy) produced from inconsistencies between an internal generative model and sensory observations. Recent work has applied the theory to planning and navigation, running simulations of navigation in a maze environment (Kaplan & Friston, 2018).

Finally, the literature on active navigation has investigated the relationship between sensory exploration and pathfinding. In a recent study, (Lakshminarasimhan et al., 2020) showed that eye movements could be used to infer latent beliefs such as the location of a hidden goal during virtual spatial navigation, and that controlling fixations had detrimental effects on navigation performance.

Swarm intelligence

This model also draws inspiration from (Trianni & Tuci, 2009), who argued that the integration of artificial life and cognitive science via “swarm cognition” could offer fruitful progress in understanding and modeling cognitive mechanisms. Swarm intelligence has demonstrated that simple unit-level behaviors (such as individual particle dynamics), when operating as a collective system, can produce complex emergent properties. In some cases, such as that of ant colony optimization, these system-level properties offer strategies effective at even NP-hard problems, such as the traveling salesman (Colorni, Dorigo, Maniezzo, et al., 1991).

4.2 Online pathfinding experiment

The task was designed to require participants to explicitly coordinate visual attention and navigation to a goal. On each trial, participants saw their present location (at the center of the screen), as well as the locations of one or more goals. A landscape of 50 ‘holds’ was initially hidden, and exposed when the particpant moved their cursor across the landscape,

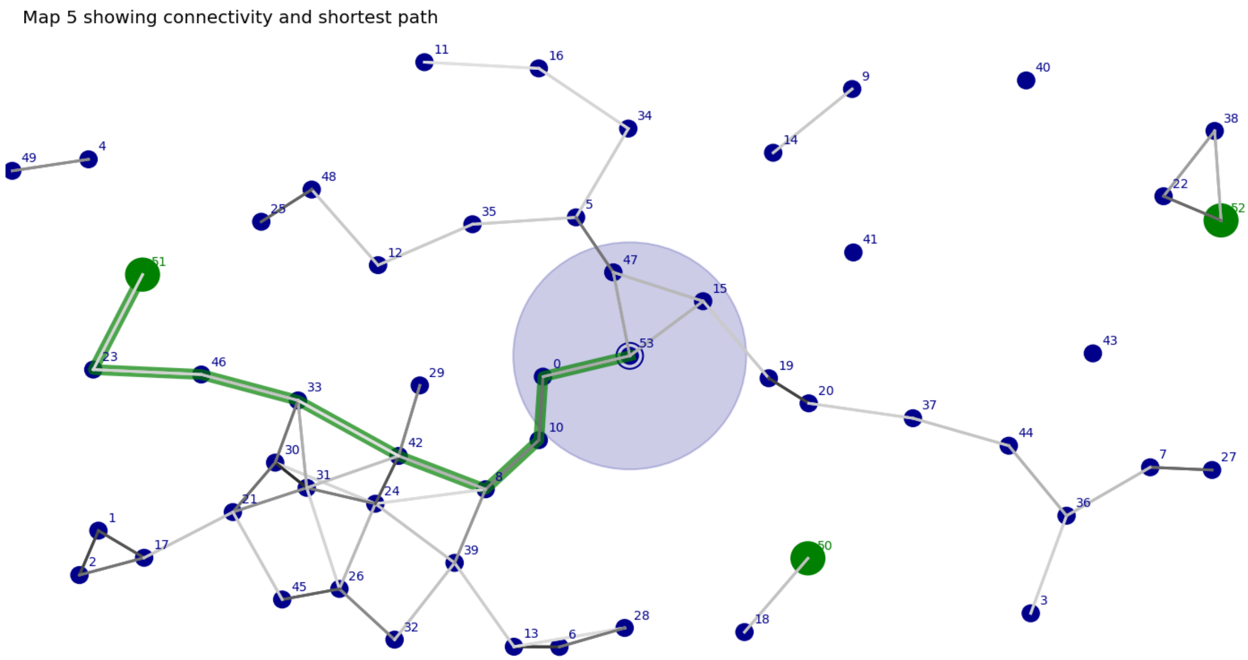


Figure 4.1: Sample map (map 5) with hold connectivity plotted as edges (lighter edges indicate gaps closer to reach limit). Optimal path to goal is plotted in green. The transparent blue circle indicates the reach radius, at the center of which, the small blue ring indicates the reach target at the agent's present location. Note that during a trial, the edges shown were not visible to the participant, and holds were only visible when near the cursor. Goal holds (green circles) were visible at all times.

exploring the map in a spotlight-like manner[2]. Holds were reachable only when within the 'reach zone', a fixed radius of the participant's present position indicated by a blue circle (as shown in Figure 4.1). To navigate to a reachable hold, the participant dragged it toward the small central target indicating their present location. During a successful drag, the full landscape shifted such that the chosen hold became centered within the egocentric space. In this way the participant was able to navigate towards and eventually reach a chosen goal.

By designing the task in this partially observable, egocentric manner, we were able to capture both movement and attention independently, and ensure that computational models of behavior in this paradigm contend with the richness of online sensorimotor exploration.

[2]While artificial attention collected in this way cannot be considered equivalent to eye-tracking metrics, this cursor-controlled spotlight method of capturing spatial attention in online experiments has been explored in other work. In particular, a validation study was performed by the authors of *MouseView.js*, a JavaScript library supporting experiments of this kind, and finding that patterns of dwell time were similar to those collected with the same stimuli in an eye-tracking experiment (Anwyl-Irvine, Armstrong, & Dalmaijer, 2021). Note that this study uses its own implementation of the cursor spotlight method since *MouseView.js* was released publicly after completion of data collection.

4.2.1 Participants

Study participants were recruited through an on-campus experimental lab at a public university in the United States.

81 participants completed the online study. Participants had a mean age of 22 ± 2.1. 61 identified themselves as women, 18 as men, 1 as non-binary or non-conforming, and 1 declined to answer. 56 reported their race as Asian, 13 white, 1 Black or African American, 1 American Indian or Alaska Native, and 7 Other, including White and Asian (2) and Middle Eastern (2). 9 participants identified as Spanish, Hispanic, or Latino. All participants were undergraduate students, graduate students or staff at the University of California, Berkeley.

4.2.2 Procedure

The experiment began with a series of instructions about the task. Participants completed a practice trial where they were guided through a trivial landscape to a nearby goal location to ensure they understood the mechanics of navigation and the trial objective. The full experiment entailed completing each of 11 predefined maps in randomized order. Each trial ended when a goal was reached, or when a 60-second trial timer expired. As an attention check, sessions were terminated early after 30 seconds of inactivity (the absence of any cursor movement). Pariticipants received a base incentive of \$6, and a performance bonus of \$0, \$2, or \$4 depending on final score as a percent of maximum (less than 60%, 60-80%, or more than 80%). The recruitment process and study protocol were approved by the local ethics review board.

4.2.3 Data structure

Two types of data were captured during each trial: navigation data, and 'attention' data. Navigation data included each attempt to navigate to a hold in the landscape, whether successful or unsuccessful, resulting in a final path through the landscape represented as a list of holds and timestamps. Attention data was recorded as a stream of 2D cursor coordinates (x, y) captured at 30 Hz.

Videos rendering all participants' navigation and attention data for all maps are available on the first author's website [3].

4.2.4 Behavioral data analysis

To compare performance metrics with map difficulty, and given the small number of maps in the dataset, I elected to define three difficulty categories (low, medium, and high-difficulty maps) based on the sample-wide success rate across all participants. In addition, we extracted a number of behavioral metrics from the raw navigation and attention data.

[3] All media is available at: `https://jgordon.io/project/adp`

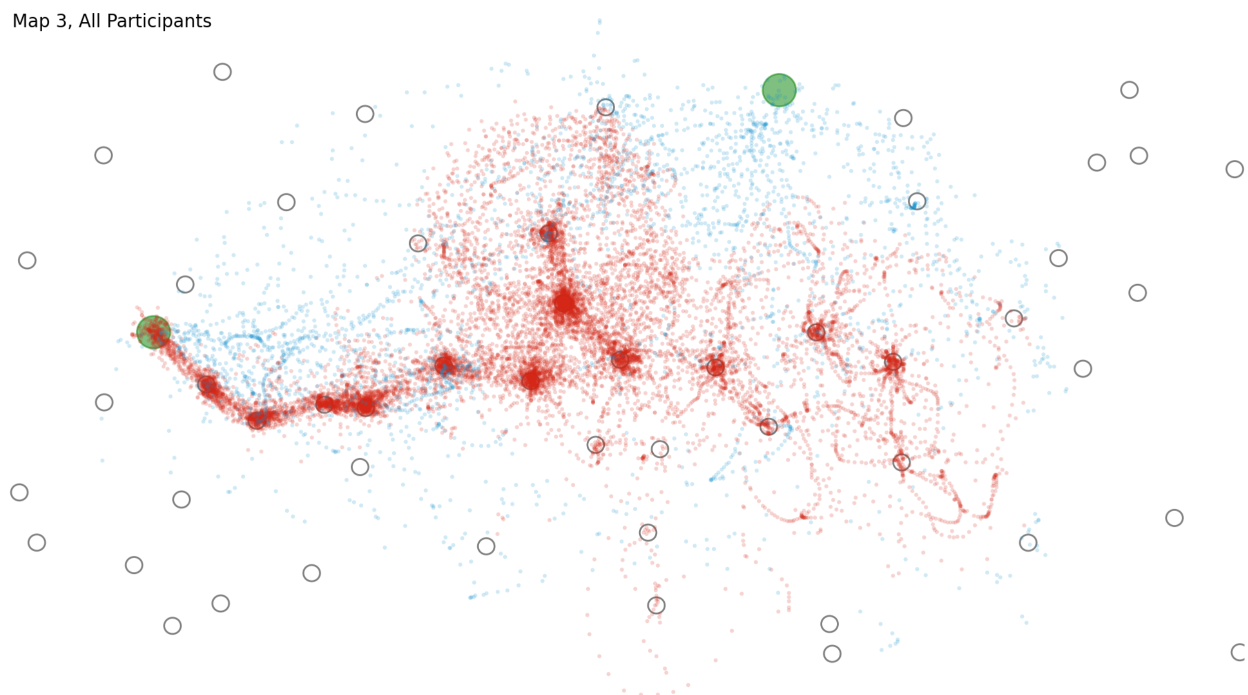


Figure 4.2: Plot of attention across all human participants (map 3). Red points indicate attention within reach zone. Blue points indicate attention beyond reach zone, a proxy for exploration.

Prospection and spatial exploration

By exploiting one feature of the task design—that cursor coordinates farther from the origin indicate visual exploration of more distant, not presently reachable, holds—we defined a key metric, attention distance, as the Euclidean distance of the cursor ("fovea") position from the agent (at screen center). Furthermore, by segmenting attentional coordinates as shown in Figure 4.2 into *reachable* (red) and *unreachable* (blue) groups, spatial patterns relating to map exploration could be directly visualized.

Trial score

In order to compare the computed behavioral metrics to an indication of both trial-level and participant-level performance, we define the score σ_{ij} for participant i and map j via the function:

$$\sigma_{ij} = f(path_{ij}, map_j) = \begin{cases} 0 & \text{if } success = 0 \\ \frac{\lambda_{min}}{\lambda_{pp}} & \text{if } success = 1 \end{cases}$$

Here, λ_{pp} was computed as the number of successful moves completed by the participant, and λ_{min} was a map-specific property defined as the minimum-length path to (any) goal. As such, $\sigma_{ij} \in [0, 1]$.

4.3 Computational model

4.3.1 Model rationale

I propose Active Dynamical Prospection (ADP), a model of planning related to active inference and augmented by ideas from swarm intelligence and dynamical systems. Following active inference, I assume that mental simulation may be leveraged to simultaneously learn, and plan within, a generative model of the agent's environment.

The computational model proposed here is guided by the following central hypothesis: that covert mental simulations supporting this task may be fruitfully modeled as Monte Carlo particle filtering[4] across a learned energy landscape, and subject to a set of precise physical dynamics aligned with the interaction capabilities of the agent within its environment. While (Tschantz, Baltieri, Seth, & Buckley, 2020) discusses the use of trajectory sampling to learn the generative density in active inference, what I propose is a stronger commitment to particle filtering as a descriptive model of simulation with possible links to covert attention.

In line with (Botvinick & Toussaint, 2012), I view pathfinding as prospective inference, or the act of reducing uncertainty over the ultimate trajectory an agent will take through its environment. In ADP, the agent learns a representation of the movement affordances in its environment, which can be thought of as a 2-dimensional energy surface. Given initial visual access only to its own location and that of the goal(s), agents begin with a sensible, but naïve, prior form for this surface, which we model as distance to nearest goal (see the gradient surrounding the goal in Figure 4.3b).

Agents leverage a set of three tools to uncover the true topology of their environment: 1) overt visual search by moving the fovea to expose hold locations, 2) navigating, by attempting to grab a nearby hold, which may be used both to confirm the true reachability, as well as to traverse the environment, and 3) simulated trajectories, modeled by particle filtering (Sequential Monte Carlo rollouts) over the present surface. The first two tools are specified by the task, and the third is the central mechanism of ADP.

Visual search As the fovea is moved across the map, precise visual information about the location (or absence of) holds is integrated into the energy surface. Specifically, regions absent of any hold are set to high energy values (indicating a vanishing probability the final path will land in this area), while the energy around regions where holds are discovered is reduced. However, the locations of holds themselves aren't sufficient to infer path passability, since one hold must be reachable from the other to allow traversal.

Dynamical prospection supports this function.

Dynamical prospection To support the learning of an energy surface suitable for navigation, prospection is modeled as parallel stochastic simulations of possible trajectories from

[4]Particle filters are a statistical technique named initially by (Del Moral, 1996) for the modeling of fluid mechanics, but applied to a wide range of domains, with their use in Bayesian inference and Hidden Markov Models most relevant to the present work. Particles represent samples from a posterior distribution over hidden states of a process, given noisy observations. By sequentially applying dynamics to each sample, trajectories (rollouts) can be generated which approximate changes to the density (belief state) over time.

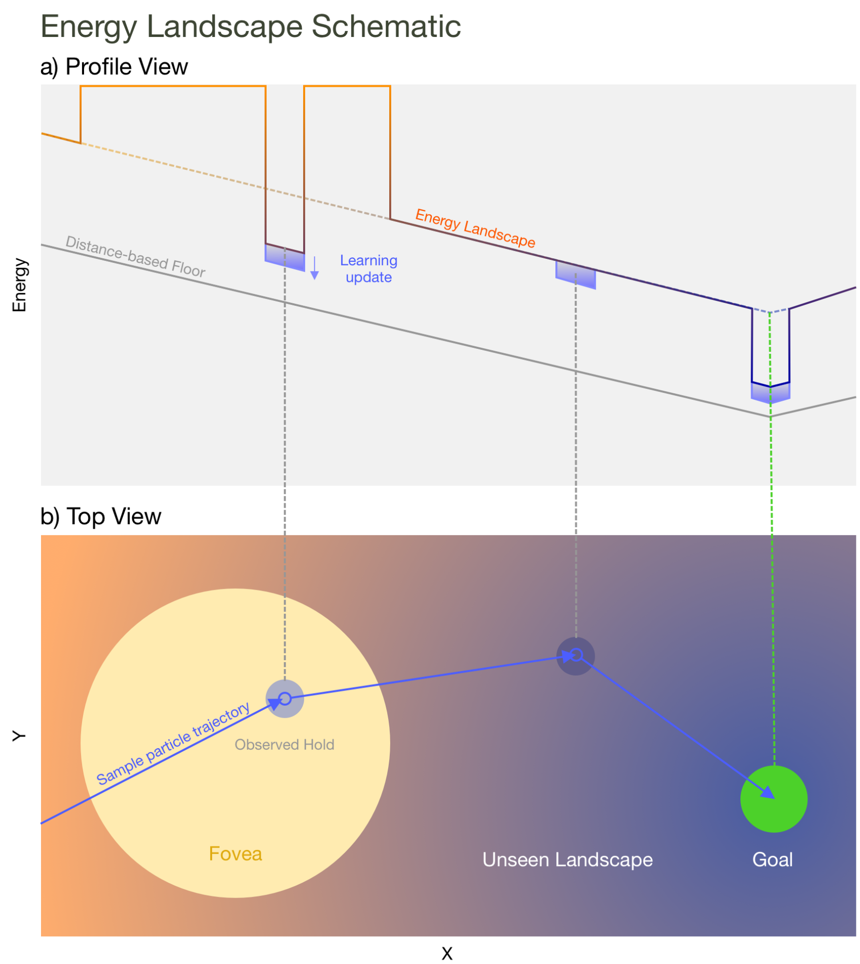


Figure 4.3: Energy landscape schematic (profile and top-down view of sample map region). Particle trajectories are influenced by an energy well (darker blue) in the landscape around the goal location. The path shown passes through the high energy location under the fovea (yellow) only because of the hold observed there.

the agent's present location. Particle dynamics sample a successor location from the present energy landscape, making hops to lower energy regions (where the agent already believes its path is likely to fall) more likely.

Rollouts—sequentially sampled trajectories—run in a latent world model, allowing simulation of paths that are not currently in the field of view. However, uncertainty about the locations and accessibility of holds in the model results in sampling variance and constraints to trajectory distance.

Central to ADP is the idea that simulated particle trajectories are most useful to an agent when governed by dynamics reflecting the agent-environment system's characteristics of interaction—both transition dynamics, and goal-seeking preferences. Specifically, particles are constrained by a simple intuitive physics: momentum and a distance-aware sample filter. Particle momentum reduces most quickly when moving up an energy gradient, and more slowly when descending, resulting in longer rollouts during descent. The sample filter limits consideration for a particle's successor location to an approximate reach-length radius, thus ensuring particle dynamics parallel the agent's own ability to traverse the landscape.

The sampled trajectories are sufficient to determine three effects which control the agent's internal representation and behavior:

1. Particles update the underlying energy at each sampled location, as a function of their terminal energy.
2. Because the agent expects to move through a low energy path, trajectories passing through high energy areas produce prediction errors. The location of these errors are used to attract visual attention (moves of the fovea), which generate observations to resolve this ambiguity.
3. The direction of the first step of each rollout determines confidence in the next (navigational) move. When directional variance drops below a threshold (parameterizing greediness), the agent attempts to reach in the consensus direction.

I hypothesized that ADP-Agent would exhibit behaviors useful to solving the pathfinding task such as: greedy exploration of direct paths to goal, focusing visual search on optimistic but still ambiguous candidate path locations, and dynamic planning demonstrated by iterative use of visual search and hold traversal. Altogether, I expected the proposed model to be capable of solving maps with similar difficulty to those solvable by humans. A detailed description of the model implementation follows.

4.3.2 Model details

Agent task paradigm The agent task paradigm was modeled to maximize consistency with the problem posed to human participants, while abstracting away low-level motor dynamics like controlling a cursor during click and drag.

Agent state is a tuple $(X_{agent}, X_{fovea}, E)$, where $X_{agent} \in \mathbb{R}^2$ is the agent's location in the map, $X_{fovea} \in \mathbb{R}^2$ is the agent's fovea location (which determines the position of the spotlight), and E is the internal model of the map as an energy landscape.

On each time step, the agent receives an observation from the local area around its fovea, which includes the positions of all holds in the map within a fixed foveal radius. The agent then chooses an action, composed of the next position for both the agent and the fovea: $A_t = (X_{agent}(t+1), X_{fovea}(t+1))$. The agent need not move itself nor its fovea on every time step. The environment updates in response to the chosen action by 1) moving the agent location to $X_{agent}(t+1)$ if this location is reachable (distance within reach radius), and 2) moving the fovea to $X_{fovea}(t+1)$, taking multiple steps if fovea distance is greater than the maximum fovea velocity. If the agent's new position lies within a goal, the trial is completed successfully.

ADP-Agent ADP-Agent is instantiated with an energy landscape represented as a 2D matrix or raster $E^{W \times H}$ where each $e_{xy} \in [0, 1]$ represents the energy at that point in the landscape. W and H are parameters specifying the resolution of the agent's energy landscape. A distance-based energy floor, E_{floor}, is separately defined, and calculated as the

Euclidean distance to the closest goal location. E is initialized to $E(t_0) = E_{floor} + C$ where C is a constant.

The following additional parameters influence various aspects of agent behavior:

- k: Number of particles to emit per step
- τ: Softmax temperature for particle location sampling
- Particle mass m: Inverse of rate at which particle momentum is reduced during rollout
- α: Learning rate for energy updates
- Move consensus threshold η: Percentage of first particle steps landing on the same hold required to attempt a move
- d: Energy decay rate (towards initial initial energy $E(t_0)$)

On each time step, ADP-Agent performs k particle rollouts, instantiating each at the agent's present location X_{agent}[5]. Rollouts are computed by considering only locations in the energy landscape within an approximate reach radius from the location of the particle (for convenience, this is implemented as a circular binary mask centered at position X, with radius r: `Mask(X, r)`), resulting in a candidate subset of the landscape E_c. Particle dynamics then follow a softmax, such that the next location is sampled as:

$$X_{p,j+1} \sim p(X_{p,j+1}|X_{p,j}, E) = \frac{exp(-E_c/\tau)}{\sum_i exp(-E_{c,i}/\tau)} \tag{4.1}$$

Particles lose momentum as a function of the change in energy of the landscape: $p_{j,mnt} \longleftarrow p_{j-1,mnt} - \frac{E[X_{p,j}]-E[X_{p,j-1}]}{m}$, and dampened by the particle mass parameter, m. The rollout continues until the particle's momentum falls to 0 or below.

Learning After each particle i terminates, the energy landscape is updated underneath each step of its trajectory $\pi_i = \{X_{i0}, X_{i1}, ..., X_{in}\}$ by an approximate momentum-discounted learning rule based on the difference between the energy at each step $E[X_{ij}]$, and that at the terminal location of the rollout $E[X_{in}]$.

$$E \longleftarrow E + \alpha p_{ij,mnt}(E[X_n] - E[\mathtt{Mask}(X_{ij})]) \tag{4.2}$$

This learning update serves to push the landscape energy towards the terminal energy as illustrated in Figure 4.3.

Following all rollouts and updates, the landscape is multiplicatively decayed (by rate d) towards its initial conditions, and clipped to the interval $[E_{floor}, 1]$ after each step:

$$E \longleftarrow clip(E + d(E(t_0) - E), E_{floor}, 1) \tag{4.3}$$

[5]Alternative strategies were also explored, including sampling an origin from the landscape, or alternating between the fovea location and the agent. The agent-origination strategy was most robust in simulations run.

To choose a new location for the fovea, an error map Ψ is computed by summing the energy under every particle trajectory step. In this way, areas of high surprise (particle trajectories moving through high energy regions) can be efficiently calculated as a direct result of the rollout computation.

$$X_{fovea}(t+1) = \arg\max_{X} \sum \Psi[\texttt{Mask(X)}] \tag{4.4}$$

The second component of action, $X_{agent}(t+1)$, is determined based on the uncertainty (entropy) of first step directions over all trajectories. The set of first step directions (for a particle batch) is defined as $\Theta = \{\theta_i\}_{i=1}^{k}$ where $\theta_i = arctan(\frac{p^i_{y,1} - p^i_{y,0}}{p^i_{x,1} - p^i_{x,0}})$. The variance is then calculated, and if $Var(\Theta) < \eta$, the agent attempts to reach the hold upon which the plurality of its first steps $(p_{x,1}, p_{y,1})$ fall.

Videos of sample agent runs can be found at the media page linked in the 'Data Structure' section above, and Python code is available as a public repository.

4.4 Results

4.4.1 Online (human) experiment results

Prospection via attention distance

I investigated both distributional and temporal characteristics of attention distance. At the trial level, a positive correlation is found between the mean of attention distance and score across all three difficulty levels. This relationship is statistically significant for medium and high-difficulty maps, but not for low-difficulty maps (see Figure 4.4 for details and OLS regression results).

To identify patterns in temporal attention data, maximum attention distance was computed and binned based on progress through trial, which allowed the standardization of longitudinal data across trials of varying duration. As shown in Figure 4.6, a general trend is seen showing reducing distance as trials progress, as well as a positive relationship between distance and map difficulty. The downward trend was shallowest for the lowest-performing participant segment.

As another perspective on prospective and exploratory behavior, I analyzed the delay prior to first move. On trials with longer delays (wherein participants explored the landscape for longer prior to navigating to their first hold) success rate overall was lower (see Figure 4.5). However, when looking only at successful trials, a statistically significant association between delay and trial score was observed.

4.4.2 Simulation results

Simulations were run using the same maps and task constraints as those used in the online experiment. 81 simulations were run on each map, with identically instantiated agents. I

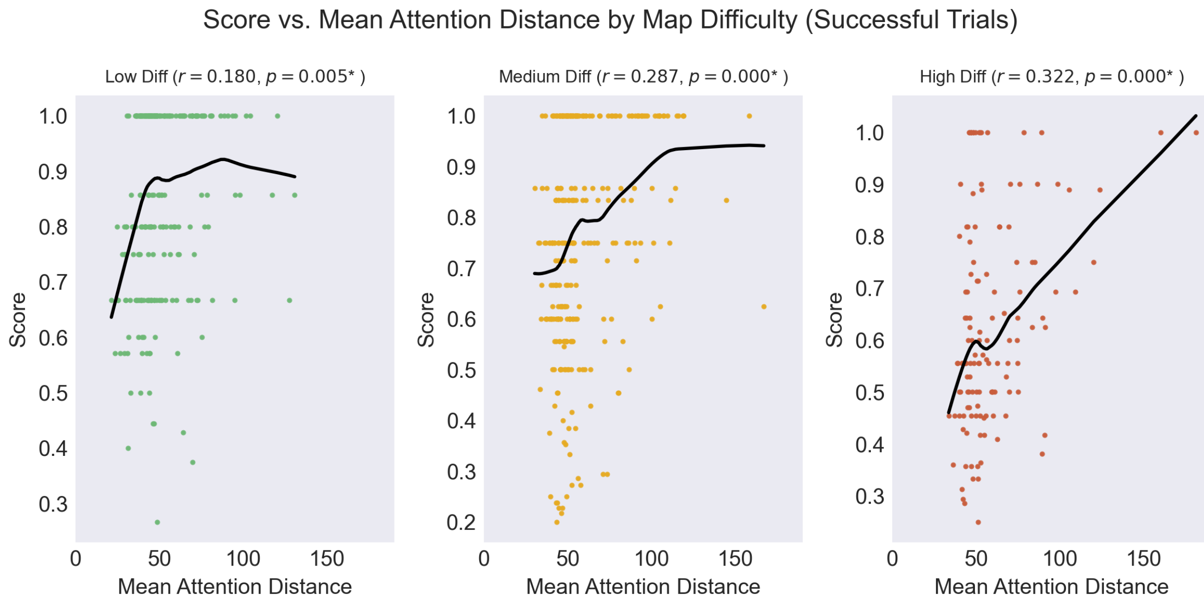


Figure 4.4: Regressions of trial-wise score versus mean attention distance in successful human trials. An increasingly strong positive correlation is seen as map difficulty increases. Linear regression results (Pearson-r and p-value) from each OLS regression are shown in title, with LOWESS plotted in black.

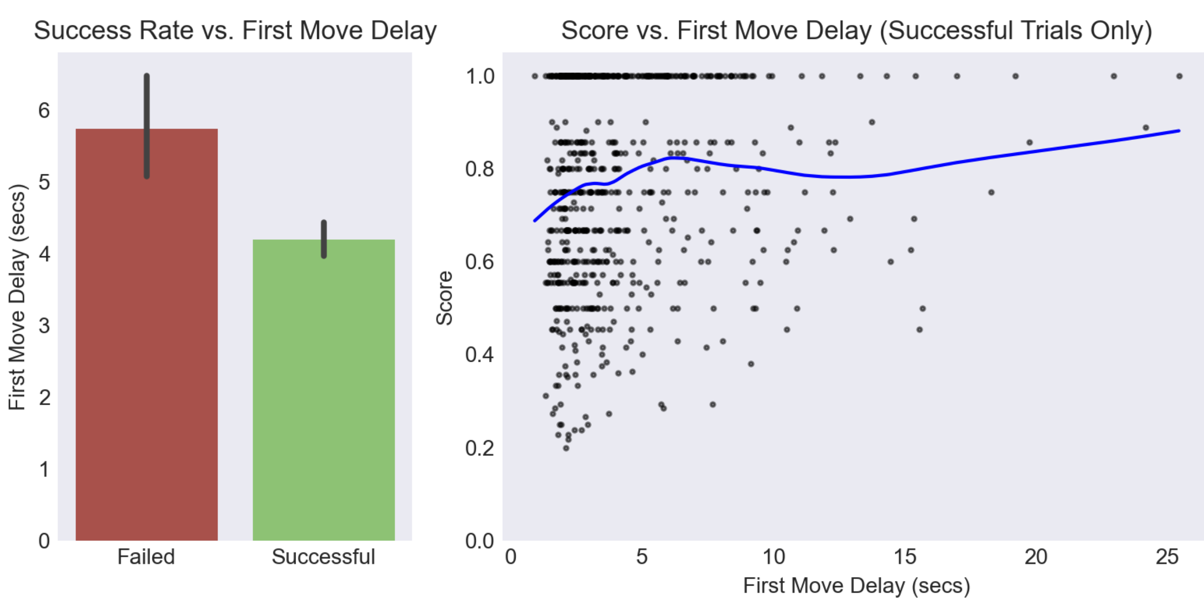


Figure 4.5: Left: Trial-wise success versus delay before first move (in seconds). Right: Scatterplot of trial score versus delay, among successful trials. Blue line shows LOWESS fit. Linear regression shows significant relationship (Pearson-$r = 0.11$, $p = 0.003$).

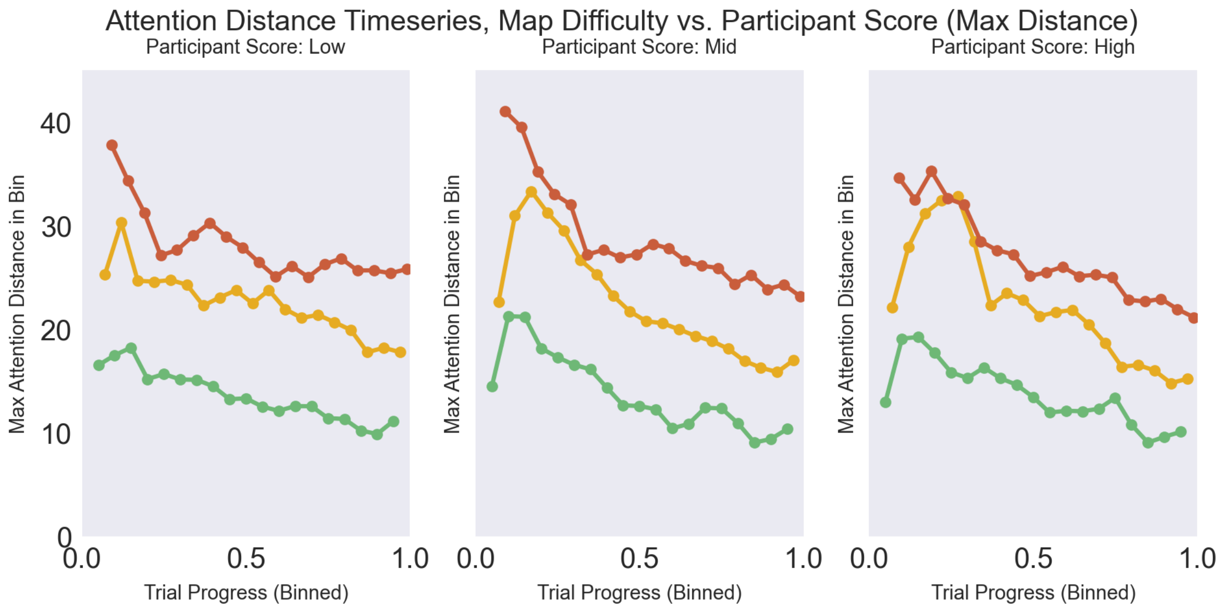


Figure 4.6: Binned longitudinal max attention distance across map difficulty (green, orange, and red line series indicating low, medium and high difficulty), and participant performance groups (left, middle, and right charts). We can see consistent declining trends across each trial duration as exploration reduces during navigation (exploitation).

compared four primary outputs of simulation runs with the results from the online experiment: success rate, as well as the distributions of goal reached (for maps with multiple reachable goals), distribution, and spatial attention.

Results show that all maps can be successfully solved by ADP-Agent, and that success rates are well correlated with that of human participants (see Figure 4.7). While some maps showed similar goal preferences, indicating related goal choice dynamics, a minority showed inversed preferences (e.g. maps 2 & 11). Trial duration was also highly correlated (Pearson-$r = 0.81$, $p < 0.005$). In aggregate, spatial attention exhibits visual similarities in its patterns of exploration and navigation. As an example, see the heatmap comparison at left in Figure 4.8, which shows both humans and agents exploring a dead end south of the starting position, before identifying a connected path to goal).

4.5 Discussion

Behavioral results showing an increasingly significant correlation between mean attention distance and trial-wise score can be interpreted as the value of exploratory distal attention. Though weaker for low difficulty maps, which might be expected since a greedy no-look-ahead policy was still effective in these cases, participants (and agents) could easily get stuck in dead ends if they didn't confirm connectivity prior to movement down a path.

The temporal trend seen in attention distance, as well as the relationship between first move delay and score, suggest that participants had to balance visual exploration with

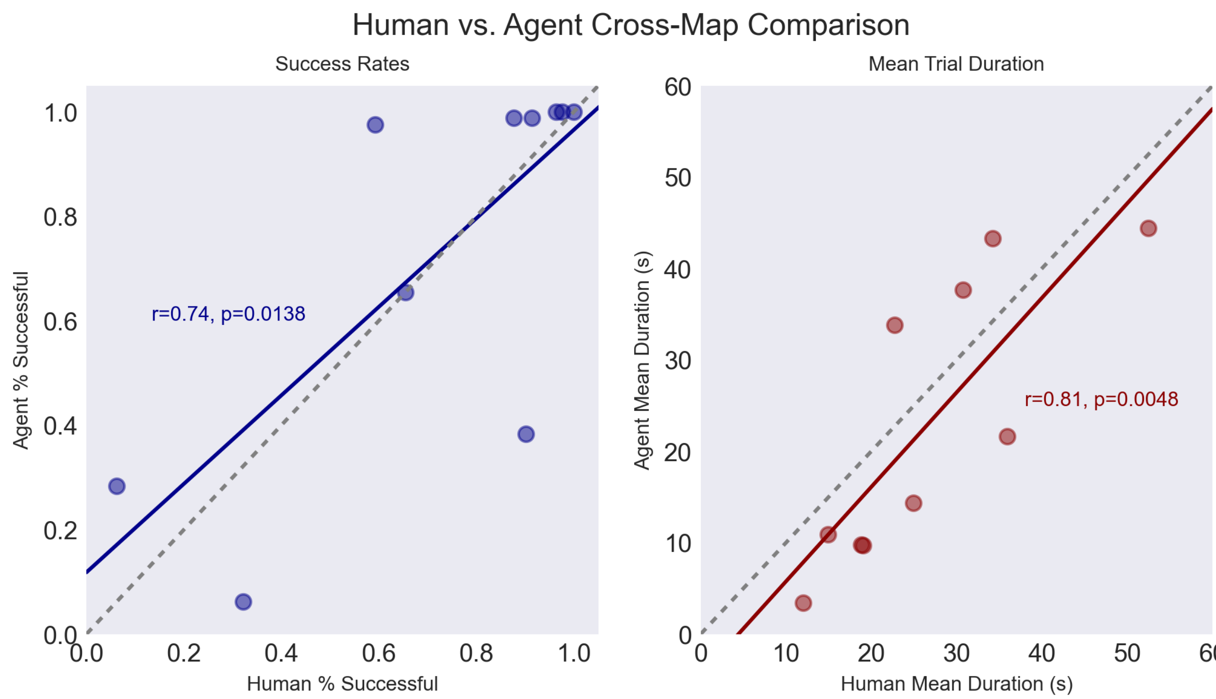


Figure 4.7: Comparison of human and agent success rates (left) and trial durations (right) across all maps.

navigation in order to succeed at the given task. With failed trials filtered out (as was done for the delay versus score regression in Figure 4.5), score can be seen as a proxy for path efficiency. While participants spending too long exploring prior to navigation were less likely to succeed (within the 60-second time limit), for those who were successful, exploratory behavior prior to committing to a spatial direction was predictive of path efficiency.

In the following sections I discuss simulation results and their comparison with human behavioral data. While quantitative comparisons of path choice, trial duration, and success rate offer some validation that simulated agents generate attributes that are, in aggregate, consistent with human planners, insights can also be derived from qualitative analysis of behaviors seen in single simulation runs.

Epistemic value via prospection As particles move through previously observed terrain (across high contrast or "well-worn" trajectory segments in the landscape), they follow predictable paths. However, when moving into unexplored terrain, the sampling dynamics generate splits and radiating branches guided only loosely by the underlying distance-based floor. Trajectories venturing into these higher energy regions produce large areas of prediction error, which suggest epistemic richness given a combination of high expected surprisal, and high path salience. ADP-Agent's foveal policy, which moves attention to the area containing maximum prediction error on the prior time step, therefore serves to expand the peripheries of the known landscape where it is most likely to yield paths to a goal.

I also find dynamics in which particles "jump off" a path of observed holds influenced by an energy well from a nearby goal—even when these jumps take particles into unobserved

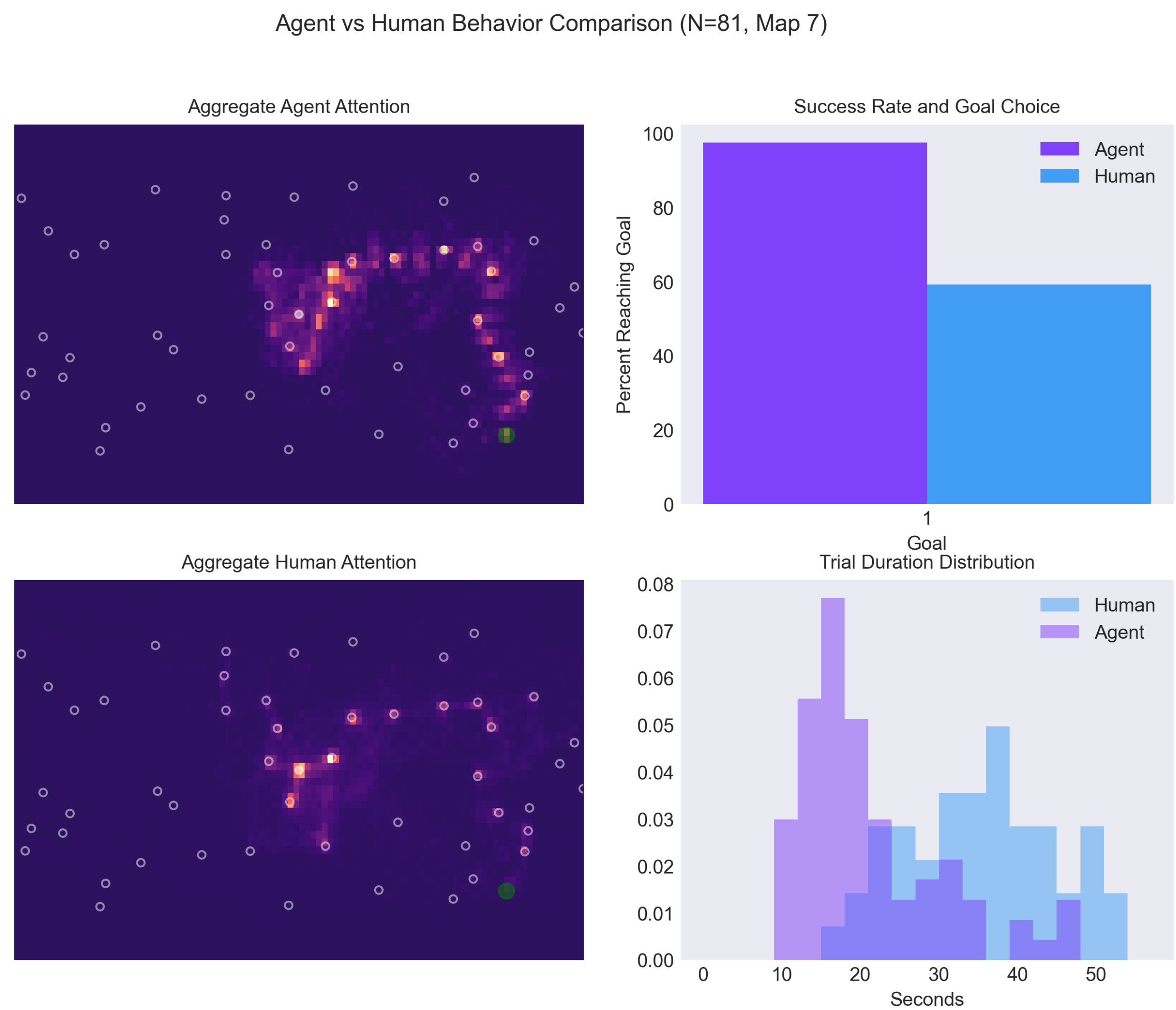


Figure 4.8: Human vs. agent comparison (map 7). Left: aggregate attention outputs from all runs. Top right: goal reached success rate. Bottom right: run duration distribution. Agent simulations solved map 7 more quickly (in simulated time), and more consistently than humans.

regions. These trajectories might be thought of as optimistic shortcuts, and the high prediction errors they produce attracts visual attention to confirm or deny the hypothesis of path connectivity.

Attention & surprise A feature shared by human and agent attention is a focus on holds that are close to the reach limit, but not in fact reachable. Though distal scans of these connections may be assessed as passable (by humans, as well as by optimistic particle trajectories), upon arriving at the hold, a failed reach attempt prompts subsequent attempts, or consideration of alternative nearby paths.

Other map attributes that are seen to attract attention across both simulations and behavioral data are symmetrical forks (in which two holds appear to lie on similarly direct paths to goal), and other regions of uncertainty caused by competing candidate trajectories. ADP-Agent fixates on these regions during increasingly long rollouts until a confidence threshold is reached[6]. In general, the agent model appears to leverage the parallelized nature of prospective simulations, with serially executed attentional movements supporting uncertainty reduction at the areas of highest error.

Forgetting An attribute of ADP-Agent not initially modeled, but discovered to be empirically valuable, was to "forget" holds upon exit (implemented by setting energy to 1 around holds following a successful reach to an adjacent hold). When forgetting was disabled, some agent simulations were seen to emit trajectories back through previously visited holds (rather than unexplored territory), despite this typically requiring ascent of the energy gradient.

Search depth & backtracking A common challenge in complex planning problems is the optimization of search depth, to avoid actions leading to dead ends. Backtracking was common in both human and agent simulations, especially in high difficulty maps including direct but ultimately disconnected paths. Forward search depth is modulated by ADP-Agent's particle mass and move consensus threshold parameters, which affect the length of trajectories, and navigational greediness, respectively. Empirical optimization of these parameters to a specific map (via grid search) was usually sufficient to achieve 100% success rate on even the most challenging problems.

4.5.1 Limitations & future work

The model presented here lacks several features inherent to human pathfinding that may limit its ability to predict and explain behavior. First, ADP-Agent is unable to generalize or treat clusters of holds or path segments as more abstract units. For example, while human participants likely perceive a sequence of closely positioned holds as a single passable route affording traversal from start to end, the landscape in this model independently represents an energy well around each hold. Secondly, some human attentional data appeared consistent with bi-directional planning (a well-known dimensionality reduction strategy long studied in psychology and artificial intelligence, e.g. (Pohl, 1971)), especially when confronted with challenging problems. In contrast, ADP-Agent's attention was seen to progress roughly monotonically towards goal locations driven by errors on the periphery of the observed landscape. Experimenting with particle emissions strategies that support inverse rollouts from goal locations may begin to address this limitation.

In this study, agent parameters were selected empirically based on a trivial success rate criterion. This was done to avoid overfitting to human behavior, and therefore making it difficult to meaningfully analyze similarity. An alternative path for future work, however, is to fit parameters to individual participant data, and seek correlations between parameter values and trial performance, as well as other individual characteristics such as spatial reasoning, working memory capacity and risk aversion.

[6]This is a dynamic consistent with theories of evidence accumulation, e.g. (Lee & Cummins, 2004).

4.6 Conclusion

In this chapter, I propose a computational model of visual exploration and control during pathfinding in a partially observable and uncertain environment. Behavioral data from an online experiment provides further insight into the range of strategies employed by humans in this task. Results from simulations show that agents can successfully solve the task by minimizing prediction error generated from particle rollouts across a learned energy landscape. Quantitative and qualitative similarities seen between human behavior and simulation results encourage further exploration of particle-based models of mental simulation.

In Chapter 5, I will present a study building off of, and overcoming limitations of this work in several ways. The experimental setup in Study B explores a kind of uncertainty more typical of naturalistic pathfinding, and does so within a more immersive setup made possible by virtual reality. By decoupling planning from navigation, and enabling the collection of eye movements across both phases, I am able to begin to answer an expanded set of research questions focusing on prospective simulation dynamics.

Chapter 5

Embodied spatial planning: a VR study of eye movements during planning under uncertainty (Study B)

5.1 Introduction

When planning future actions, e.g. finding the best route to a novel destination or the best sequencing for our daily tasks, we are often confronted with very large state spaces that we cannot search exhaustively. These exist among a class of problems we frequently encounter as humans which involves searching for a sequence of candidate actions likely to reliably result in a desired "goal" state, while also contending with limited (cognitive and motor) resources. Prospective search problems of this type are common in the realm of spatial navigation, but also apply to much more abstract domains of cognition (Newell & Simon, 1972; Hills et al., 2015).

From a formal perspective, searching in large state spaces is different from searching in small ones that permit exhaustive analysis, and as a result, requires a different class of algorithms (Browne et al., 2012). Indeed, in discrete state spaces with known dynamics, it is possible to use simple decision trees that capture a complete set of possible action sequences, to identify paths resulting in a desired state, as well as an evaluation of their relative efficiency (e.g. by weighted length). However, in large state spaces (including tasks where states are continuous rather than discrete), the combinatorial nature of possible action sequences results in a dimensionality explosion that makes exhaustive search intractable. Even when loops are excluded (to avoid an infinite search space), the number of sequences to check exceeds the computational capabilities of both biological and artificial problem solvers. In machine learning, the problem of searching in large state spaces has lead to the development of various effective techniques to *sample* the most promising candidate actions (Browne et al., 2012). For example, Monte Carlo Tree Search (MCTS) algorithms, when coupled with deep neural networks, provide extremely effective solutions to challenging high dimensional games such as chess and GO (Silver et al., 2016). In parallel, the identification of algorithms and heuristics organisms use to overcome these challenges has become a rich target of inquiry for cognitive science and psychology.

When faced with large search problems, living organisms (as well as artificial systems) must contend with the *breadth-depth dilemma*, in which the problem solver must decide how to direct limited resources (e.g. exploratory actions which consume time and energy). Typically, spending more time reasoning about (simulating) sequences within a particular region of a state space offers more precise estimates of its value (i.e. in pathfinding, the likelihood of particular positions being part of an efficient route to destination), at the cost of reduced precision in other regions. A number of studies have attempted to model and empirically measure human handling of this dilemma. In a study in which participants had to select between 121 options that provided different rewards, Wu et al. found that people sometimes adopt a sampling strategy similar to the ones used in machine learning and that they are able to use spatial structure (when present) to generalize effectively (C. M. Wu, Schulz, Speekenbrink, Nelson, & Meder, 2018). In a problem solving study resembling the traveling salesman problem, Eluchans et al. found that people adapt their planning strategy to the minimal depth required by the problem, but use a much lower depth when the problem becomes too cognitively demanding for exhaustive search (Eluchans et al., 2023). In a patch foraging-like experiment where the allocation of samples needed to be determined prior to sample feedback, Moreno et al. found that individuals prefer breadth when capacity is low, and depth when capacity is high (Moreno-Bote, Ramírez-Ruiz, Drugowitsch, & Hayden, 2020). Interestingly, even when sufficient capacity is available to sample from all states, optimal behavior involved ignoring the majority of states in order to more deeply explore a subset. In a related study, the number of states sampled increased with capacity approximately according to a power law (with exponent $\approx \frac{3}{4}$) (Vidal, Soto-Faraco, & Moreno-Bote, 2022).

The study of behavior (and embodied intelligence) offers other perspectives on how organisms contend with the challenge of search in vast state spaces. One fundamental strategy that has emerged both from behavioral studies, as well as proven successful in computational models and reinforcement learning, is the use of forecasting from the embodied present (Chaslot et al., 2008; Dayan, 1993; Momennejad & Howard, 2018). Indeed, it is perhaps unsurprising that the current state of an organism, as well as sensory observations available to it in the present, provide a valuable, grounded starting point to search in more distal regions of the problem space. Studies have shown that when planning routes in novel mazes, people focus their attention and search resources on *salient* locations or landmarks, which provide more information gain. For example, during virtual navigation, people display more deliberative behavior (as indexed by a behavioral index called *vicarious trial-and-error* (VTE) first studied by Tolman during rodent navigation (Tolman, 1948)) in locations of the maze where making optimal decisions required using more information (Santos-Pata & Verschure, 2018; Lancia, Eluchans, D'Alessandro, Spiers, & Pezzulo, 2023). When people have access to an external representation of the search space (e.g., a map of the maze) prior to navigation and can therefore form a prospective plan, they again focus their resources to encode only a subset of map elements that are relevant for their (optimal) plan, therefore constructing a simplified mental representation of the problem (Ho et al., 2022). Similarly, people may be able to adaptively *compress* simulations of potential routes during prospective pathfinding (Arnold et al., 2016), using mechanisms that are potentially analogous to the (time-compressed) '(p)replay' of experience found in the rodent hippocampus and thought to support spatial planning (Buzsaki, 2019; Pfeiffer & Foster, 2013; Pezzulo, Van der Meer, Lansink, & Pennartz, 2014; Pezzulo, Kemere, & Van Der Meer, 2017).

Another useful approach to search in large state spaces consists of decomposing the problem into smaller and more manageable sub-problems, thus using a hierarchical approach to planning as opposed to "flat" tree search. Various studies indicate that people use such hierarchical planning strategies when the environment permits this sort of decomposition. During navigation in a maze structured similarly to the "Hanoi tower" problem (Solway et al., 2014), participants faced with a choice between two paths of the same length preferred the path that crossed fewer community boundaries (as identified by graph theoretic measures), which is "shorter" as quantified by information-theoretic measures that reflect the hierarchical clustering of space (Donnarumma, Maisto, & Pezzulo, 2016; McNamee, Wolpert, & Lengyel, 2016). Another study reported that people spontaneously organizes space into clusters that support hierarchical planning and subgoaling (Tomov, Yagati, Kumar, Yang, & Gershman, 2020). A neuroimaging study suggested that when planning routes in a virtual subway network, people exploit the problem structure—and in particular the division into lines and stations—to form hierarchically organized plans (Balaguer et al., 2016). Interestingly, besides hierarchical problem decomposition, individuals often exhibit sensitivity to the structure of the environment in their planning and navigation strategies. For example, a large scale virtual navigation study showed that people are better at wayfinding in mazes having the same general structure as the place where they grew up (Coutrot et al., 2022).

Many of the studies reviewed above measure navigation behavior to infer planning strategies. However, planning is a largely covert process, implying that key dynamics may be missed by studying only wayfinding decisions. An alternative approach to study planning in large state spaces consists of recording eye movements while people visually scan the environment prior to beginning navigation. Given its flexibility and speed, the visual system affords a rapid virtual simulation of candidate routes prior to actual movement. Therefore, looking at eye movements may provide privileged access to the covert stages of planning. Along these lines, in a recent study of visual search prior to a navigation task, Zhu et al. showed that participants' eyes made fast 'sweeps' that traced the trajectories that participants later selected, and that gaze targets balanced fixations on reward locations and critical transitions (walls or openings) in the environment—findings which resemble results in neural replay (Zhu et al., 2022). Another study of visual search in humans has shown that eye movements themselves exhibit behavior consistent with multi-step planning (in this case, planning of a gaze sequence matched to the shape of a given search space), rather than greedy selection of informationally rich or otherwise salient locations (Hoppe & Rothkopf, 2019). In a primate study, Lakshminarasimhan et. al. 2020 showed that eye movements could be used to infer latent beliefs such as the location of a hidden goal during virtual spatial navigation (Lakshminarasimhan et al., 2020). Controlling fixations also showed detrimental effects on performance, indicating the involvement of free visual search in effective navigation.

Taken together, this literature indicates that when planning in large state spaces that cannot be searched exhaustively, people (and other mammals) use a range of effective strategies: they allocate limited resources by privileging information-rich states, they attend to the most salient aspects of the problem, and they split problems into more manageable sub-problems. However, despite these achievements, we still lack a comprehensive account of the planning strategies that people adopt when confronted with the exploration-exploitation dilemma inherent to visual exploration of large, novel and uncertain environments. Interestingly, while behavioral studies have found signatures of hierarchical planning during navigation, we still do not know how these hierarchical navigation plans might be established during visual exploration prior to movement.

Through the work presented in this chapter, I focus on a number of open questions related to the temporal dynamics of planning in large, informationally rich, but uncertainty-laden, environments. Specifically, I am interested in the following:

1. How do people search for future-oriented information when a very large number of options is available, and time resources are limited?
2. What can overt motor behaviors such as eye and arm movements tell us about hierarchical aspects of the internal representations supporting planning?
3. How do participants generate and adapt plans to contend with predictable regions of uncertainty?
4. How do information and uncertainty-related attributes of the environment impact navigation plans and visual search dynamics?

To investigate these questions, I designed a novel Virtual Reality-based spatial navigation task where planning is explicitly required, and collected eye tracking data while participants searched the landscape prior to navigating through it. In contrast to standard wayfinding experiments in which only one goal location is present, the maps were designed with multiple candidate goal locations ("caches"), and participants did not know in advance which ones would be populated by actual rewards. This design was selected to enable the study of uncertainty-laden contexts common to naturalistic planning.

To preview my results, signatures of hierarchical planning are observed in gaze dynamics, which likely aid participants in efficiently transforming rich spatial information into a simpler, contingency-aware, plan of action. Examples of such signatures include a decreasing trend seen in gaze distance from origin, and a broad to narrow shift (with reducing saccade distances and longer fixation durations) as plans are established. In line with prior work, "critical tiles" to which landscape connectivity is most sensitive were the strongest predictors of visual attention. Finally, performance-group-level analyses show that the most effective participants leveraged offloading of initial navigation direction through orientation of the arm, navigated paths closer to regions of the map visually attended during planning, and exhibited greater eye movements with shorter fixations, as well as a focus on more peripheral tiles during planning.

5.2 Methods

5.2.1 Participants

Participants were undergraduates, graduate students and staff at the University of California, Berkeley. Participants were invited by email to complete a screening questionnaire. I excluded participants who were deemed high risk for adverse effects of using VR. This include those who have used VR in the past and experienced severe dizziness or nausea during usage, as well as those with epilepsy, a history of seizures, heart ailments, or those who respond that

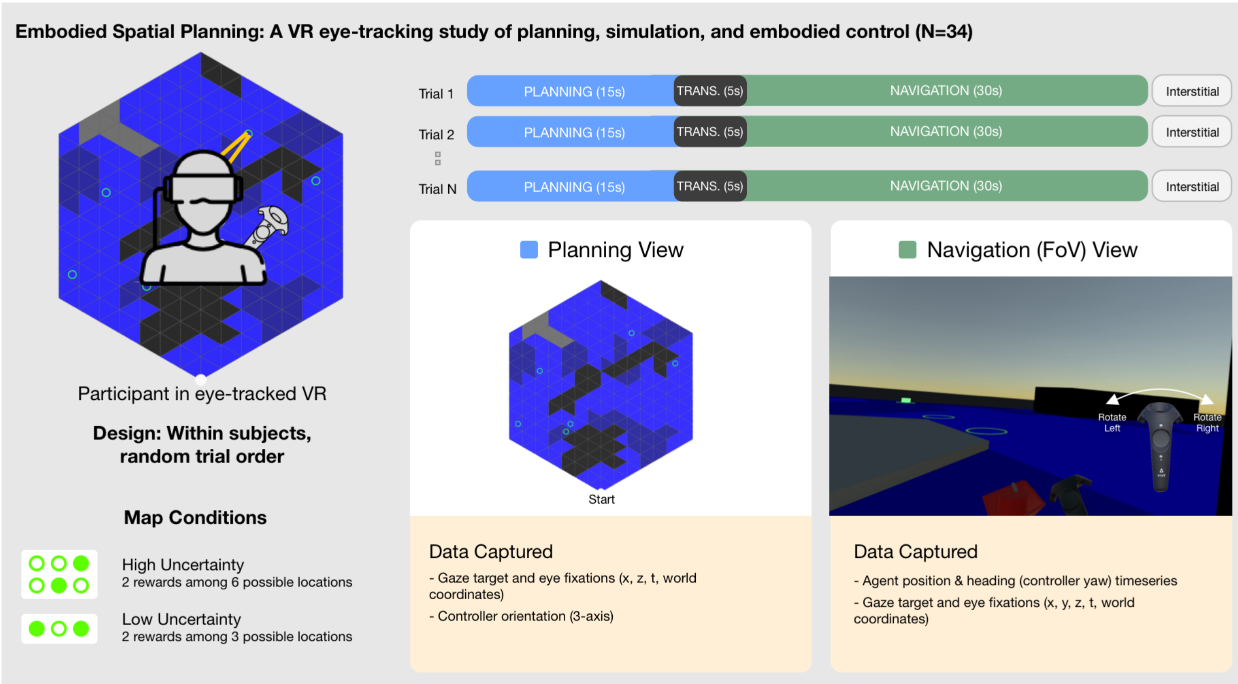


Figure 5.1: Experiment design. Each trial was divided between a planning phase, where participants viewed a top-down map of a novel spatial landscape and a navigation phase, where they navigated a vehicle through the landscape in immersive virtual reality. Map order was randomized, with each map containing either 3 or 6 goal-cache locations.

they are prone to motion sickness. Participants who are left handed or ambidextrous were also excluded in order to ensure standardized analysis of controller motion in the dominant hand. Participants were also required to be 18 years old or older, and speak English (native English was not a requirement).

Respondents qualifying for the study were then invited to participate in the main in-person experiment. 40 participants began the study, but several had to stop early due to motion sickness. In total, 34 participants completed the study, and form the sample for this analysis. Mean participant age was 21.3 ± 3.0. 75% were female, 21% were male, and 4% declined to state. 62% identified as Asian, 21% as White, 3% as Black or African American, and 8% identified as multiple races or Other.

5.2.2 Experimental design

The study used a single-condition within subjects design. Maps were procedurally generated (see Section A.1.1 for details), and balanced between high uncertainty (2 true goals among 6 possible goal locations), and low uncertainty (2 true goals among 3 possible goal locations). The same group of 33 maps was shown to all participants. The order of the first 30 (the procedurally generated set) was randomized uniformly. The final 3 hand-designed maps were shown at the end of the study, in the same order for all participants. See Figure 5.1 for an overview of study design.

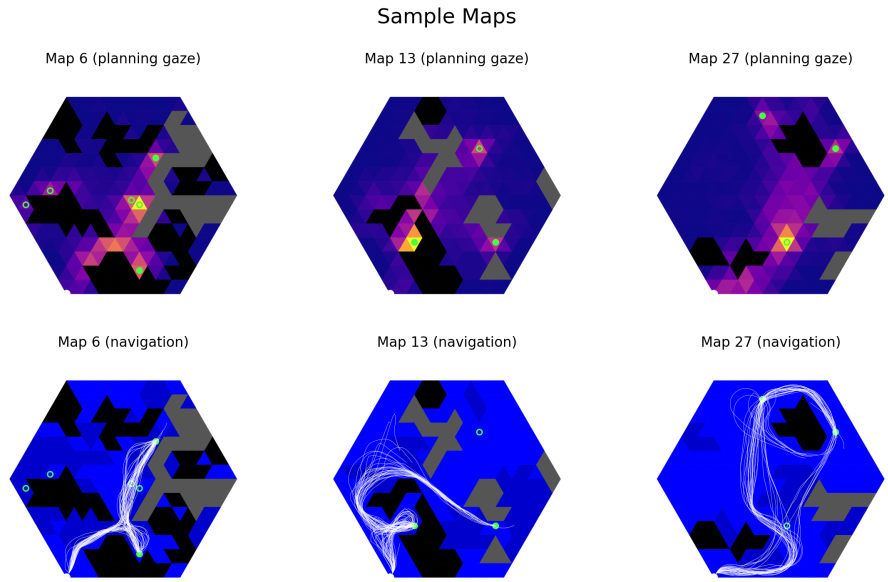


Figure 5.2: Sample maps from experiment. Top row shows map with aggregate planning gaze across participants. Bottom row shows navigated routes for each participant. Blue tiles indicate water (with darker blue indicating slower, "muddy water"), black tiles indicate walls (tall obstacles blocking visual access), and gray tiles indicate low barriers which prevent navigation but do not obstruct the participant's view during navigation. Green rings show the locations of goal caches, exactly two of which were populated at navigation time. The lower left vertex of the hexagon (indicated by a white circle) was the origin for all maps. Note that maps are rotated for convenience in visualization, but were presented to participants during planning such that the origin appeared centered at the bottom of the map.

5.2.3 Apparatus

An HTC Vive Pro Eye head mounted display (HMD), which was preconfigured with two Tobii eye-tracking cameras, was the primary apparatus supporting this study. The HMD was driven by an Alienware laptop running experiment scenes in Unity. A single HTC Vive controller was given to each participant to hold in the dominant hand. A "room-scale" setup with two lighthouses positioned on either side of the participant's seat was used.

5.2.4 Task & procedure

Prior to beginning the experiment, participants were briefed, signed an informed consent form, and were fitted with the VR headset. An eye-tracking calibration exercise was performed prior to starting the experiment. Then, a series of 3 practice trials began. In each, a simple map was shown to the participant, to ensure familiarity with the interaction and the task. I explained the interaction and task goals as the participant completed the three practice trials, answering any questions that were raised.

Following the three practice trials, the main experiment begins. Each trial is constituted by a unique map, and progresses in three phases. In the *planning phase* (15 seconds), participants see a top-down view of the landscape in front of them (see a few examples in

Figure 5.2). This map view shows the starting location of a boat—a player representation which the participant controls—at the bottom of the map, as well as the colored features of the landscape. Possible goal locations are shown as green rings. Tall obstacles ("walls") are shown as black tiles, low obstacles (which can be seen over but not passed through) are shown as gray tiles, and blue and dark blue tiles indicate water. During the planning phase, the participant is able to view the map, but cannot yet take any action. 3 chimes indicate the approach of the end of the planning phase. Next, in the *transition phase* (5 seconds), the map is removed and participants wait for the beginning of the next phase. In the third and final phase, the *navigation phase* (30 seconds), participants' view shifts and they are now situated in the landscape just behind the boat they control. They can look in any direction to explore the landscape from the agent's perspective. By rotating the controller around the vertical axis (yaw), the participant controls the direction of the boat. By holding down the trigger, the participant is able to move the boat forward in its present orientation. In this way, participants are able to navigate the landscape with the aim of reaching and collecting both goals in the map. If the boat contacts either type of obstacle, or falls off the edge of the map, the trial ends immediately. If both goals are collected, the trial ends early, several seconds after the second goal was collected. If both goals are not collected in time, the trial ends when the 30-second navigation timer is exhausted (which is again anticipated by a series of 3 chimes).

Following the end of each trial, an interstitial screen appears telling the participant that they can pull the trigger to proceed to the next trial. The study continues this way, until all trials are completed. At the end of the study, a screen appears providing feedback on the participant's performance. This screen summarizes the number of goals collected, as well as the bonus compensation earned based on the participant's performance. Base compensation was \$20. Bonus was added as follows. 0-20% of all possible rewards: \$0; 21-40%: \$1; 41-60%: \$2; 61-80%: \$3; 81-100%: \$5. After having read this information, I helped the participant remove the head-mounted display.

Finally, participants respond to a brief survey using a laptop. When finished, I released the participant to receive their total compensation, and exit the study room.

The experimental protocol was approved by the Committee for Protection of Human Subjects, under CPHS protocol number 2021-06-14394. Study methods and provisional hypotheses were registered and can be viewed on the study's OSF project page (`https://osf.io/5tacn/`). A video capture of the point of view of a sample participant can be downloaded from the "Participant POV" folder on the project page.

5.2.5 Analyses performed

Performance metric and other dependent variables

The simplest indicator of performance is the score obtained during the experiment (equal to the number of goals collected). Our analysis indicates that most participants devised planning strategies that were effective for the goal collection task. Mean score overall was 1.50 ± 0.63. Mean score per map varied, and in some maps, no participants successfully collected both goals. Hence, to account for this diversity, we defined a more fine-grained metric of performance that also takes into account map difficulty, and spatial progress towards the second goal when not collected (see Table 5.1 for details).

Based on this latter performance metric, we split participants into three equally sized percentile-based performance group bins (bottom, mid, and top-performers).

Furthermore, to study the strategies followed by participants, we defined additional dependent variables, which consider the sum of eye movements, and mean fixation duration (see Table 5.1).

DV	Definition
Performance	Computed based on goals collected (n) as proportion of optimal goal collection (n_{opt}), seconds remaining (τ), and path-progress from last goal (g_0) to remaining goal cache location (g_1). Formally: $performance = \frac{n}{n_{opt}} + \frac{\tau}{\tau_{opt}} + \frac{\lvert Path(g_0, agent_f)\rvert}{\lvert Path(g_0, g_1)\rvert}$. Optimal values are defined by highest performing individual participant.
Sum of eye movements	$\sum_i \lvert f_i - f_{i-1}\rvert$
Mean fixation duration	Mean duration of fixations, defined as gaze target remaining on the same object for over 50ms.

Table 5.1: Description of key dependent variables.

Salience of map features during planning

To identify which features of the map were considered to be more salient during the planning phase, we calculated an aggregate gaze heatmap for each map, across all participants, by summing tile-level gaze data during planning. Two methods of aggregation were explored: one based on fixation counts, and the other on total fixation duration. In both cases the resultant aggregate was standardized as an N-dimensional probability distribution (where N is the number of tiles in the base map). Notably, these aggregate heatmaps are sequence agnostic, and therefore are interpreted as a map of static visual salience during planning.

Then, the flattened gaze distributions were compared with a number of 10 "feature maps" that were calculated based on the spatial and informational geometry of each map. A description of each feature map tested is included in Table 5.2. Furthermore, Figure 5.3 provides example feature maps computed for map 2, as well as aggregated participant gaze on this map.

At the map level, I compared the distributional distance between each feature-map and the aggregate gaze distribution using the 2D Euclidean earth mover distance (EMD)[1]. The EMD provides a symmetrical measure of how far probability weights must be transported to match two distributions, and was chosen based on the intuition that gaze targets farther from predicted gaze density in 2D coordinates indicate a worse fit between the generated agent dynamics and empirical data. Feature maps with lower mean distance (across all maps) therefore are interpreted as supplying greater predictive power over preferred gaze locations.

[1]The POT python optimal transport library (Flamary et al., 2021) was used for EMD calculations.

Feature Map	Definition	Mean EMD
Line-of-sight (LoS)	Tiles valued by number of goals visible (linear access)	22.5
SP-LoS	Shortest-path tiles weighted by goals visible	17.5
Goal Distance (2)	Tiles valued as per distance to second-nearest goal	18.4
Short Paths	Shortest paths to each goal location	19.3
Goals	Goal locations and adjacent tiles	25.8
Path Through 2	Shortest path through origin and two goals	25.7
Path Through 3	Shortest path through origin and three goals	19.3
Critical (GD-Origin)	Change in goal-distance at origin (see "Goal Distance (2)" above) when tile and adjacent have passability toggled	23.8
Critical (GD-Map)	Change in sum of map-wide goal-distance when tile and adjacent have passability toggled	16.5
Critical (Shortest-Path)	Number of changes to short paths map when tile and adjacent have passability toggled	17.6

Table 5.2: Description of feature maps used in prediction of planning gaze. The third column shows the mean cross-map Earth Mover Distance (EMD) between the feature map and the map-aggregate gaze heatmaps. See Figure 5.7b for visual comparison of feature-map EMDs.

Temporal gaze analysis and RQA

In addition to considering "static" gaze distribution, I extracted a number of temporal features from planning-time gaze data, including time-binned total saccadic distance (which approximates the total amount of eye movement during a trial), distance from origin, distance from nearest goal, and a range of metrics derived from Recurrence Quantification Analysis (RQA).

RQA is an analytic method common in dynamical systems, and increasingly used to gain insights into temporal dynamics of eye tracking data (Tancredi, Abdu, Abrahamson, & Balasubramaniam, 2021). I generated recurrence plots from the sequential tile fixation data during the planning phase, and extracted an array of common RQA metrics measuring attributes of the timeseries such as the likelihood of a fixation being recurrent (recurrence rate, RR), the longest recurrent subsequence (L_{max}), and others. These were used as independent factors in the ANOVA analyses, to identify statistically significant differences correlated with both performance, as well as geometric map attributes.

An additional temporal analysis of fixations on hypothesized subgoal regions was also performed, see Section A.1.2 for a description, and Section A.2 for results.

Map analysis and classification

To address my research questions about the influence of information and uncertainty-related attributes of particular maps on planning processes, I ran ANOVAs on a number of gaze-related attributes as dependent variables, against a list of map features which were hypothesized to be relevant to navigation planning. These included the length of the shortest path

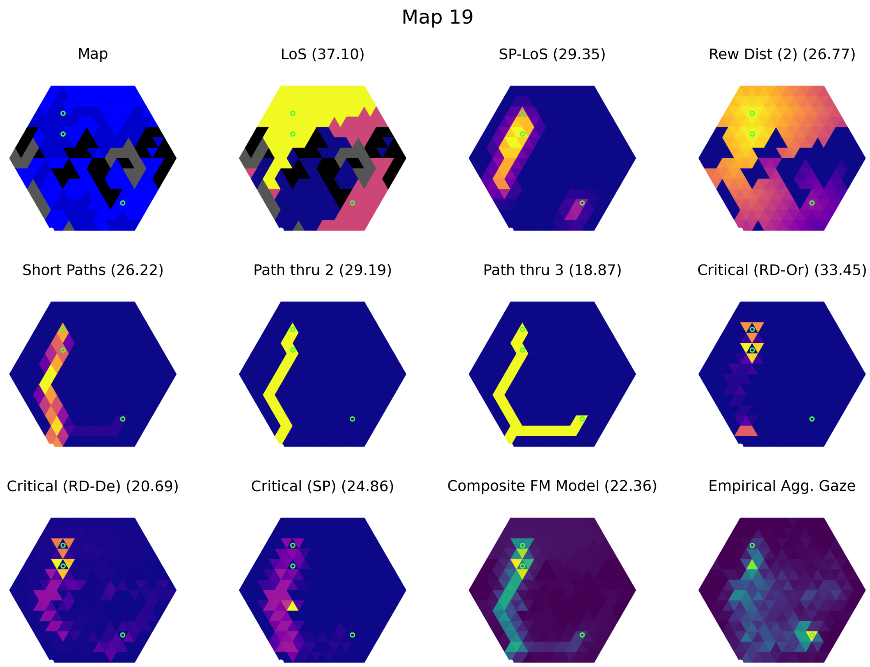


Figure 5.3: Sample feature maps for map 19. See Table 5.2 for details on each feature map. Numbers shown in each map title indicates aggregate gaze earth mover distance (EMD).

through 2, 3, and 4 goal cache locations (note that the last could only be computed for 6-cache maps), as well as a handful of metrics classified manually for each map which aimed to capture information access (e.g. where goal caches become visible in the landscape) and a simplified path-segment representation of each map. The coding process involved first producing a tree illustrating key segments and decision points, and then exhaustively considering goal cache observation outcomes (without knowledge of true goal locations).

I refer to the resultant representations as "info-graphical," given their tree like graphical structure, and sensitivity to visual information availability in the landscape. For an example, consider map 7 shown on the right side of Figure 5.4. At the origin (as with all maps by construction), no goal cache locations are visible. Participants must remember to navigate to the left, and by the time they reach about the midpoint of their journey to the narrow corridor at the top of the map, they will have visual access to two of the goal caches (square info nodes). At planning time, they contend with two outcomes: either one present goal is observed, or both are. Both outcomes result in full goal location knowledge, but in the first case (bottom fork in info-graph), after collecting the first goal they'll need to continue around the low obstacle in the center of the map to collect the second. The target goal during this journey is fully observed, but some memory of the map structure is required to remember which way to navigate around the obstacle, and hence this latter segment is marked as memory-dependent (three hatch marks). Alternatively, if both closer goals are present at the info node, they can continue around and collect both with no significant memory dependence (un-hatched edge).

After producing info-graphical representations for all maps, these were used to calculate

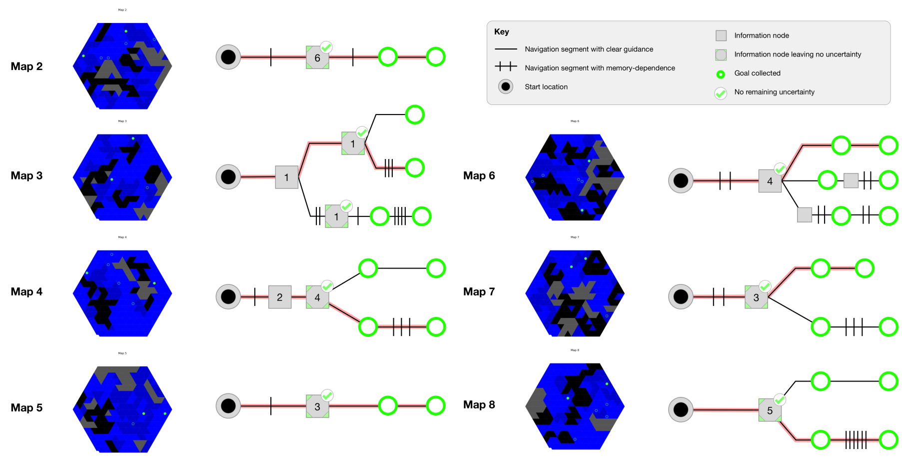


Figure 5.4: Map info-graphical classification for selected maps. See key for symbol semantics. Red line indicates expected path through graphical structure according to true goal locations.

a number of key metrics, including: branching factor, number of forks, steps to certainty, and expected number of memory-dependent segments. See Figure 5.4 for sample info-graphical representations of selected maps, and Table 5.3 for a full list of attributes computed.

Attribute	Definition
Branching factor	Average number of children of each node in the info-graphical path tree.
Number of forks	Count of forks in which a node splits into two or more children in the info-graphical path tree.
Number of information nodes	Number of informational nodes (represented as squares) in which goal-cache locations are hypothetically visible at navigation-time.
Steps to certainty	Expected number of nodes at which point all goal-cache presence statuses will have been observed.
Expected memory-dependent segments (planning)	Expected number of tree segments that require some memory of map structure to successfully navigate (indicated by hash marks through path segment).
Expected memory-dependent segments navigated	Expected number of tree segments to be navigated conditioned on true goal locations.

Table 5.3: Key info-graphical map attributes analyzed.

5.3 Results

5.3.1 Initial heading correlation with arm movements

To explore potential embodied dynamics during planning, I first asked whether arm movements during the planning phase are predictive of some of the very first decisions made—

which direction to begin navigating. For this analysis, I defined initial navigation direction based on the position of the participant in relation to the map center-line ("left" or "right"), 1 second after the beginning of navigation.

I find that, indeed, arm movements are directly correlated with initial heading. See controller yaw traces, segmented by initial navigation direction, and performance group, in Figure 5.5.

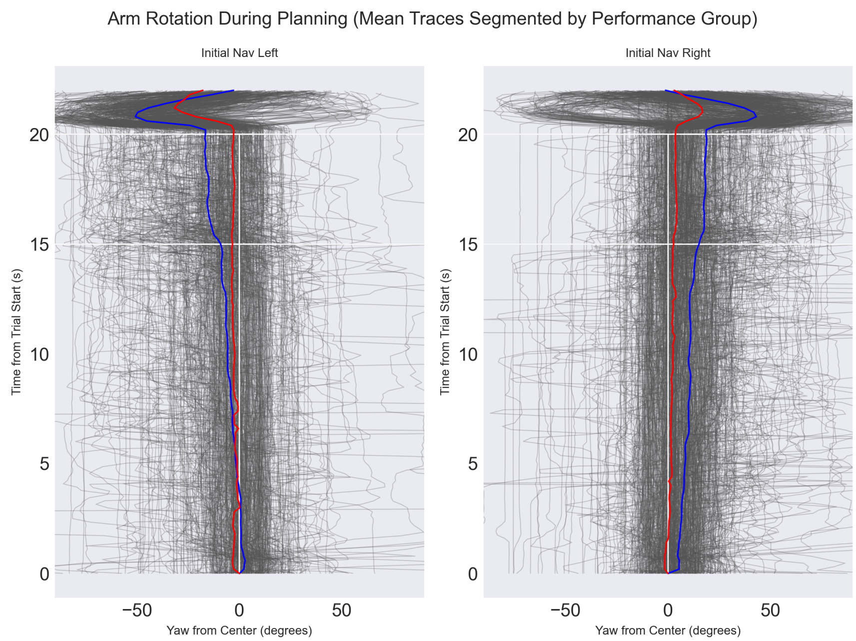


Figure 5.5: Arm movement traces as measured by controller yaw, segmented by initial navigation direction. Time from trial start progresses vertically from bottom. Navigation direction groups are defined as the sign of the offset from center-line 1000ms after navigation start. The blue and red lines show mean across trial-level traces among top and bottom-performers respectively. Planning phase (bottom 15 seconds), transition (following 5 seconds), and first 5 seconds of navigation phase are shown, separated by white lines.

5.3.2 Trends in gaze dynamics through planning

Next, I explored whether gaze dynamics show signatures of hierarchical planning and subgoaling. For this, I focus on temporal gaze patterns prior to movement.

Figure 5.6 shows that consistent trends in gaze dynamics can be found across the 15-second planning phase. On average, total saccade distance decreases while fixation duration increases. Fixation targets converge to tiles closer to the path ultimately taken, with minimum distance seen at the end of planning (see also Supplementary Figure A.9) for gaze dynamics in some example maps.

Additionally, a U-shaped trend is observed in distance from fixated tiles to the nearest goal location; tiles close to goals are more frequently fixated in the middle of the planning phase, while more distal tiles are viewed at the start and end of planning. Furthermore, These dynamics are largely consistent across top and bottom performers (top panels) and across 3 and 6-goal maps (bottom panels), but among top performers, gaze exhibits lower fixation durations and greater movement distances.

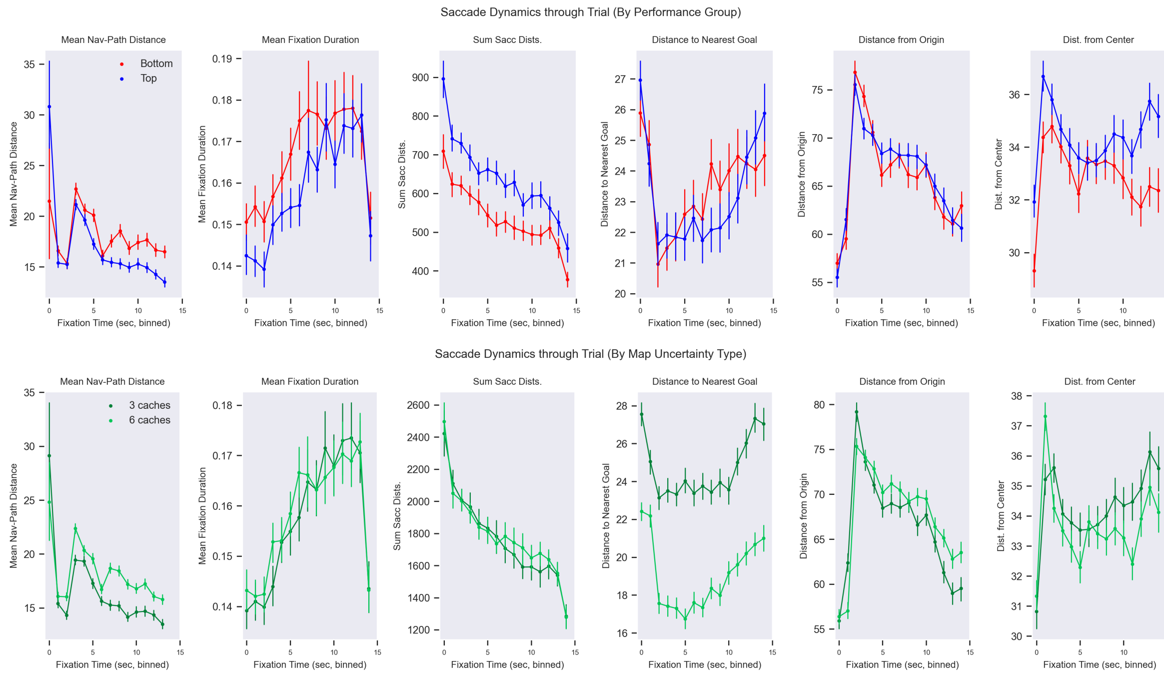


Figure 5.6: Trends in eye movement dynamics through planning phase. The six plots show the distance between gaze and navigation path, the saccade distance, saccade duration, distance from nearest goal, distance from origin, and distance from center. Top panel: comparison of top and bottom performers. Bottom panel: comparison of 3 and 6-goal maps.

Finally, when gaze metrics were analyzed against map attributes, results show a consistent relationship between key map info-graphical attributes and eye movements. Specifically, maps with more forks, a higher branch-factor, and more memory-dependent segments are associated with longer eye movements and reduced fixation duration during planning. Furthermore, I found associations between map attributes and planning-time gaze recurrence metrics. The map branch factor, number of memory-dependent segments, and number of forks were associated with lower recurrence rate (RR). These results may indicate wider novel exploration (that therefore exhibit fewer repetitions during search) in more complex maps. See Table 5.4 for a summary of ANOVA results.

5.3.3 Gaze distribution and map attribute salience

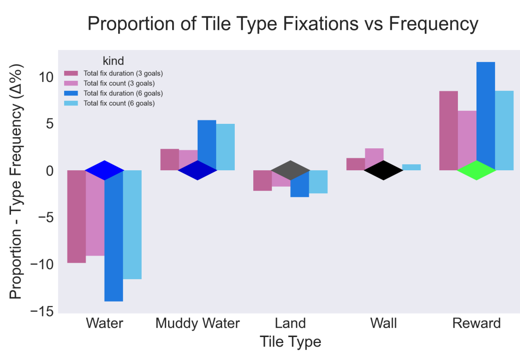
Another question of interest in this study is what features of the map are considered most salient during planning, and therefore attract the most visual attention. That is, I asked: how well do factors such as tile type, spatial and geometric factors, path-based, and other information-geometry-based attributes, predict the overall flattened (non-temporal) distribution of gaze at planning time?

I first considered the visual salience of map tile types through a static analysis of aggregate (cross-participant) gaze on each map. Figure 5.7a shows the fixation counts and duration

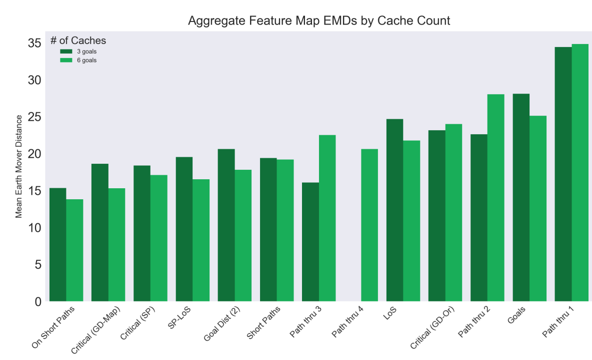
DV	Factor	DF	F	p	η_p^2	Dir.
Sum of eye movements	Branch factor	$F(2,66)$	9.86	$< .001$	.011	+
Sum of eye movements	No. forks	$F(2,66)$	13.3	$< .001$	.013	+
Sum of eye movements	No. memory-dependent segs.	$F(1,33)$	4.9	.033	.004	+
Mean fixation duration	Branch factor	$F(2,66)$	4.99	.010	.006	-
Mean fixation duration	No. forks	$F(2,66)$	6.47	.003	.006	-
Mean fixation duration	No. memory-dependent segs.	$F(1,33)$	10.2	.003	.004	-
Planning RR	No. forks	$F(2,66)$	8.2	$< .001$	.027	-
Planning RR	No. memory-dependent segs.	$F(1,33)$	22.6	$< .001$	.032	-
Planning RR	Branch factor	$F(2,66)$	4.1	.020	.015	-

Table 5.4: Summary of ANOVA results comparing gaze dynamics with map-level info-graphical metrics. "Dir" indicates whether the relationship is positive (+) or negative (-). Planning RR is gaze recurrence rate during the planning phase.

in relation to the five different tile types (water, muddy water, land, wall or goal-cache tiles). The value of zero in the plot indicates expected fixation duration and time if visual attention was perfectly uniform for a given map. This analysis indicates that both reward locations and muddy water are disproportionately fixated at planning time. Goal locations are obviously salient for the task, whereas muddy water (and especially unavoidable muddy water along a chosen path) might require more careful consideration. Furthermore, I found that regular water tiles were under-fixated by both metrics, likely due to regions of each map easily identified as not relevant to the navigation task, and therefore pruned early on. I also found that fixations on muddy water and goal tiles was even more disproportionately favored on 6-cache as compared to 3-cache maps. Finally, I observed decreasing attention on obstacles in the second half of the planning phase.



(a) Tile type fixation proportionality. The value of zero in the plot indicates expected fixation duration and time if visual attention was perfectly uniform for a given map.



(b) Feature-map predictors of aggregate gaze maps. Lower EMDs correspond to greater distributional similarity between feature map and aggregated participant gaze. Bars are plotted in ascending order from lower to higher distances.

Figure 5.7: Tile type attention proportionality changes and feature-map-based gaze prediction.

Figure 5.7b shows the fixation counts and duration in relation to the static feature-maps described in Section 5.2.5. I found that the feature maps were able to predict planning-time gaze in aggregate (across participants) significantly better than a uniform baseline. Feature-maps based on tile criticality, or the sensitivity of navigation paths and goal distances to the passability of particular locations, had lowest average EMD, implying they offer the highest predictive value (see Table 5.2 and Figure 5.7b). See also Supplementary Figure A.10 for illustrative examples of how feature maps reflect gaze allocation.

5.3.4 Path planning depth and performance

Finally, I explored how the planning strategy of participants prior to movement influenced their subsequent performance.

Included in the feature-map analysis of Section 5.2.5 were maps generated by finding the shortest path through both the origin and either 1, 2, 3 or 4 goal locations (SP-1, SP-2, SP-3, and SP-4 respectively). Using these maps, I inferred each participant's "planning depth" – or the putative depth of the plans that participants developed during the planning phase – on each trial, by choosing the SP-depth map with the minimum distance (EMD) to trial-level planning gaze.

I therefore analyzed the relationships between these four levels of planning depth (SP-1 to SP-4) and both performance and various map attributes. As shown in Figure 5.8, in three of the info-graphical attributes most directly correlated with map difficulty (number of forks, branch factor, and number of informational nodes), we see a relative agnosticism to depth class for simpler maps (e.g. those with less than 2 forks, or less than 3 info nodes). However, in the highest complexity group, a stronger performance correlation appears. Participants with planning depth most similar to SP-3 performed best on maps with the largest number of forks, while SP-4 participants performed best on maps with the largest number of info nodes, and largest branching factor.

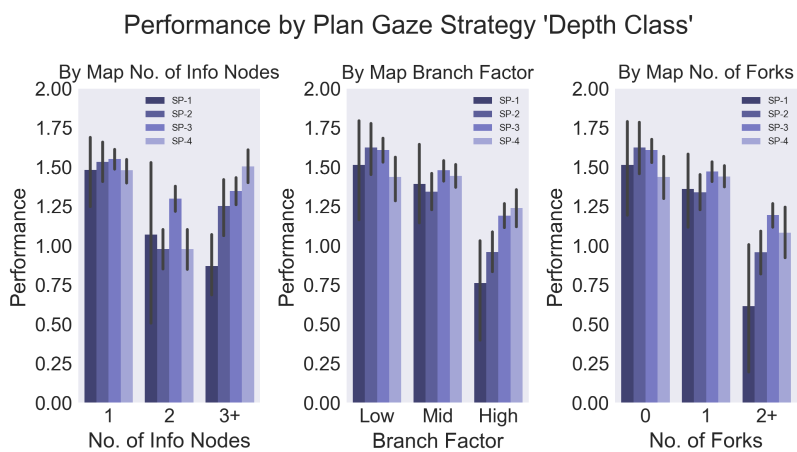


Figure 5.8: Trial-level performance comparison showing the relationship between planning depth (as inferred from planning time gaze distance from shortest-paths through 1, 2, 3, and 4 goal caches), and info-graphical map attributes. A positive, nearly monotonic, relationship is seen in the most complex maps (as determined by number of forks, branch factor, and number of info nodes) between planning depth and performance.

An additional perspective on planning gaze can be observed by comparison with true navigated paths. To explore this relationship, I first ran a validation analysis of planning gaze and visitation statistics on the level of individual tiles. Visit probability indexes the probability, across all trials and all maps, that a given tile falls on the navigated route. This metric is then related to fixation count (the number of times a tile was fixated during planning within a trial), and distance from origin. We find the expected negative relationship showing that tiles more distant from the start were less likely to be visited, but that fixated tiles had a higher visit probability. Fixation count was also predictive of visit probability, and fixation duration was negatively correlated with nearest visited tile (Pearson-$r = -.145$, $p < .001$). These results, in aggregate, support an expected relationship between tiles that attracted visual attention during planning, and the likelihood that each (or nearby neighbors) would be visited during navigation. See Figure 5.9 and its caption for details.

Following this validation, I explored the converse relation, that is, how close navigated routes were to tiles fixated during planning. Figure 5.10 compares the distance from the player's location to the nearest planning-time fixation at each second of navigation. I hypothesized both an increase in distance over time (as players navigate into regions farther from where they looked during planning), as well as potential performance-level differences. Both were observed, with larger distances (navigation farther from planned tiles) for lower performers, and generally increasing divergence from planned routes as navigation progresses. Interestingly, the increasing trend is somewhat inverted for top performers on 6-cache maps, where we observe a reduction of distance towards the end of navigation.

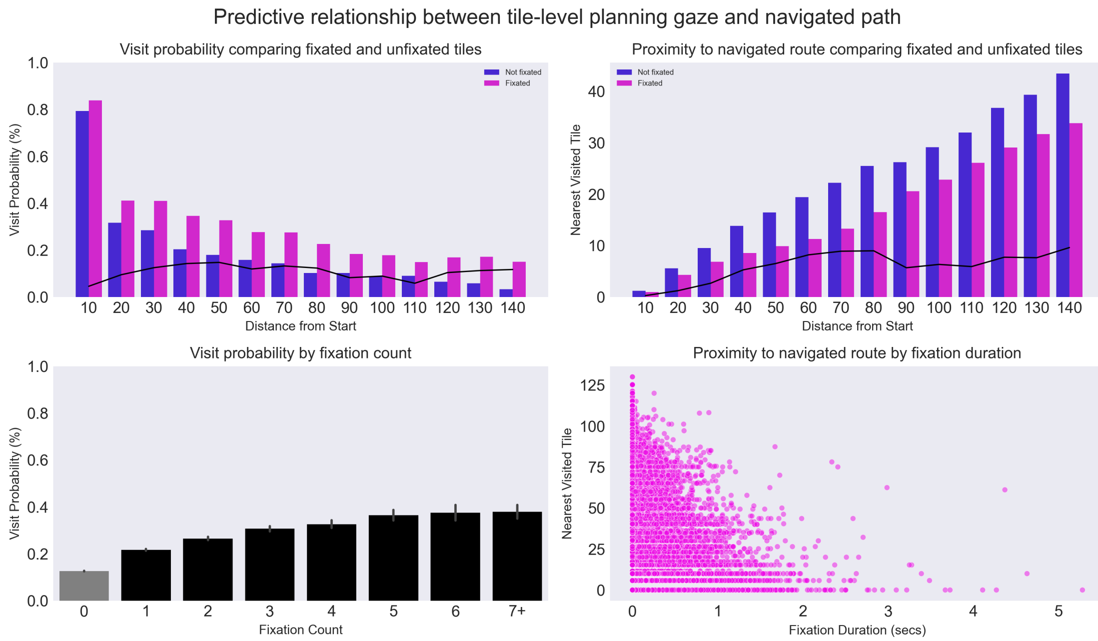


Figure 5.9: Plots relating tile-level planning time fixations with navigation decisions. Top left: visit probability decreases with distance from start (as expected due to the geometry of the task), with binary fixation status resulting in 10-20% increases in the probability that a given tile is visited during navigation. Top right: a similar (inverted) trend is seen in nearest visited tile, which is a more flexible measure of the planning gaze versus navigation relationship. In both top panels, the black line plots the difference between fixated and non-fixated tiles. Bottom left: visit probability versus fixation count shows a monotonically increasing trend, with more fixation instances correlating with increased probability of visit. Bottom right: a scatter plot showing a clear negative relationship between the duration a tile is fixated and the proximity to the navigated route. Longer durations were associated with closer proximity to the nearest tile in the visited path (Pearson-$r = -.145$, $p < .001$).

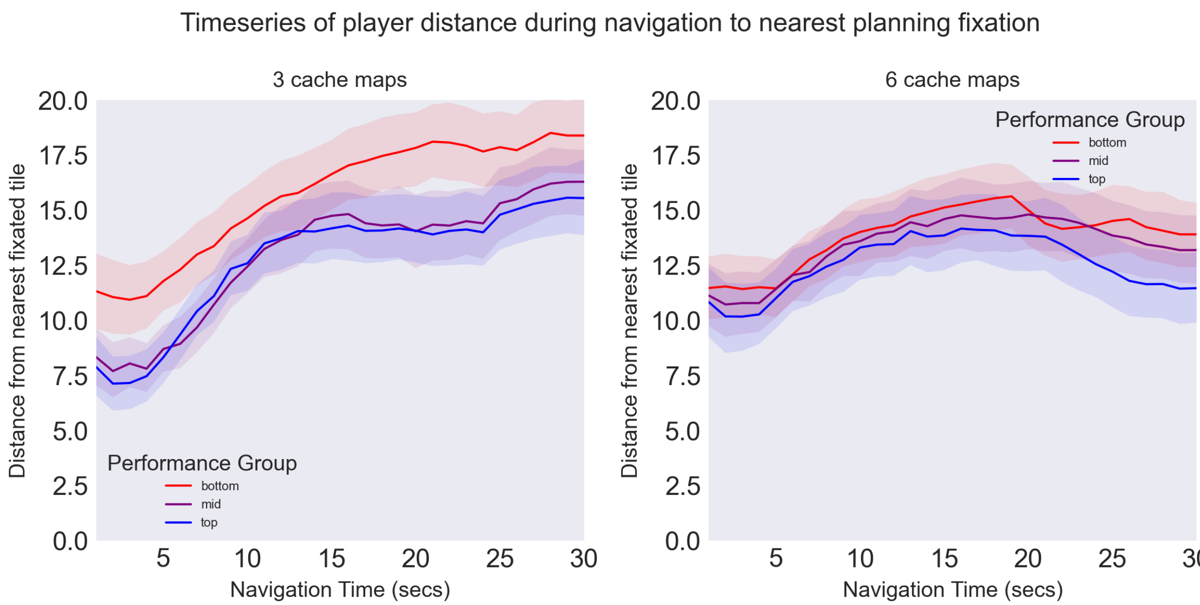


Figure 5.10: Timeseries during navigation of distance from nearest tile fixated during planning. The increasing trend seen in 3-cache maps (left), shows greater divergence from planned regions of the map during navigation, and consistently higher divergence among the worse performers (red). In contrast, top performers returned to planned locations towards the end of navigation on 6-cache maps, which may indicate more successful anticipation of terminal locations.

5.3.5 Other performance correlations

As a validation for the manual map info-graphical classification described in Section 5.2.5, I first confirmed that attributes hypothesized to index difficulty were correlated with trial-level performance. One way ANOVA shows a strong relationship between the number of memory-dependent segments and performance, with reduced performance on maps with a greater number of such segments ($F(1, 33) = 245.6, p < 0.001, \eta_p^2 = .356$), as well as a greater number of true memory-dependent segments during navigation ($F(2, 66) = 280.7, p < 0.001, \eta_p^2 = .673$)[2]; this finding is seen across both 3 and 6-cache maps. Similar results are seen between the number of forks and performance ($F(2, 66) = 131.6, p < 0.001, \eta_p^2 = .414$), as well as steps to certainty ($F(1, 33) = 103.6, p < 0.001, \eta_p^2 = .233$). Branching factor and number of information nodes were not predictive of performance, likely due to the indirect nature of their relationship with navigation-time outcomes; that is, while both planning behavior and navigation contributed to overall performance, map attributes effecting the difficulty of the path ultimately navigated were the strongest predictors of performance.

At the participant level, I ran linear regressions of mean planning-time gaze RQA metrics against mean (cross-trial) performance. I find a significant positive performance correlation with trapping time (TT, Pearson-$r = 0.36$, $p = .03$), and a corresponding negative correlation

[2]Note that actual memory-dependent segments, which were extracted by the length of the red line in info-graphical characterizations of each map (e.g. those shown in Figure 5.4), are an indication of the difficulty of segments that had to be navigated by participants following near-optimal navigation routes.

with recurrence rate (RR, Pearson-$r = -0.51$, $p < .001$).

I also find that side to side head movement, known as IdPhi and defined as $IdPhi = \int |unwrap(\dot{(\theta)})|$ as per (Santos-Pata & Verschure, 2018), both during planning (Pearson-$r = 0.61$, $p < .001$) and navigation (Pearson-$r = 0.53$, $p < .001$), was predictive of performance at the participant level.

5.4 Discussion

Planning as resource-rational optimal control It may be instructive to first consider what constitutes optimal behavior, as context for a discussion of empirical findings. I take the *resource rationality* view of optimality as suggested by Callaway et al. (2018), that is, that people make rational use of their limited cognitive resources in planning tasks, and *satisfice* by executing the best good-enough solution they find (Simon, 1956).

In this human behavioral study, participants were asked to view an overview of a landscape for 15 seconds, prior to immersively navigating through it to find and collect up to two goals. Because the number of possible goal locations ("caches") exceeds the number of true goals, successful participants had to develop a navigation plan robust to this uncertainty. A risk-minimizing planner might identify a path through the origin and all goal cache locations (to ensure both true goal locations are visited in any goal configuration). In this case, the problem is equivalent to the well-studied Hamiltonian path problem, which is NP-complete (Itai, Papadimitriou, & Szwarcfiter, 1982), and would require consideration of $n!$ paths (where n is the number of goal caches) to exhaustively compute. In contrast, the task studied here differs from formal graph-based pathfinding due to the visual geometry afforded by immersive spatial navigation through an open environment. Specifically, any given position in the map affords visual access to other locations (subject to the constraints of visual obstructions and orientation), thus allowing participants to verify true goal locations without visiting each cache directly. As a result, solution of the given task could plausibly be supported by simulation—the internal generation of possible sensory trajectories produced by candidate action sequences. Time constraints imposed on the planning phase introduce an additional limitation on the number of sequences that can be considered. Given these task dynamics, what constitutes a resource rational output of planning within this structure?

If we frame planning prior to navigation as the simulation-driven generation of a control policy for a dynamical system (ones own anticipated movements through the environment), we can cast the visual search process during planning as a problem of optimal control. Optimal control theories have been leveraged to explain simpler tasks like visuomotor coordination and control of a stochastic system (e.g. bouncing a ball on a racket Schaal, Mohajerian, and Ijspeert (2007); Schaal, Atkeson, and Sternad (1996)), but can also be applied to navigation in uncertain environments (e.g. to support robust robot navigation (Bansal, Tolani, Gupta, Malik, & Tomlin, 2020)). Applied to the task studied here, the time-varying affordance landscape constructed during planning is seen as a candidate control law, and the optimality criterion can be approximated by aggregating returns across multiple simulated trajectories conditioned on the present plan. However, since the control law developed during planning can only be leveraged in the future (after some delay), working memory constraints must also be considered. That is, plans must balance not only robustness to uncertainty (inclusion of multiple possible scenarios), and expressiveness (e.g. guidance through key visual milestones), but also memorability.

A primary motivation in this design is to better understand and characterize the dynamics of the internal representations developed during planning, and the handling of uncertainty during this process. However, these representations were not, themselves, directly measurable. Instead, I analyzed various biometric data streams collected during the planning phase of the task (namely eye, arm and head movements), which putatively reflect some dimensions of the control policy of interest, and enable quantitative comparison with an artificial agent's behavioral outputs.

Overall, the spatial and temporal analyses of data collected in this study, as well as correlations between gaze patterns and quantitative map-level attributes and task performance metrics, help to illustrate a planning process that uses a multi-level approach that is both hierarchical and bi-directional. Signatures of hierarchical planning in gaze dynamics likely aided participants in efficiently transforming rich spatial information into a simpler, contingency-aware, plan of action. I also find additional embodied dynamics leveraged in preparation for action, which exhibit behavioral divergence between top and bottom performers. Each of these main findings is discussed in the sections that follow.

5.4.1 Hierarchical planning

Temporal gaze dynamics start broad and narrow as plans are established In a temporal analysis of gaze dynamics, I find consistent trends indicative of hierarchical planning across the planning phase (see Figure 5.6). Total saccade distance decreases while (in line with this result) fixation duration increases over time. Fixation targets converge to tiles closer to the path ultimately taken, with minimum distance seen at the end of planning. A U-shaped trend is observed in distance from fixated tiles to the nearest goal location; tiles close to goals are more frequently fixated in the middle of the planning phase, while more distal tiles are viewed at the start and end of planning. Finally, gaze distance from origin shows a quick radiation outwards in the first several seconds, followed by a consistent negative trend for the remainder of the planning phase. Together, these results illustrate a visually guided planning process that exhibits a time-evolving hierarchical nature. Visual exploration initially favors proximal locations, and begins wide and fast, exhibiting saccades that move between easily-identified salient locations such as goal caches and other subgoals. Gaze then narrows (in terms of spatial spread), and deepens (seen as longer fixation durations) through the planning phase as a navigation plan is converged upon. Finally, performance analyses show that among the most successful participants, gaze exhibits lower fixation durations and greater movement distances, attributes correlated with a higher throughput of information processing at planning time.

Results from recurrence quantification show greater trapping time (TT), and reduced recurrence rate (RR), among higher performers, which might indicate the value of deeper visual inspection (seen as longer vertical lines in recurrence plots measured by TT), as well as a larger working memory capacity (which might reduce the need for repeated gaze sequences).

Salient tiles, subgoal locations, and critical regions predict attention during planning Findings from map-based spatial prediction of gaze preferences suggest that visual search strategies are highly sensitive to not only goal locations, obstacles and other gross aspects of spatial geometry, but also to aspects of graphical connectivity and information geometry (e.g. goal presence information available at a particular tile).

Overall, goal locations and muddy water were most disproportionately fixated. I observe increasing attention on water, and decreasing attention on obstacles in the second half of the planning phase (see Figure 5.7a). This finding may indicate that early saccades to obstacles support the development of a simplified graphical representation of map connectivity (and information access), i.e. pruning, while later gaze sequences support evaluation of candidate paths through the structure identified.

Gaze prediction analysis using static geometry-based feature-maps enabled a more nuanced analysis of the types of features that attracted visual attention. Feature-maps were effective predictors of planning-time gaze in aggregate (across participants). Comparison of predictive power for each feature-map analyzed is suggestive of the landscape features that most consistently attracted planning-time visual attention (see Table 5.2 and Figure 5.7b). Specifically, feature-maps based on tile criticality, which identify regions for which a change to connectivity (i.e. switching a passable region for an obstacle) most significantly modify navigation considerations such as the lengths of goal-paths (Critical Shortest-Path), or mean goal distance map-wide (Critical GD-Map), were especially effective predictors. This finding is consistent with results described as edge "toggling" in (Zhu et al., 2022).

Deeper planning gaze predicts success on high-uncertainty maps Planning depth results (see Section 5.3.4) showed that the SP-3 map was the most predictive of aggregate gaze across participants (on both 3- and 6-cache maps), indicating a possible balance of (cognitively costly) planning depth and risk of planned caches being empty.

Performance correlations with gaze depth-class (as determined by similarity with "SP-X" feature maps) indicate that some map attributes rendered deeper plans particularly helpful. In particular, planning gaze similar to SP-3 was associated with highest performance on maps with at least one fork. Similarly, maps with fewer information-nodes (locations at which goal existence information would be obtained) saw greatest performance when gaze resembled SP-3. In contrast, SP-4 performed best in maps with the largest number (3 or more) informational nodes.

See Figure 5.8 for comparisons, and Section 5.2.5 for details on the determination of map classification attributes. Another perspective on this result can also be seen in the performance-binned gaze preferences in Figure A.9, which show a bias for more peripheral, and later path segments among the top performing group.

I interpret these findings as evidence that deeper plans predicted performance, but only on more challenging maps where multiple branches, or multiple stages of information access, required the development of a more robust plan visiting more goal caches. On the other hand, a greedier (and riskier) strategy was satisfactory on maps with simpler information geometry and less path uncertainty.

Correspondence between planning-time attention and navigated routes predicts performance Results on the distance between navigated routes and the nearest tiles viewed at planning-time (as seen in Figure 5.10) illustrates a finding in line with a number of prior studies indicating that more effective gaze-based forecasting (as measured by predictive eye movements) go with increased performance and greater expertise (Huang, Velarde, Ma, & Baldassano, 2023). Note that this result is mirrored in Figure 5.6 (upper left), which shows a performance-group gap in distance to navigated path throughout the general convergence seen during the planning phase. Together, these findings highlight that visual attention during planning more directly predicted areas of the map that would be visited during navigation, and that successful (visual) anticipation of eventual navigation choices predicted success on the task.

The reduction in distance in 6-cache maps seen in Figure 5.10, and especially clear among top-performers may indicate that higher-performing participants did a better job of anticipating deeper path segments likely to arise towards the end of navigation. One interpretation might suggest this is evidence of greater hierarchical planning in high performers, where path segments towards the end of the route were successfully predicted, while earlier intermediary locations were more likely to be skipped over. However, it is not clear why a corresponding trend is not seen in 3-cache maps, where we instead observe a divergence from planned locations towards the end of navigation. Perhaps the higher certainty (fewer possible reward configurations), and sparser distribution of the 3-cache maps led participants to overlook some possible goal configurations, causing them to more often traverse parts of the map they hadn't inspected during planning. In this case, the the explicitly greater ambiguity in 6-cache maps may have encouraged more flexible planning since it was clear the true configuration of rewards couldn't be well predicted.

5.4.2 Embodied dynamics

Though arm movements (measured by the controller) were used primarily as an interface for navigation, I collected arm position and orientation during planning and transition phases as well. I hypothesized that arm movements prior to navigation might be predictive of early navigational decisions[3], if anticipatory movements indicative of action simulation, or offloading were observed. I find that, indeed, arm movements are directly correlated with initial navigation heading; see controller yaw traces, segmented by initial navigation direction, in Figure 5.5. I interpret the increased directional bias seen just following the end of planning (beginning of transition phase) as action preparation that may also indicate the use of decision-offloading, that is, storage of the initial navigation plan in the orientation of the arm so as to reduce cognitive memory demands. I also note that mean yaw begins to diverge from neutrality within the first 5-10 seconds of planning, indicating that arm movement is predictive of the ultimate direction choice even in the early stages of planning.

Interestingly, directional bias is significantly greater among top-performers (blue mean trace) as compared to bottom-performers (red mean trace). This difference may indicate that effective participants were more likely to exhibit embodied offloading (in line with Kirsh and Maglio (1994)), or perhaps anticipated their planned initial rotation to improve navigational efficiency.

[3]Note that controller orientation had no effect on the agent's heading until the beginning of the navigation phase (at the moment when the participant is immersed in the landscape, and begins observing the agent in perspective). As such, controller movements during planning are not inherently connected to navigation heading.

5.4.3 Limitations & future work

In this study, uncertainty was primarily manipulated by requiring participants plan with incomplete information about true goal locations. While this paradigm corresponds with certain naturalistic settings where rewarding states often fall within a set of discrete locations, other types of uncertainty are frequent in real-world situations requiring planning and navigation. As one example, path connectivity may differ from maps or other representations available when a route is being selected, e.g. due to unforeseen obstacles, the presence and influence of other individuals (e.g. crowds that slow one's progress), and other modifications to the environment (e.g. a transit line discovered to be down for maintenance). These environmental uncertainties likely require even more robust contingency planning, and perhaps prompt the development of different internal representations better suited to more continuous and dynamic changes. Such a design would likely limit the efficacy of heuristic pathfinding strategies such as "find a short path through 4 goal caches," likely used by some participants in this study, since path length would itself be unpredictable during planning. Eye movements and other biometric data would likely shed light on both correspondence and divergence in the planning processes used by individuals encountering these other types of uncertainty.

Similarly, while the plan-then-execute paradigm employed here reflects common daily scenarios where planning may progress prior to direct engagement in the environment, as discussed in (Gordon, Maselli, et al., 2021), more embodied planning and action scenarios require the constant interplay of perception and action as the state of an individual with respect to their environment continuously evolves. The meta-cognitive and resource-rational aspects of how to coordinate movement with sensory sampling, and when and how to leverage deeper simulations for more distal planning processes, even during ongoing activity, is not well understood; these questions offer many rich avenues for further study.

5.5 Conclusion

This work presented a virtual-reality based behavioral experiment which collected eye movements during a pre-navigation planning phase, with the aim of capturing mental simulation and reasoning processes related to pathfinding through a landscape with uncertain goal locations. I highlighted various ways the body, both in the form of arm movements and eye fixations at planning time, are predictive of navigation performance, and indicative of a search process that exhibits hierarchical dynamics. These include a decreasing trend seen in gaze distance from origin, a broad to narrow shift (with reducing saccade distances and longer fixation durations) as plans are established. In line with prior work, critical tiles to which landscape connectivity is most sensitive were the strongest predictors of visual attention. I also find that deeper planning was correlated with success only on the most complex maps (e.g. those with a larger number of information-nodes, higher branching factor, and more forks, according to an info-graphical map analysis).

In the next chapter, I will present work shifting attention beyond planning as a behavior of individuals, and turn to the joint action paradigm. In this important domain of collaborative problem solving, I will explore the ways simulation and counterfactual reasoning likely play a role in the selection of strategies to decompose a task into pieces that lend themselves to independent completion.

Chapter 6

Shared prospective representations in collaborative planning (Study C)[1]

6.1 Introduction

We humans often engage in cooperative tasks, such as collecting a list of items at the market (Pacherie, 2008; Rand & Nowak, 2013; Sebanz et al., 2006; Sebanz & Knoblich, 2009; van der Wel, Becchio, Curioni, & Wolf, 2021). Such tasks require the coordination of actions and plans of two (or more) people for prolonged periods of time and can be decomposed into multiple subtasks (such as grouping vegetables and cheeses found in different parts of the market) that can be executed—at least in part—in parallel by different people, resulting in improved efficiency through cooperation (H. H. Clark, 1996; Henrich, 2017; Tomasello, 2014; Tomasello, Carpenter, Call, Behne, & Moll, 2005; S. A. Wu et al., 2021). Such a decomposition requires decisions about how, specifically, to divide the work, i.e., who does what and when. Collaborators who devise and coordinate on good strategies to split their work can increase the benefits of cooperation (e.g., not buying duplicate items).

Prior work on collective decision-making and action coordination has illuminated some of the approaches used by collaborators working towards shared goals (see Section 2.6.1). Additionally, work in computational cognitive science has attempted to explore whether agents endowed with a human-inspired theory-of-mind are better able to effectively share work in complex tasks containing a hierarchy of sub-tasks to complete successfully. One recent example is a model called Bayesian Delegation, trained to solve a joint cooking task, which infers which sub-tasks collaborators are working on, and then chooses whether to help (cooperate), or divide and conquer (work independently) (S. A. Wu et al., 2021). This line of work highlights the fact that improving our understanding of human collaboration techniques is not only valuable in its own right, but potentially critical to developing safe and effective AI and robotics applications intended to work in conjunction with human collaborators (a key objective in the fields of HRI and more broadly HCI).

Despite this progress, and the fundamental importance of division of labor strategies to human cooperation, we lack a good understanding of how individuals split work during

[1] This chapter is based on work published in Cognitive Science under the title "Strategic task decomposition in joint action" (Gordon, Knoblich, & Pezzulo, 2023)

cooperative planning (Braun, Ortega, & Wolpert, 2009; Candidi et al., 2015; Chackochan & Sanguineti, 2019; Chouhan & Niyogi, 2017; Clodic, Pacherie, Alami, & Chatila, 2017; Curioni et al., 2022; Kourtis, Sebanz, & Knoblich, 2013; Pezzulo, Candidi, et al., 2013; Pezzulo & Dindo, 2011; Shum, Kleiman-Weiner, Littman, & Tenenbaum, 2019; S. A. Wu et al., 2021). To successfully decompose a problem, group members could benefit from agreeing upon a single shared strategy and implementing it in a consistent way. Importantly, there are costs and benefits associated with strategy selection and maintenance. Accordingly, the establishment cost of selecting an appropriate strategy—which I define as the cumulative cognitive resources required to jointly discover, communicate, and enact a strategy—can be significant. At the same time, successful strategies alleviate the cognitive cost of monitoring the others' actions (the cognitive resources unavailable to perform a subtask due to attention on a partner's actions), by enabling collaborators to independently complete their own subtasks. Notably for the present work, these monitoring costs are directly related to the amount of information required about the activity of a collaborator in order to perform one's own subtask successfully — that is, the extent to which participants' subtasks are coupled. Hence, in planning a joint strategy, collaborators must weigh the expected establishment cost of forming a strategy against the cost of performing the task without it (i.e., "cognitive monitoring" cost). Which parameters affect this balance in a joint decision-making process is not fully understood (Ang, 2021; Meyer, van der Wel, & Hunnius, 2013; Sacheli, Arcangeli, & Paulesu, 2018; Török, Pomiechowska, Csibra, & Sebanz, 2019).

To test what strategies people use to divide work my collaborators and I investigated (1) how collaborators balance the establishment and monitoring costs of decomposition strategies, (2) under which conditions they maintain existing strategies when cost-benefits change, (3) whether they adapt their strategies when they lead to an unfair allocation of effort, and (4) how strategy selection influences overall task performance. To address these questions, new joint planning task was designed in which two players navigated in a shared grid world to collect "gems" of two different colors in a limited amount of time. For each trial, the dyad was assigned a joint goal (e.g., collect exactly 6 red gems and 2 blue gems) and was free to decide how to accomplish the task, without being able to communicate verbally. Players could only move to connected nodes and they had to press a button to collect a gem at their current location.

In a basic (non-strategic) approach, this task can be completed by monitoring another player's movements and tracking the quantity of gems of each color collected so far, to meet the joint goal of collecting an exact number of gems of each color. However, I hypothesized that dyads would develop a "color polarization" strategy to decompose the whole set of gems into subsets of gems of the same color. Crucially, forming such a strategy substantially reduces the cognitive monitoring costs but incurs some establishment cost. To study the trade-offs implicit in strategy formation and maintenance, I manipulated three experimental variables (see Figure 6.1).

First, I varied the monitoring costs of joint performance by introducing two different color counters. In the low monitoring cost (LMC) condition, participants could monitor progress towards the joint goal with the aid of a color-specific progress counter that showed the number and color of gems collected so far. In the high monitoring cost (HMC) condition, a similar progress counter (shown in gray) tracked only the number of gems collected but not their colors. The presence of the LMC counter should significantly decrease cognitive

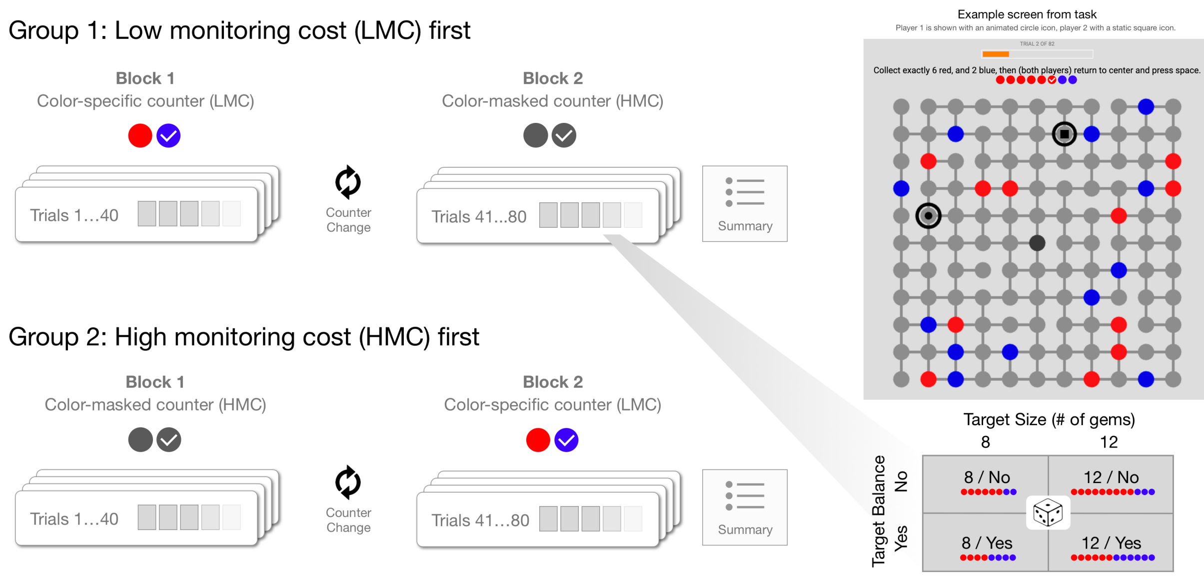


Figure 6.1: Experimental design. A between-subjects factor manipulated which counter type each dyad saw first (the grey "high monitoring cost" counter, or the colored "low monitoring cost" counter). Additionally, a 2x2 within-subjects manipulation randomized both the target size (8 or 12 gems total), and the target's color balance (either balanced with an even number or red and blue gems, or not balanced with 75% of gems red).

monitoring costs due to the shared, color-specific informational aid displaying current task needs (e.g. "I can pick up this blue gem since we still need 2 more"). With the non-color specific HMC counter, players' subtasks were tightly coupled, and, therefore, required costly continuous monitoring ("my teammate is about to pick up the last blue gem, so I'll look for a red"). Dyads were randomized into two groups: group 1 saw the LMC block first before switching to HMC trials, and group 2 saw the HMC block first before switching to LMC trials. Comparing the behavior of each group during block 1, I expected fewer dyads in the LMC block to invest in establishing a task decomposition strategy since the task was feasible without it. Due to high monitoring costs HMC dyads should be more motivated to establish such a strategy, e.g., to split the task by color.

Second, each group of dyads were shown a second block, following the first, using the alternate counter type. I hypothesized that if color polarization had already been established in the first block, and it remained effective at reducing cognitive costs under the new dynamics of the second block's task, then it would remain in place, effectively becoming a "convention" (Centola & Baronchelli, 2015; Galantucci & Garrod, 2011; Hechter & Opp, 2001; Young, 1993). Therefore, HMC-first dyads should conventionalize their previously developed strategy into the second block, despite the inherent reduction of monitoring costs due to the LMC counter. LMC-first dyads would need to develop a color polarization strategy once confronted with the more cognitively demanding HMC block.

Third, a potential imbalance of effort was introduced by creating trials in which the joint target was to collect either the same number (balanced, e.g., 4 red, 4 blue) or an imbalanced number of gems of each color (imbalanced, e.g., 6 red, 2 blue). Note that color imbalanced trials were always imbalanced in the same direction, i.e. there were no trials with a goal to collect more blue than red gems. I hypothesized that in trials with an unequal allocation, dyads who had developed color polarization would loosen this strategy to increase fairness, especially in the LMC condition where monitoring costs are low.

6.2 Methods

6.2.1 Map generation

Each of the 80 unique maps was generated by the following process. An 11 by 11 grid of nodes was created. Nodes could be connected to adjacent neighbours in each of the four cardinal directions. To produce more interesting map connectivity, a p-coin (with $p_{connected} = 0.9$) determined whether to connect each pair of adjacent nodes. To ensure that all nodes could be visited, I confirmed that the resultant graph was composed of a single connected component.

To ensure gems were well distributed across each map, and not too close to the central starting location, the grid was divided into four 4x4 quadrants (containing 25 nodes each), with centres at coordinates (0-indexed x & y): $[2, 2]$, $[2, 8]$, $[8, 2]$ and $[8, 8]$. Finally, 12 gems of each color were placed in random order, such that exactly 6 gems would appear in each quadrant, and both the color distribution, and location of gems within each quadrant, would be fully random. A 3x3 middle region, and each of the 4 corner nodes were also disallowed for gem placement. Finally, I assigned each generated, valid, map a task condition, based on the two by two by two design: number of gems to collect (8 or 12), color balance of target gems (even or 75% red), and counter type (easy monitoring, colored, or difficult monitoring, grey). While this task condition assignment influenced the dyad's goal for a trial, as communicated by the in-game task message, as well as appearance of the counter, the map generation process was consistent across all conditions (i.e. all shared the same size, connectivity, gem color balance, and randomized placement process for gems). Note that map design was fixed, such that all dyads saw the same set of trials in pseudo-randomized order based on which counter group they were assigned to. See Figure 6.1 for a schematic of the experimental design.

6.2.2 Recruitment & study procedure

Participants were recruited using Prolific (www.prolific.co) between August and October of 2021. The only exclusion criteria were that participants be of at least 18 years of age, and list English as a primary language. A sample of 24 dyads (48 participants) completed the study during the data collection period, with an equal number randomly assigned to each between-subjects condition (12 in the easy monitoring condition and 12 in the difficult monitoring condition). This sample size is consistent with prior work on joint action coordination, which typically looks for large effect sizes (Curioni, Vesper, Knoblich, & Sebanz, 2019; Konvalinka, Sebanz, & Knoblich, 2023; Vesper et al., 2013). No dyads completing the experiment were excluded. The mean age was 26.2 ± 6.1. 60.4% of participants identified as male, and 37.5% identified as female.

A study coordinator scheduled each dyad for a particular time window allowing us to match each dyad through the web interface, and initiate the two-player experiment.

In preparation for the experiment, both members of each dyad were instructed to visit the experiment website independently to complete the online consent form, respond to basic demographic questions, and read detailed instructions introducing the game dynamics (including navigating their player around the map, collecting gems, monitoring progress using the counter, and then finishing each trial). A unique dyad code was provided to participants, identifying each participant's dyad and randomized study condition. As a result, the instructions could be customized based on which block the dyad would see first. Instructions included only information about the type of color that would appear in the first block, to avoid biasing strategic choices based on knowledge of the eventual counter-based manipulation. Following the instructions was an invitation to try out several independent (one player) practice trials, to ensure participants entered the main experiment with sufficient familiarity with the game controls and dynamics. The practice trials, like the instructions, were customized based on the first counter block to be seen in the main experiment.

At the scheduled time, both members of each dyad signed into the study website from independent locations. When both players were present, they were automatically matched and the main two-player study began. On each trial, each participant used the keyboard to navigate a single, shared map for up to 30 seconds. The positions of both players were visible on the map at all times (See Figure 6.1 and Figure 6.8). Each player was assigned a unique visual marker, and the player's own position was animated to clearly differentiate it from their collaborator's marker. A trial completed when either a) both players returned to the starting location (map centre) and pressed space, b) the trial's goal was exceeded (too many red or blue gems were collected), or c) the 30-second trial timer expired. If successful, the dyad received points for each trial equal to the number of seconds remaining in the timer. If unsuccessful (i.e. either color target was not exactly met), no points were awarded. Upon completion of the first block of 40 trials, an interstitial was presented to both members of the dyad explaining that in the next block, the counter's presentation would be changing, but that the task would otherwise remain the same. This interstitial was customized to describe and show the changed appearance of the counter, based on the dyad's block order. After acknowledging the information in the interstitial, and completing a single practice round with the new counter, the experiment resumed through the remaining 40 trials in the second block.

At the end of the experiment, a summary page was presented to the dyad, reporting the dyad's score on each trial, and cumulative final score. A unique confirmation code was also generated for each participant, allowing him or her to confirm the completion of the study to the coordinator. All participants who completed the study were compensated 10 euros for their participation.

6.2.3 Data collected

The web-based experiment allowed telemetry to be collected and stored in an online database, which constituted the primary dataset for analysis. Telemetry included: the timestamp and parameters of each player's movements, and gem collection actions, and the results of each trial (number of each gem color collected by each participant, trial success, and points received). This data allowed us to reconstruct a precise trajectory of each player's movements, and therefore absolute and relative spatial locations, which was necessary for various analyses.

6.2.4 Dependent variables

The key dependent variables for timeseries and mixed ANOVA analyses were defined as follows.

color polarization: $\frac{1}{2}[\frac{max(G_(R,1),G_(B,1)}{G_1} + \frac{max(G_(R,2),G_(B,2)}{G_2}]$, where $G_(C,P)$ is the single-trial count of all collected gems with color C collected by player P. This trial-level metric is most easily interpreted as the average percent of gems collected by each player in their preferred color. Range: 0.5 when each player collects 50% of the gems of each color, 1.0 when each player collects 100% of the gems of one color and 0% of the gems of the other color.

Work balance ("fairness"): $\frac{1}{2}[1 - |\gamma_1 - \gamma_2|]$, where γ_P is the decimal percentage of all collected gems, on a given trial, collected by player P. Range: 0 when a single player collects 100% of gems and 0.5 when each player collects 50

Success rate: This is the percentage of trials completed successfully (in which dyads collected the exact target number of gems of each color). Range: 0 when no trials are completed successfully, and 1.0 when all trials are completed successfully.

Trial score was also analyzed. Results were in line with trends and group differences, and are reported in the Supplementary materials (see table 6.5 and Figure 6.9).

6.3 Results

For the main analysis, I devised two high-level variables that index key aspects of the behavioral patterns and strategies expected: color strategy ("color polarization") and fairness ("work balance"). For definitions of each of these dependent variables, see the Methods section. I additionally considered the "success rate" (i.e., percent of successful trials) as a dependent variable. Figure 6.2 visualizes the dynamics of these three variables, as they evolve throughout the course of the experiment.

Analysis used a mixed ANOVA with group (i.e. block order: high monitoring first "HMC-first" vs. low monitoring first "LMC-first") as between-subjects factor, and task color balance (balanced vs. not balanced, i.e. more red) and counter type (high monitoring cost "HMC" or low monitoring cost "LMC") as within-subjects factors. The ANOVA has model DF of 1 and error DF of 22 for all results. ANOVA analyses were followed up with pairwise t-test post hoc tests when ANOVA showed significant interactions at $p < 0.05$.

Table 6.1 shows the results of the analysis of the "color polarization" variable, which indexes the choice of a color strategy to split the task. This variable takes a minimum at 0.5 when both players take half of each color, and a maximum of 1.0 when each player collects only one color. I find a main effect of counter type ($F(1, 22) = 13.4, p < 0.001, \eta_p^2 = .378$) showing increased split by color in the HMC trials, which was expected given the challenges of monitoring without a convention. This main effect can be seen especially clearly in comparison of block 1 alone (prior to any effects of task switching), as shown in the bottom-left panel in Figure 6.3 ($T = -6.1, p - uncorrected < 0.001, \eta_p^2 = .163, BayesFactor > 100$). Comparison within the second block alone shows a weaker effect in the same direction

($T = -2.3, p_u = .021, \eta_p^2 = .027, BF = 1.9$). Additionally, an interaction between group and counter type ($F(1,22) = 5.7, p = 0.026, \eta_p^2 = .206$) is observed, showing that participants seeing the LMC counter first were significantly less likely to exhibit color polarization during the first block of LMC trials, but increased their color polarization during the second block of HMC trials ($T = -3.6, p_u = .004, \eta_p^2 = .323, BF = 11.86$).

Table 6.1: color polarization: an index of use of a color-based strategy for splitting work. Means are reported, with standard deviation shown in parentheses. These results are plotted in Figure 6.1.

Block Order	Balanced LMC	Imbalanced LMC	Balanced HMC	Imbalanced HMC
LMC-first	0.762 (.120)	0.784 (.114)	0.928 (.087)	0.867 (.143)
HMC-first	0.873 (.158)	0.833 (.143)	0.878 (.135)	0.880 (.129)

Table 2 shows the results of the analysis of the “work balance” variable, which indexes fairness. This variable takes a minimum of 0.0 when one player collects all gems, and a maximum of 0.5 when each player collects exactly half the gems. A main effect of counter is observed ($F(1,22) = 7.2, p = 0.014, \eta_p^2 = .246$), with a greater work balance for LMC; a main effect of target balance ($F(1,22) = 33.2, p < 0.001, \eta_p^2 = .602$), with greater fairness when target colors are balanced; and an interaction ($F(1,22) = 7.3, p = 0.013, \eta_p^2 = .249$), i.e., effect of target imbalance is larger for HMC trials where splitting colors is the main strategy ($T = 3.9, p_u < .001, BF = 50.35$).

Table 6.2: Work balance: an index of fairness in the split of the work at every trial. These results are plotted in Figure 6.3.

Block Order	Balanced LMC	Imbalanced LMC	Balanced HMC	Imbalanced HMC
LMC-first	0.420 (.049)	0.376 (.051)	0.419 (.121)	0.306 (.074)
HMC-first	0.446 (.050)	0.348 (.075)	0.429 (.070)	0.297 (.048)

Table 3 shows the results of the analysis of the “success rate” variable. There are two 2-way interaction effects, and one 3-way effect. An interaction between counter type and group ($F(1,22) = 7.7, p = 0.011, \eta_p^2 = .260$) indicates that participants seeing the high monitoring cost block first were more successful during the low monitoring cost block ($T = 2.97, p_u = .013, BayesBF = 4.87$). An interaction between target color balance and group ($F(1,22) = 5.4, p = 0.029, \eta_p^2 = .198$) shows that the HMC-first group succeeded more often in imbalanced trials, while the LMC-first group succeeded more often in balanced trials. Finally, a three way interaction of group, counter type, and color balance ($F(1,22) = 8.4, p = 0.008, \eta_p^2 = .276$) shows that the HMC-first group struggled most with balanced trials in the HMC block.

Finally, as an additional way to evaluate the efficacy of developing a color polarization strategy, I analyzed the relationship between dyad-level success rate and mean color polarization using linear regression (see Figure 6.4). When segmented by counter type, a positive relationship is found in both trial types (Pearson $r = .41$, $p = .045$, LMC trials, $r = .48$,

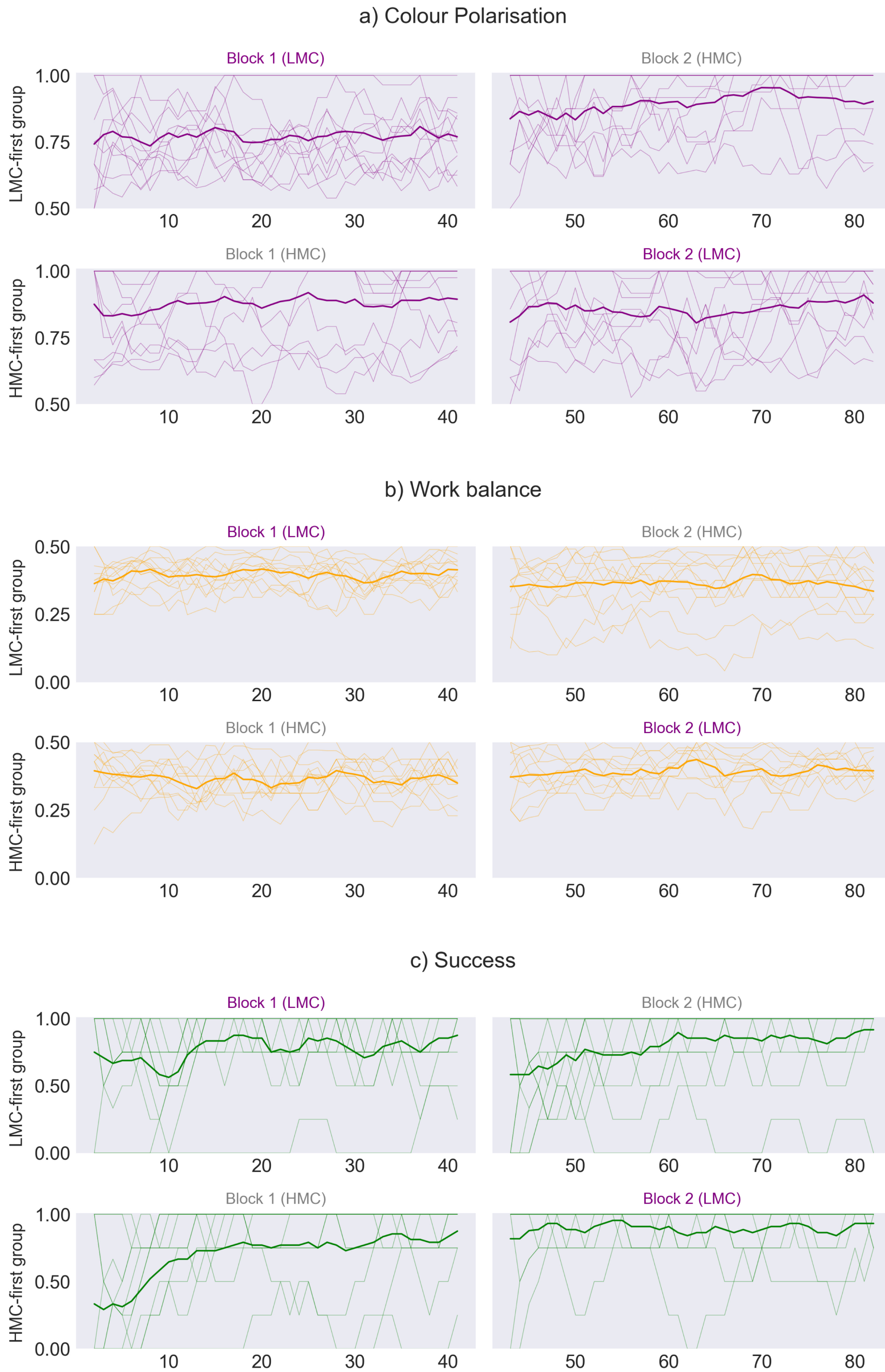


Figure 6.2: Timeseries plots of three key dependent variables, split into blocks. Each 2x2 panel shows a complete trial progression (1-80) for LMC-first group (top row) and HMC-first group (bottom row). Variables in each panel are as follows: a) color polarization, b) work balance or "fairness", and c) success rate. The plots show both the average time series (in bold) and the time series trace for each dyad. For simplicity, these figures are not further segmented by the color-balance condition; these results are reported in Tables 1-3.

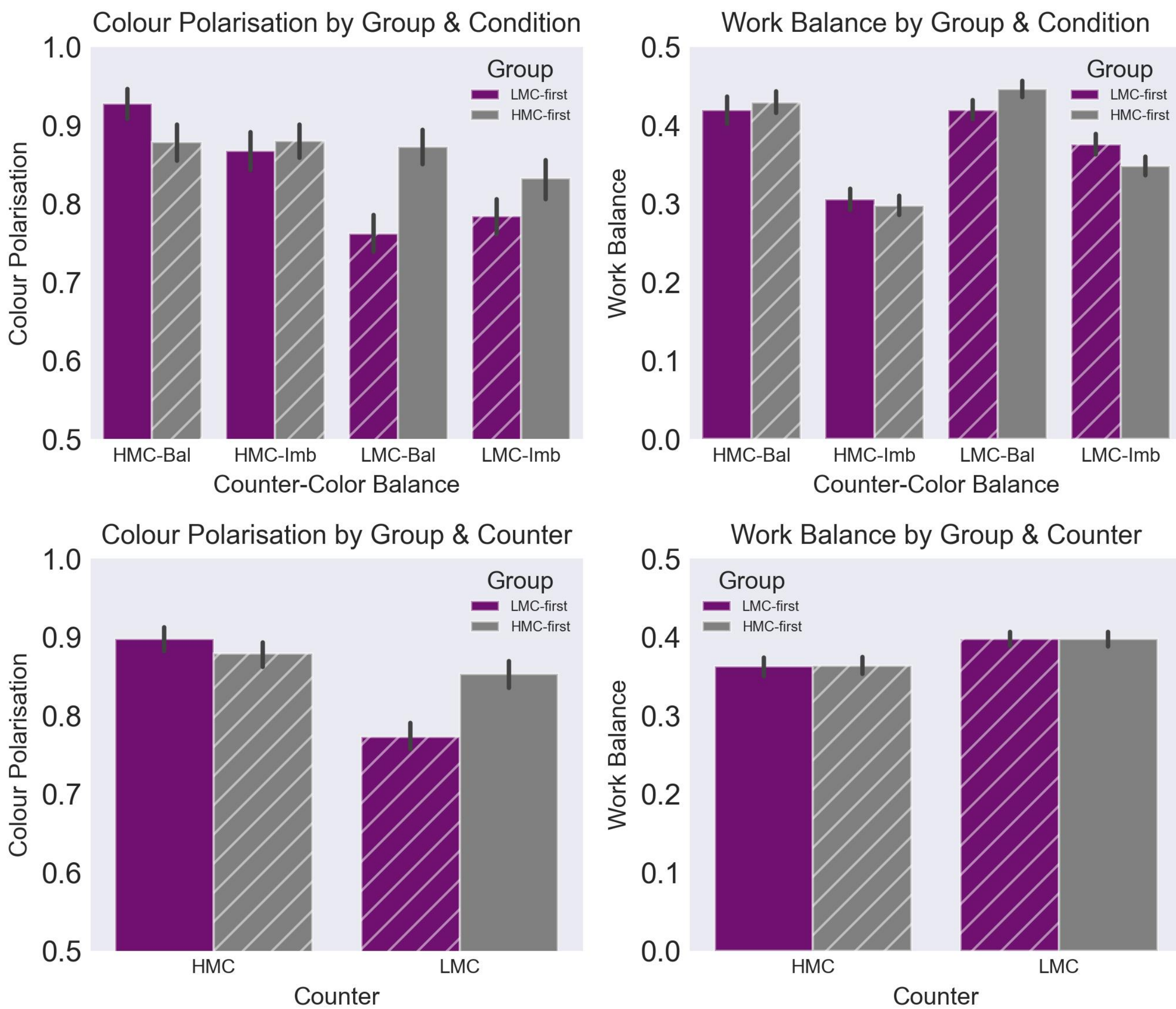


Figure 6.3: Plots of selected ANOVA results. Upper-left: color polarization by group and condition. Upper-right: work balance by group and condition. Lower-left: color polarization by counter type only. Lower-right: work balance by counter type only. The LMC-first group is plotted in purple, and the HMC-first group in gray. Bars are grouped as per condition (trial counter, and trial color balance for the top row, and counter type only for the bottom row). Angled hash marks indicate trials within the first block. Note that axis limits differ and are based on the range of each dependent variable.

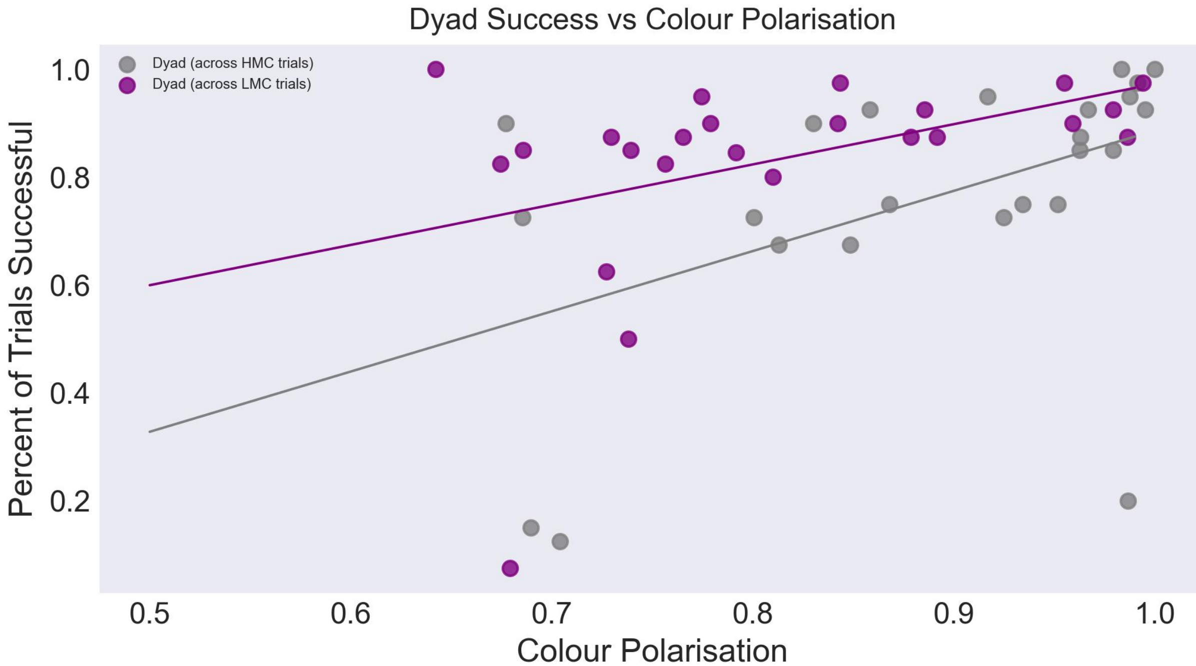


Figure 6.4: Dyad-level mean color polarization was predictive of success across all trials. The figure shows a dyad-level scatter plot of mean trial success against mean color polarization segmented by trial type (dyad mean within HMC trials in gray, and within LMC trials in purple). I find a stronger positive relationship in HMC trials, where the color polarization strategy was especially effective. Linear regression shows a slope of 0.75 (Pearson $r = 0.412$, $p = 0.045$) for LMC trials, and a slope of 1.12 (Pearson $r = 0.481$, $p = 0.017$) for HMC trials.

p = .017, HMC trials), but a steeper regression coefficient among HMC trials (plotted in gray).

6.4 Discussion

In this study, I investigated how collaborators, when decomposing a joint task, balance the costs of establishing a joint strategy of "color polarization," and the expected benefits of its execution. To understand this balance, I manipulated task parameters in such a way as to "tilt" this trade-off, and then measured the effect on dyads' behavioral and strategic responses.

Table 6.3: Success rate: an index of task performance calculated as the percent of trials for which the exact target number of blue and red gems was collected.

Block Order	Balanced LMC	Imbalanced LMC	Balanced HMC	Imbalanced HMC
LMC-first	0.771 (.227)	0.779 (.243)	0.838 (.209)	0.771 (.262)
HMC-first	0.888 (.115)	0.899 (.165)	0.679 (.270)	0.758 (.300)

As pointed out by (S. A. Wu et al., 2021), successful strategy selection in such a joint task requires an understanding of the future actions and intentions of a partner. It can be seen, then, as a joint planning exercise, and as such, may entail the prospective mental simulation (akin to the motor simulation posited in (Vesper et al., 2013)) of various counterfactual strategy choices (Klein & Crandall, 1995). As an example, let's consider a dyad having established a strong color polarization strategy (participant A takes red, and participant B takes blue) during the first, HMC, block. Confronted with a balanced trial (collect 4 of each color), participant A notices that three of the blue gems their partner will likely pursue are clustered, but the others are positioned far away, and several are in regions easily collected by A's likely path. A may well evaluate a decision to break from their color-based work division by collecting a blue gem, and thus earning additional points for the dyad through time savings (and thereby establish a new, more lenient joint strategy). Evaluation of this decision may be supported by mental simulation (prior to commitment) of plausible trajectories for each through the map. Perhaps B will notice their (strategy breaking) collection of the final blue gem, and continue home with the dyad's target achieved. Alternatively, B may be distracted by their own navigation and gem collection, and confident in the efficacy of the color strategy used to date, and move to collect an additional blue gem, thus overshooting the target and failing the trial. The probability of these two outcomes (and perhaps others) may be approximated during such counterfactual simulation (Kahneman & Tversky, 1981b), as well as the anticipated cognitive costs (monitoring B's movements) and value of each outcome. In this example, value may be assessed as the additional points earned for time savings in the first outcome, and the zero value of a failure in the second. Of course, due to the repeated nature of this task, such simulation-based reasoning might well be replaced, or complemented by, a more on-line approach (try a policy change, and check on the resultant joint performance), enabling strategy adaptation through experience across trials.

In the present work, four main novel findings are presented, consistent with the evaluation and balancing of expected costs and benefits when changing strategy. First, task demands can alter the trade-off between the execution costs of a candidate strategy, and the expected costs of its establishment. Second, these establishment costs create an asymmetry between groups at the transition between blocks: I observe persistence (of the established strategy) for one group, and adaptation (to develop a more effective strategy) for the other. Third, considerations of equal allocation of work were present but difficult to prioritize when in conflict with overall task performance. Fourth, "color polarization" is found to be beneficial for task performance. I discuss these four findings in order.

First, though the task is simple and could be potentially solved using a strategy that is insensitive to task manipulations (e.g., to collect every gem in the surrounding until the counter is filled in), the unique establishment-monitoring cost trade-off confronted by each group in the first block drove differences in strategy selection. Specifically, in the first block, the LMC group primarily satisfied the task's demands with a simple "collect-while-monitoring" behavior and only a loose tendency towards color preferences. The LMC counter made it possible for this group to avoid the costs of establishing a strict convention. In contrast, the HMC group, contending with the large cognitive costs of monitoring and remembering task progress with the less specific counter, exhibited stronger use of a color polarization strategy to render the task feasible. Though establishment costs were significant (e.g. see the delayed rise of mean task success at the start of the first block, as well as several individual dyad examples of early failures prior to convention establishment in Figure 6.7), once adopted, this strategy allowed dyads to "decouple" the task in a way that significantly lowers monitoring costs and increases success rate (see 6.4). Overall, these findings illustrate that task demands significantly affect strategy selection. It is worth noting, however, that within these trends there is substantial variation both in dyad-level strategies, as well as individual behaviors (see the Supplementary material in Section 6.6 for analysis and a brief discussion).

Second, I find a history-dependence of strategy selection when tasks dynamics changed. This implies an asymmetry in either the expected establishment cost, cognitive monitoring cost, or both. Indeed, participants in the HMC-first condition largely reused the color polarization strategy that they had developed in the first block through the second, despite the changed demands. Their behavior was markedly different from participants who encountered the LMC block first, who did not develop strong color polarization. The color polarization of the HMC-first condition could therefore be considered a "convention" that was maintained even if no longer strictly necessary under the new LMC task dynamics (i.e. akin to a carryover effect of strategy, but which influences joint rather than individual behavior in this cooperative task paradigm). In principle, these dyads could have benefitted from switching to a more liberal strategy that did not include a strict color assignment, or a strategy permitting a fairer allocation of work, but such considerations did not justify the costs of switching to a new convention.

Participants in the LMC-first group, who initially avoided establishment of a strict color polarization strategy, faced a different trade-off. Satisfactory performance could no longer be achieved within a tightly coupled teammate-monitoring strategy, and thus a strategy shift, despite the costs of establishing it, was necessary. The significant drop in success following the transition for this group illustrates the cost of a weak strategy (resulting in highly coupled actions), under the HMC task dynamics. Like their counterparts encountering the HMC counter in block 1, the majority of dyads in this group established a strict color polarization strategy, despite their success without it in the prior block. Having said this, a strict color polarization strategy was not truly mandatory to solve the task. Indeed, several dyads used spatial preferences as an alternative to (or augmentation of) the more common color polarization (for an analysis of this spatial polarization see the Supplementary materials).

Notably, by the second half of the HMC block, the LMC-first group exhibits the highest rate of color polarization seen. These results imply that the cost of establishing a strict color polarization strategy was significant, and that dyads were only willing to make this investment after seeing they couldn't be successful without it. This result may also indicate that establishing an initial strategy is less costly than moving away from an already established convention. A carryover of old strategies to novel situations could be common to both individual and joint action setups. However, while in individual settings the carryover could be linked to the cost of devising a novel strategy (Lancia et al., 2023; Todorov, 2009), in joint action settings it could be additionally due to (or amplified by) the communication costs inherent to strategy change. Distinguishing between these possibilities is an intriguing question for future experimental study.

Third, work balance (fairness) was lower when target colors were imbalanced and during the most difficult (HMC) blocks. Furthermore, fairness was lowest in color-imbalanced HMC trials, in which strict color polarization does not allow an equal split. In these cases, it seems that considerations about the effectiveness of the strategy prevail over fair work distribution. Previous studies of fairness during joint action allocation reported inconsistent findings about the prioritization of efficiency and fairness. One study reported findings consistent with those presented here, suggesting that efficiency is prioritized over fairness during joint action allocation (Strachan & Török, 2020). However, another study reported that planning during joint action was more difficult and took longer time if the selected strategy implied an unequal allocation of effort between partners (Ang, 2021), suggesting that people's perception of fairness of work allocation influences decisions about joint plans. Furthermore, fairness could motivate the selection of apparently counterproductive actions, such as those that alter others' incomes at a personal cost (Dawes, Fowler, Johnson, McElreath, & Smirnov, 2007), suggesting that people care about fairness despite its potential costs (Fehr & Schmidt, 1999). Finally, considerations of fairness could help strategy selection; for example, the could help group members coordinating their decisions tacitly in conditions of uncertainty (Schelling, 1980; van Dijk & Wilke, 1996). A deeper exploration is still needed of how collaborators handle trade-offs between the fairness of individual contributions to a jointly performed task and the cognitive limitations related to its planning and performance.

Finally, the costs and benefits of selecting "color polarization" are well reflected in performance measures (see Figure 6.2c). While this result might seem trivial, it is important to remark that participants following a color polarization strategy must sometimes skip gems that are close to them but have the wrong color – which might potentially slow down task completion. Despite this, in the presented task, the benefits of the strategy (in terms of lowering monitoring costs and the uncertainty about which gems to collect) surpass the costs of skipping gems. In the first block, participants who saw the LMC counter exhibited improved performance, whereas participants who saw the HMC counter performed poorly in the first half of the block, prior to developing a strong color polarization strategy. Across blocks, the performance of the HMC-first group remained stable, as the color convention established by this group remained effective. In contrast, the performance of the LMC-first group dropped significantly at the beginning of the second block; this drop of performance was likely a key driver of the strategy shift across blocks. Overall, the most successful block was the HMC-first group during their second block, which is the only block that benefited from an already learned (color polarization) strategy. Interestingly, color polarization was predictive of success in both HMC and LMC trials, with a stronger relationship seen among HMC trials (see 6.4).

6.5 Conclusion

In sum, findings reported in this chapter suggest that subtask allocation strategies are modulated by costs and benefits of joint actions (Curioni et al., 2022; Pezzulo et al., 2018; Sebanz et al., 2006) and that strict decomposition strategies emerge more often when simpler (though cognitively costly) behaviors fail to satisfy task demands. However, once a certain decomposition strategy has been adopted, it continues to be used even in cases where simpler behaviors could satisfy task demands. This raises the interesting possibility that repeated interactions result in an increased use of decomposition strategies across a range of task and cost-benefit parameters that is stable enough to be culturally transmitted (Sperber &

Hirschfeld, 2004).

An open question is how such a stability trades off with the necessity to change individual and joint behavior in a continuously changing social environment (Becchio, Sartori, & Castiello, 2010; Pezzulo et al., 2018; Tomasello, 2014). Understanding how the human mind deals with such trade-offs might be a critical ingredient in future multi-agent AI systems that have to face complex problems of joint planning and coordination (Belhassein et al., 2022; Castelfranchi & Falcone, 2010; P. R. Cohen & Levesque, 1991; Donnarumma et al., 2017; Pezzulo, Donnarumma, & Dindo, 2013; Ramirez & Geffner, 2010; Tambe, 1997). Another open question regards the generalization of findings about decomposition strategies to situations in which participants can communicate explicitly (e.g., linguistically). Converging on a specific decomposition strategy (e.g., color polarization) would be plausibly much simpler if the agreement can be reached verbally, but explicit communication could be redundant when the strategies to be followed are already established or obvious. Future studies are needed to assess under which conditions people prefer using implicit or explicit forms of communication to allocate subtasks.

This chapter presented results from an experimental paradigm designed to better understand cooperation and coordination strategies when two individuals take on a task with a joint goal. The presented manipulation, by design, enabled dyads to decompose and thereby decouple each individual's actions, and to do so within a vaguely embodied, and fully non-verbal, paradigm in which overt actions (in this case movements) were the only available communication channel. But a variety of embodied signals that can, in some circumstances, convey intentions and other internal states of mind, have been explored in in previous chapters. In the next chapter, I consider a variety of applications for communication strategies, and the alignment of representations of the future in general, which might benefit from a more nuanced conception of the body's role in cognition.

6.6 Supplementary materials

6.6.1 Spatial preferences

While the task design promoted a fairly intuitive single strategy of color polarization, in general dyads could develop other strategies that are either alternative to color polarization or operate in parallel with it. As an example of the latter, dyads are observed mixing multiple decomposition strategies even within a single trial (e.g., players developing a strong spatial preference to collect gems to the north on every trial, while maintaining a loose preference for red; see the Supplementary materials for an illustration of the spatial strategy). Given the task design, a pure spatial strategy to collect gems in different regions of the maze fails to fully decouple each player's actions, and therefore contends with significant monitoring costs. However, mixing in spatial preferences may have contributed to improved motor efficiency and conflict avoidance, without discarding the primary joint strategy. To further analyze the dynamics of the observed spatial preferences I defined a fourth dependent variable to index this trend.

Spatial polarization measures the extent to which each player preferred to move to some regions of the grid over others. To define this variable, I first generated occupancy heatmaps

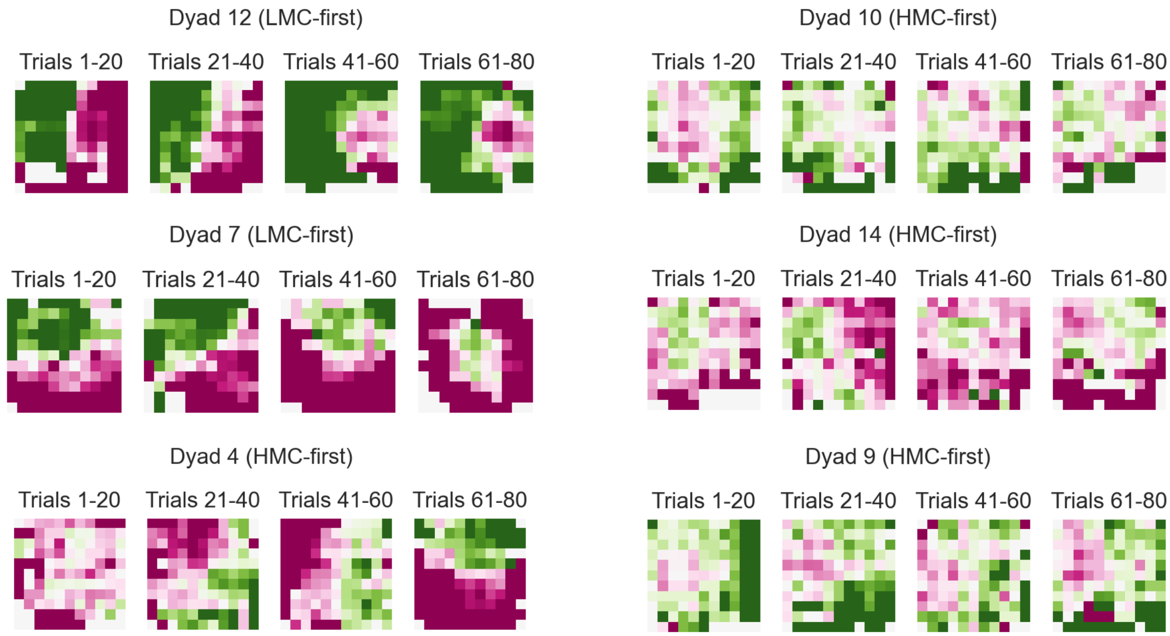


Figure 6.5: Dyad-level spatial preferences. Occupancy heatmaps from the top 3 (left) and bottom 3 (right) dyads ordered by spatial polarization. Heatmaps are rendered based on the percent of each cell's occupancy (visitation count) contributed by each player. Green: majority of occupancy by player 1, magenta: majority of occupancy by player 2, white: balanced occupancy (or not visited in this trial range). Preferences are bucketed into 4 groups of 20 trials (1-40 compose block 1, and 41-80 compose block 2).

(see examples in Figure 6.5), based on the proportional occupancy of each player on each grid cell across a group of trials. See below for a complete definition.

Spatial polarization: $\frac{1}{2}\frac{m_{1,1}}{(m_{1,1}+m_{1,2}} + \frac{m_{2,2}}{m_{2,1}+m_{2,2}}$, where $m_{P,C}$ is the number of moves in a given trial where player P moves onto a grid cell that has been majority occupied by P in the nearest k trials. The nearest k trials for trial t are those in the range $[t - k/2, t + k/2]$, and offset as needed by the bounds of the current block. k is set to 10 to ensure a robust spatial preference estimate, while remaining flexible to changing preferences throughout the course of a block. Spatial polarization can be interpreted as the percent of the time each player visits locations they have preferred in recent trials. In this way, it serves to measure the extent to which spatial preferences guide movements. Please note that the range of spatial polarization approaches 0 when each player visits only the other player's preferred cells, and 1 when each player visits only cells which they have the majority occupancy in the surrounding trials.

Figure 6.5 illustrates the spatial preferences seen for a selection of dyads, as well as how these preferences evolved over the course of the experiment. In Dyad 15, we see a strong bias to occupy the upper left region for player 2, and the lower and lower right region for player 1, and note that these preferences remain roughly unchanged throughout both blocks. Dyad 13 shows a weak but visible left-right spatial division, which dissolves by the time the second block is reached. In contrast, Dyad 16 exhibit unclear spatial preferences in the first (HMC) block, but with the eased monitoring demands of the second block, begin to develop a clear spatial strategy (upper left, vs. lower right).

To see how strategies shift over time, I plotted an aggregate timeseries of the "spatial polarization" variable, and compared this with the other dominant strategy observed, "color polarization" (see Figure 6.6). Of particular interest is a U-shaped trend in the HMC-first grou—spatial polarization reduces during block 1 as a color strategy strengthens and takes precedence. However, under the lower monitoring costs of the LMC counter, average spatial polarization begins to rise again, eventually matching its previous highest level (seen at the outset of the experiment).

Table 6.4 summarizes the results of the mixed ANOVA for spatial polarization. I found a main effect of group ($F(1, 22) = 5.6, p = .027$) resulting in more spatial polarization among dyads seeing the LMC-counter first. An interaction between group and counter type is also observed ($F(1, 22) = 9.0, p = .007$) indicating the highest spatial polarization during LMC trials for the group seeing this block first (LMC first), and lowest during LMC trials for the group seeing this block second (HMC first).

Table 6.4: Spatial polarization: an index of use of a location-based strategy for splitting work.

Block Order	Balanced LMC	Imbalanced LMC	Balanced HMC	Imbalanced HMC
LMC-first	0.802	0.800	0.775	0.755
HMC-first	0.724	0.739	0.745	0.740

6.6.2 Individual and dyad-level differences

Despite the constraints of the task dynamics, variability is seen in both dyad-level strategies and individual behaviors. This variability includes patterns in leader-follower behaviors (e.g. as seen in which player takes the first gem, or makes the first move), color preference consistency or inconsistency, and trends related to strategy selection and communication between individuals, especially during the early trials in a given block. See Figure 6.7 for a selection of dyad-level timeseries which illustrates some of these individual differences and variability.

In the first 6 trials of Dyad 13's timeseries (upper left), a notable difference in behavior can be seen. While player 1 (bottom) collects both gem colors, player 2 uses a fully polarised strategy collecting a single gem color in each trial. These early trials are not successful, but in the following (7th) trial, player 1 mirrors player 2's color polarization, taking only blue, and the dyad is for the first time successful. The remaining trials in both HMC and LMC blocks are nearly fully polarised, with player 2 taking a red gem first in nearly every trial.

Dyad 12 (upper right) avoids color polarization entirely throughout the experiment, but manages to achieve a high success rate despite this in both blocks. This dyad is particularly imbalanced in gem collection, especially in the HMC block.

Dyad 5 (lower left) is color polarised from the first trial, and initially (due to player 1's initial gem choice in trial 3) experiments with swapping colors between trials. Several failures might explain the loosening of polarization towards the end of the LMC block (where player 1 helps out with some red gems). This balancing of work continues into the HMC block,

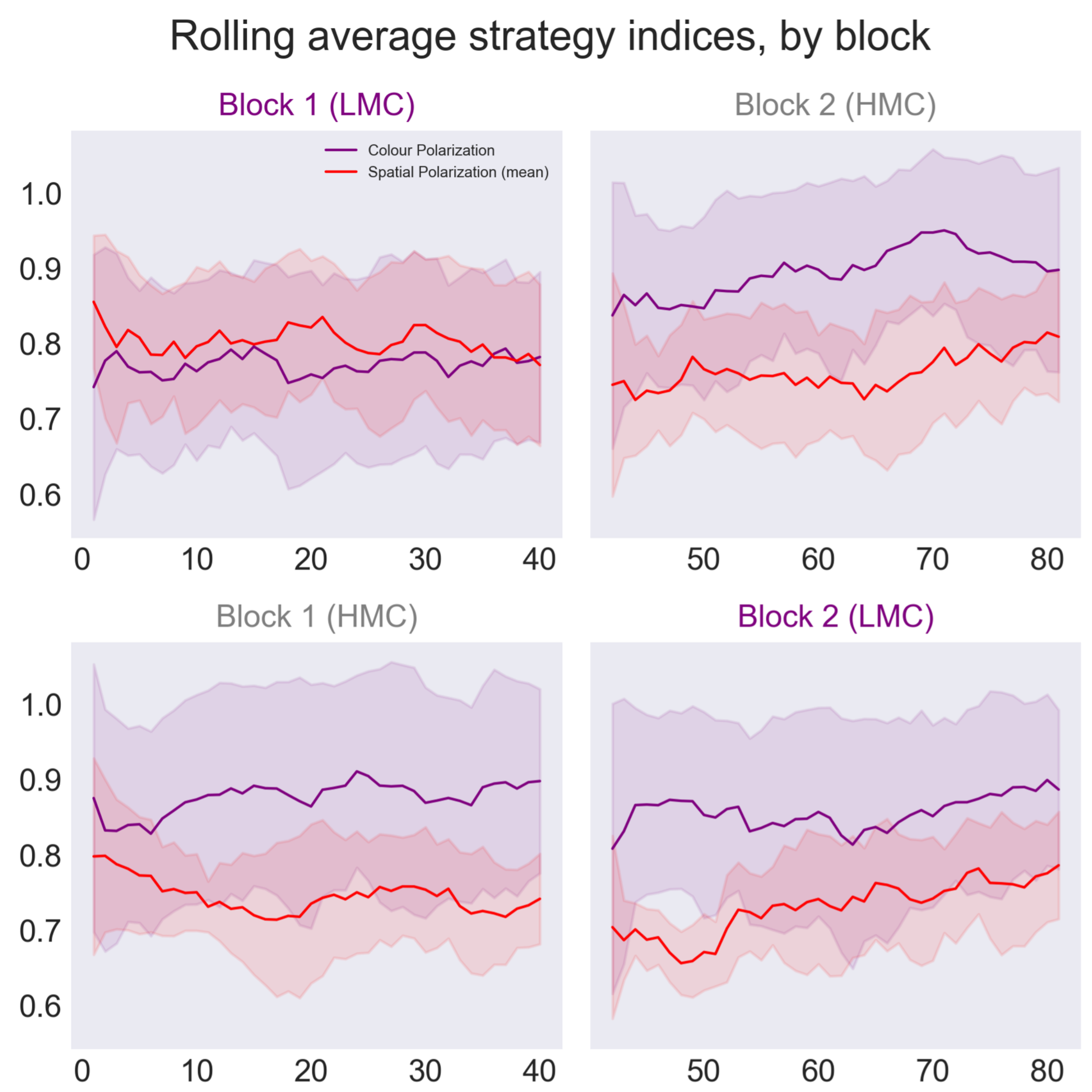


Figure 6.6: Timeseries plots comparing color polarization and spatial polarization. The lines show rolling averages (window size 4) of each index's cross-dyad mean. Standard deviation is shown as a shadow.

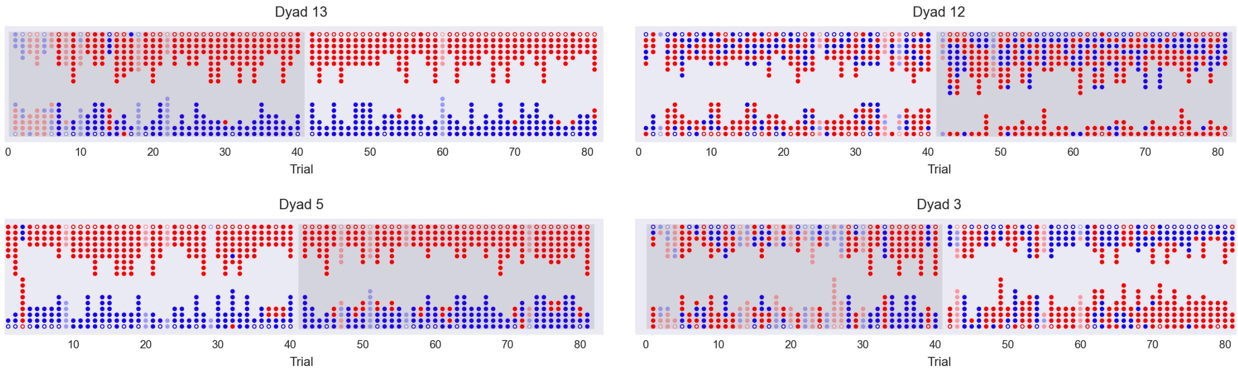


Figure 6.7: Selected dyad-level gem collection timeseries. Gems collected by each player are clustered at top and bottom, preserving order. White markers indicate the first gem collected for each trial. Gray background indicates the HMC counter block. Trials that were unsuccessful are shown as faded. See main text for discussion.

with player 2 taking only red, and player 1 preferring blue but taking reds towards the end of many trials.

Dyad 3 (lower right) is one of the few dyads that developed a fairly successful color polarization strategy, after many failures in their first HMC block, but then frequently swaps color preferences throughout the LMC block. color switches are instigated by first-gem selection by both players, and by the end of the second block, player 1 has settled on nearly exclusively collecting red, while player 2 begins with blue, and takes additional red as necessary (perhaps to balance work).

This selection of dyad-level timeseries demonstrates the richness of both individual and joint behavioral choices that were exhibited by participants taking on the task; see also (Krichmar & He, 2023; Lancia et al., 2023) for evidence of a wide variability of strategies during individual problem solving.

6.6.3 Analysis of score

Table 6.5: Score, a measure of dyad performance awarding higher scores for successful trials with more time remaining.

Block Order	Balanced LMC	Imbalanced LMC	Balanced HMC	Imbalanced HMC
LMC-first	12.47 (4.35)	12.12 (4.27)	14.02 (4.76)	12.22 (4.94)
HMC-first	13.67 (3.55)	13.38 (4.24)	9.95 (5.51)	10.61 (5.67)

ANOVA results for score were similar to those for success; I find two 2-way interactions, and one 3-way interaction. Counter and group ($F(1, 22) = 12.96, p = .002, \eta_p^2 = .37$); Target balance and group ($F(1, 22) = 9.54, p = .033, \eta_p^2 = .19$); Three-way interaction of counter, target balance, and group ($F(1, 22) = 8.60, p = .024, \eta_p^2 = .21$). However, findings also include a main effect of counter type ($F(1, 22) = 4.6, p = 0.044, \eta_p^2 = .173$), indicating

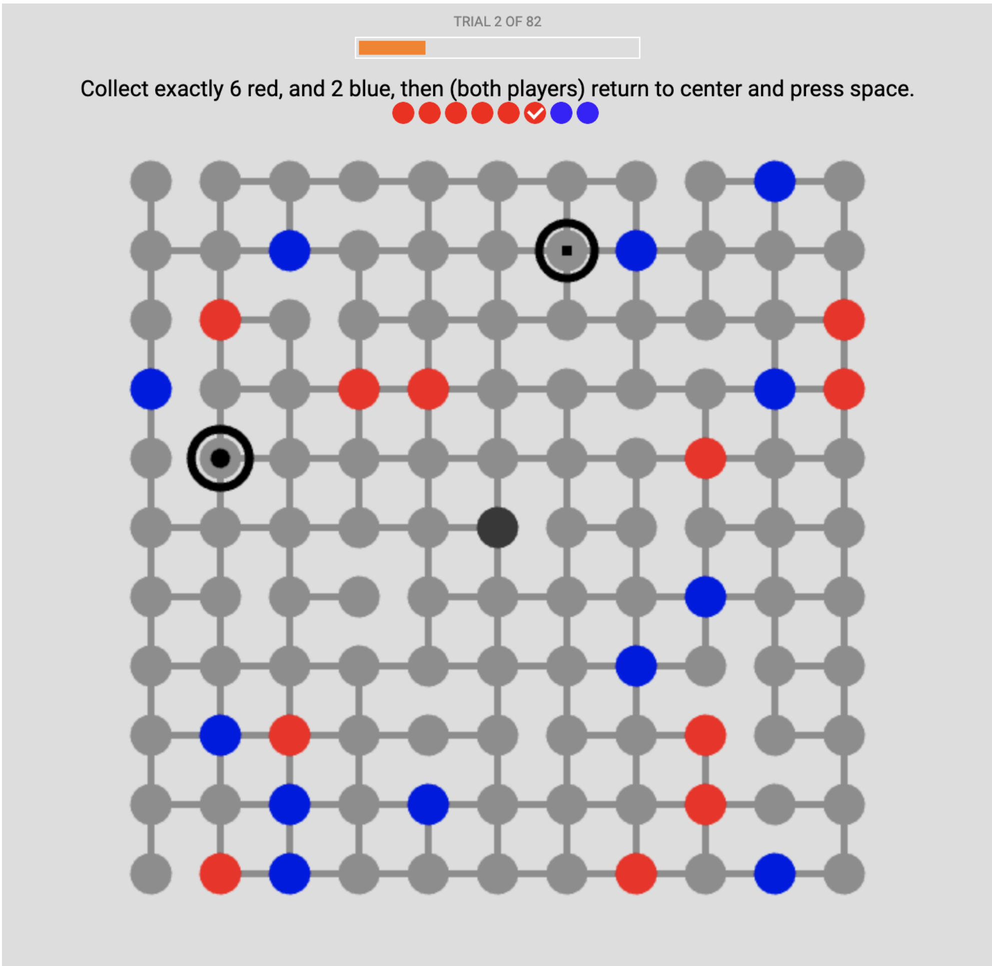


Figure 6.8: The figure shows a screenshot of the game. At the beginning of the trial, the two participants are at the initial (center) position. A short screen capture can also be viewed from this website `https://jgordon.io/project/planning_together`

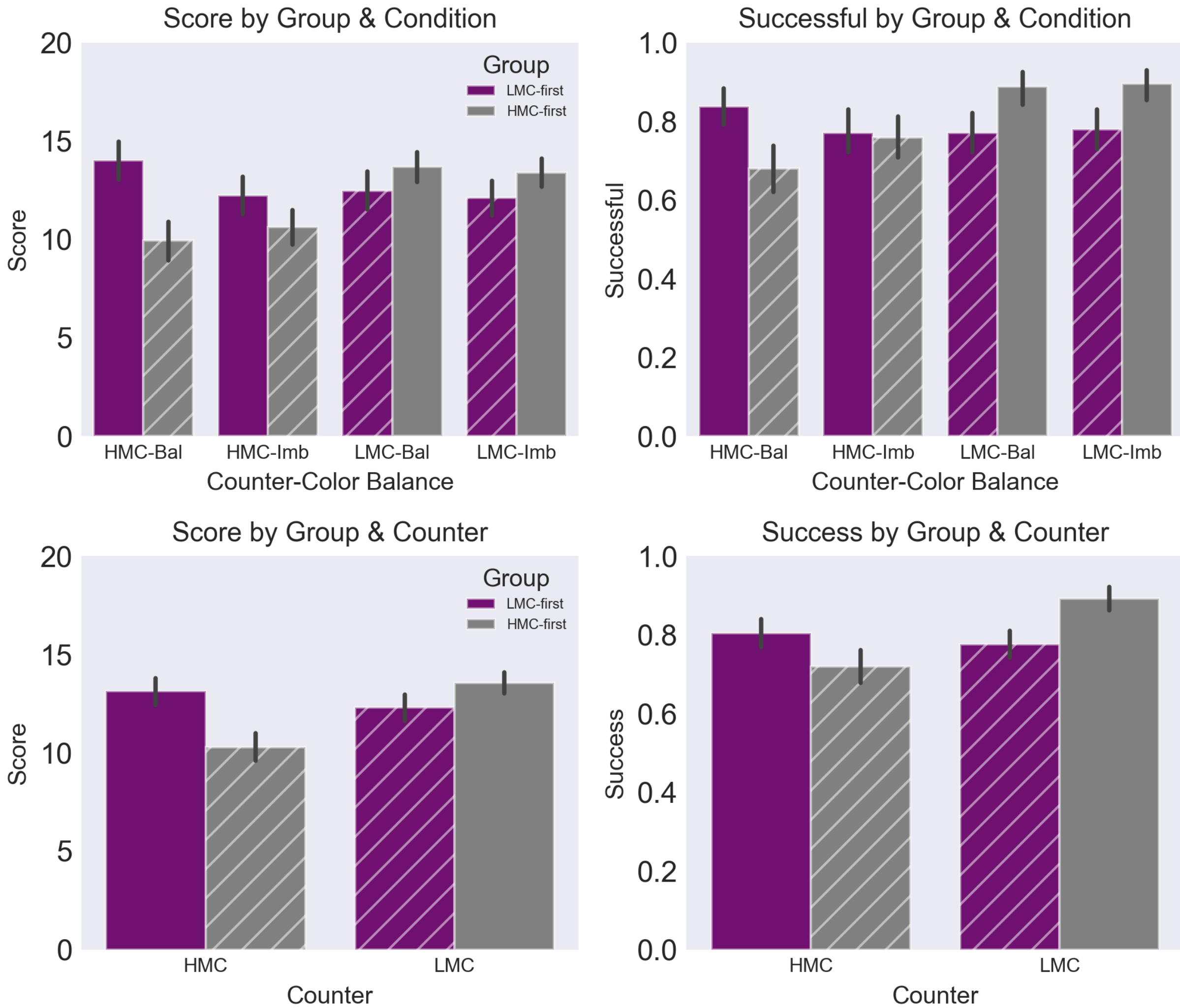


Figure 6.9: Side-by-side comparison of two performance-related dependent variables. Upper-left and lower-left show score (which takes into account time efficiency). Upper-right and lower-right show success rate.

improved score (which is a measure of performance that includes time efficiency) on LMC trials. A side-by-side comparison of the success and score dependent variables can be seen in Figure 6.9 below.

Data availability The data that support the findings of this study are openly available in Open Science Framework (OSF) at `https://osf.io/db32m/`.

Chapter 7

Prospects

7.1 Beyond theory

Having entered the academic world from a prior career designing software tools aiming to address real-world problems, the theoretical, and in parts, philosophical, nature of my chosen research path has sometimes felt like an uncharacteristically intellectual exercise. Though it has been satisfying in many ways to learn methods, historical trajectories, and key results across a variety of disciplines, and then search for opportunities to contribute to these massive collective efforts, questions like "who will use this?" have never been far below the surface. Moreover, in recent years amidst an overdue societal conversation about the role and politics of technologies that inherently aim to learn about us, predict our behavior, and build models of not only how we will act, but how those actions can be manipulated (Zuboff, 2015; Noble, 2018)—a more pressing question arises: "how can this work be responsibly and productively applied?"

In this chapter, I try to offer several answers to this question. This is a speculative chapter, in which I consider a variety of ways the concepts developed in this project might be applied by designers of technical systems interested in helping people reason, plan, and collaborate effectively.

7.1.1 Privacy & sensitivity

To the extent that internal representations of the future exist in the mind, and correlate with relatively easily collected biometric signals, we must contend with the reality that our expectations will, increasingly, be the target of surveillance systems. When these expectations relate to one's own anticipated actions, they convey, among other things, *intentions*, and when they relate to the outside world, they shed light on *beliefs*. Even without direct access to these sensitive attributes, it has been shown that manipulation of visual attention alone can bias decision-making (Armel, Beaumel, & Rangel, 2008). As such, attempts to measure and more accurately infer these internal states pose significant threats to personal autonomy.

In one exploration into subjective responses to this kind of surveillance, I asked how individuals respond to predictive systems taking biometric tracking data as input in an adversarial context. My collaborators and I designed a paradigm dubbed Covert Embodied

Choice, which was composed of a game-like task played in virtual reality that asked participants to select a card on the table without an algorithmic actor correctly predicting which card they would collect (Gordon, Curran, Chuang, & Cheshire, 2021). I examined how individuals adjust their behavior when incentivized to avoid the algorithmic prediction of their intent when gaze, movement, and other physiological signals were tracked. Our results indicated that while participants use a variety of strategies, data collected was highly predictive of choice (card selection was accurately predicted in 80% of trials). Notably, a significant portion of participants became *more* predictable despite efforts to obfuscate, possibly indicating mistaken priors about the dynamics of algorithmic prediction.

This work highlights the potency and sensitivity of the kinds of biometric signals studied in this dissertation, as well as the incongruity of many users' beliefs about the mechanisms and abilities of surveillance systems.

It is worth noting that similar pitfalls exist even among systems which fail to make accurate predictions. In this case, one source of danger comes to us in the form of automation bias, which leads individuals to overestimate the accuracy of algorithmic predictions (Goddard, Roudsari, & Wyatt, 2012). This risk is in some sense even more pressing given the perverse incentives in the private sector to launch new systems, even before they can be fully understood, to benefit from first-mover advantages, and collect additional data deemed critical to improving performance in the future. Such risks therefore carry no dependency on the likely, but difficult to anticipate, timelines of advancing algorithmic efficacy (not to mention theoretical limits on what can be meaningfully predicted).

7.2 Opportunities

With these important considerations in mind, it is worthwhile to consider a number of domains in which an improved conceptual understanding of the embodied dynamics of (future oriented) simulation might be productively applied. I group these opportunities into two categories. First, those that aid individuals in *externalizing* their own cognitive models or affordance landscapes in order to improve legibility and independent reasoning. And second, *synchronizing* these internal representations with other co-actors with whom an individual may wish to coordinate, whether these are students, collaborators, or perhaps software tools and artificial intelligence-based systems.

7.2.1 Externalizing

Augmented planning & embodied reasoning Present work exploring the potential value of visualizing affordance landscapes (such as those introduced in Section 3.2.3) as training aids in climbing (Maselli, Gordon, & Pezzulo, n.d.) suggests there may be potential in more abstract planning domains as well. Might systems enabling planners to externalize representations of a landscape of future trajectories, whether relating to personal and professional decisions, or business opportunities and risks, be deployed as tools to augment and ground reasoning and the consideration of likely (or unlikely) contingencies? Though encoding the causal relations between actions and uncertainties in these abstract domains would pose significant technical challenges, designers could leverage a deep literature exploring ways to elicit priors and beliefs about conditional dependencies, and map these onto an interactive visual landscape affording interactive experiments in a preplay like fashion (S. R. Johnson, Tomlinson, Hawker, Granton, & Feldman, 2010).

A number of studies have shown that the technology-enabled visualization of distant future events can encourage individuals into more future-oriented action (that is, effectively lowering the well-studied temporal discounting effects seen in psychology and economics). As one example, aged versions of one's own face induced higher confidence in response to financial literacy questions, and seemed to induce greater interest in long-term financial planning resources (Sims, Raposo, Bailenson, & Carstensen, 2020). The robustness and longevity of such effects remains to be seen, but the notion of using multimodal technologies to help ground future considerations within the embodied present is intriguing. Other work in a similar vein has explored the use of image generation models to produce photo-realistic composites of familiar urban locations confronted with the effects of climate catastrophes (S. Zhou, Luccioni, Cosne, Bernstein, & Bengio, 2020). While trust in artificially generated media promises to be an increasing challenge of the coming years, helping individuals and communities imagine, or even experientially preview, some of the futures forecast by climate researchers, may offer a valuable communication tool.

Prosociality Though it is alluded to only indirectly in the previously mentioned work, another recent study has shown that future oriented cognition can increase prosocial behavior (Cernadas Curotto, Sander, d'Argembeau, & Klimecki, 2022). Given the relatively simple intervention tested in this study—prompts were to imagine future episodes for a short duration—it may be interesting to see whether guiding individuals into more richly constructed simulations of a range of time-spans, perhaps in combination with visual imagery, might improve efficacy or the longevity of the effect.

Though the authors note that the reason for this effect is unclear, I might speculate that the widening of the affordance landscape prompted by future-oriented cognition necessarily highlights its interrelation with individuals and social contacts in the wider community, and perhaps forefronts the importance of those relations to long-term homeostasis.

Mental health Links between prospective abilities and mental health disorders such as depression and anxiety have been well established. Beck proposed one of the first frameworks in cognitive psychology — the negative cognitive triad — to explain not only the effects but potentially the causes of depression as negative biases across three domains of cognition (Beck, Weissman, Lester, & Trexler, 1974). These components are: a negative view of the world, the self, and, notably, the future, which together mark the primary symptoms of depression. Building on this work, Roepke et al. suggested that negative future thinking is the key causal factor in depression (Roepke & Seligman, 2016), stemming from systematic biases in the generation of possible futures, their evaluation, and attitudes and beliefs towards the outputs of these processes. They argue, then, that treatment should deepen its focus on the rehabilitation of prospective abilities.

Similar to the framing of prospection in depression, it is likely that biases in the generation and evaluation of future episodes helps to explain the mechanisms underlying a variety of anxiety disorders (Miloyan, Pachana, & Suddendorf, 2014). One of the central claims of the present work is that the ongoing anticipatory projection of the future is inherently embodied; that is, simulations are run in, and inherently experienced by, the body. Because these simulations are generated, and not constrained by any notion of true probability, increased attention to negative outcomes that are typical of anxiety could result in a "salient now" that may contain a host of overlayed, unlikely (or perhaps completely confabulated), and potentially negative-affect-imbued, sensory experiences. In this case, the anticipatory response — a typically adaptive behavior to evaluate and prepare motor sequences that may soon need to be deployed — instead triggers maladaptive responses that constitute deeply experienced, stress-inducing, previews of highly implausible timelines.

In depression, anxiety, and plausibly other related psychoaffective disorders, perhaps methods of inducing expected futures, making them legible as objects of reasoning, and highlighting their differences from those generated by others, e.g. to help investigate their veracity, may provide a useful new perspective, and new methods, for clinicians and individuals[1].

As alluded to in this section, comparing the nature of internal prospective representations may offer valuable perspective and a focus on the divergences of our subjective experience. To explore this further, I next consider the potential value of prospection towards strengthened social interaction, within a group of related opportunities focused on *synchronizing* internal representations of the future with others.

7.2.2 Synchronizing

With experts Related to the climbing training example discussed in Section 7.2.1 above, is an idea based on the intuition that perception of effective task-aligned affordance landscapes is a key skill developed with expertise. If so, perhaps training a model with the ability to generate such landscapes based on expert performance (e.g. through methods in line with the empirical approaches discussed in Section 3.2.3) and visually sharing these landscapes during training, may provide a valuable learning tool. An example of a visualization of such landscapes in the climbing context is shown in Figure 7.1, with link to the video in the caption.

[1] Such externalization of thought patterns, and questioning of veracity, is in line with common techniques used in cognitive behavioral therapy. Perhaps a stronger commitment to the anticipated future as primary constituent of the conscious present would help understand and design strategies around such mental health challenges.

Figure 7.1: Still frame from animation of a climber on the MoonBoard, with affordance landscapes calculated and projected prior to each move. The video is hosted at `https://osf.io/req2c/` in the "Climbing Affordances" folder. Note that affordance landscapes were computed and projected onto the climbing video retroactively.

Even without additional processing or model development, raw signals such as eye gaze have been explored for some time as ways to capture expert performance (Tien et al., 2014), evaluate learners (Harezlak & Kasprowski, 2018), and augment training (Chetwood et al., 2012). In Chetwood et al. 2012, a visual UI successfully augmented trainee performance in identifying a target site, when a supervisor's "point-of-regard" was projected onto their laparoscopic screen during a simulated surgery (Chetwood et al., 2012).

There may also be opportunities to augment results from studies exploring the use of surprise and the illumination of "ignorance boundaries" as a learning tool, by bringing embodied sensory aspects to the experience of surprise. Ranney et al. demonstrated that eliciting numerical predictions from individuals prior to reporting true statistics can often produce a sense of surprise that is critical to producing enduring belief updates, especially in the realm of political or deeply held ideological beliefs such as those related to climate change

(Ranney, Munnich, & Lamprey, 2016). In particular, the authors believe that an immediate back-dating of beliefs after learning the true statistics (when predictions are not elicited) may undermine the experience of surprise, and work against an acknowledgement of ignorance useful to learning. Separately, within the embodied learning literature, a number of results have demonstrated that improved conceptual understanding can be achieved when learners are guided through tasks requiring them to embody, e.g. through gross limb movements, even abstract concepts such as proportionality (Abrahamson, Lee, Negrete, & Gutiérrez, 2014).

Combining these two approaches may inspire novel techniques for communicating and improving reasoning about non-intuitive time evolving quantities, perhaps leading to more effective prospection of future outcomes. Imagine an exercise in which educators elicit predictions about future national statistics of the types used in Ranney et al. (2016) (e.g. demographic shifts, or carbon dioxide rates in the atmosphere), by asking learners to approximate historical and future trajectories for these numbers. By having learners embody their estimates, e.g. using both arms to indicate estimated proportional changes over time in line with Abrahamson et al. (2014), and then guiding movements consistent with true statistics, learners whose estimates differed substantially from true figures might produce a multi-sensory memory of dissonance (an embodied experience of surprise) useful to facilitating lasting belief updates.

With collaborators As considered in Chapter 6, joint projects and effective collaboration require the development of shared representations of a task, and the strategies which compose its solution. These representations need not (and indeed, cannot be) identical, but incongruous representations (such as the convention mismatches that sometimes undermined joint success in Study C) threaten the efficacy of collaboration.

An underexplored avenue of research and application involves the sharing of biometric and behavioural variables, either as “raw” signals (e.g. gaze targets), or through inferences about internal representations made based on these data, supporting the synchronization of task-relevant reasoning states between one or more collaborators. As seen in Study A and B, as well as numerous other prior empirical and computational contributions, a significant portion of task representation comes in the form of salience filtering, or the masking of task-irrelevant stimuli. To the extent that gaze correlates with subjective beliefs about salience, then sharing such masks may be, in itself, a valuable signal by which to converge upon a joint representation of the problem.

Such concepts become grounded when considering the possible increase in the use of mixed reality (MR) applications. Given the increasingly common inclusion of eye-tracking systems in modern MR devices—and use cases ranging from hybrid or remote workplace collaboration, to medicine, education and training—it is easy to conceive of attention-awareness as an additional ambient communication channel. Extending this, under the assumption that humans naturally generate task-relevant representations (e.g. the simulated symbolic manipulations studied in Shimojima and Katagiri (2013)) to augment their perceptual experience and possibly improve reasoning, perhaps these representations could be made explicit and shared in real-time. Imagine a joint interface enabling users to add context to an object in a shared environment (whether real-world, as in augmented reality, or virtual), by looking at it and enacting a gesture, or speaking a verbal keyword. Perhaps by externalizing mid-computation contextual cues or transformations (e.g. similar to how participants offloaded perceptual rotations to the environment during Tetris play in Kirsh and Maglio (1994)), a more tightly synchronized belief state could be achieved, fruitful for collaborative exercises. Indeed, given the future-oriented cognition lens, positional or object-oriented intentions (e.g. "I'm planning to pick up *[this]* shared tool" or to use the task studied in Chapter 6, "I'm going to shift and collect red gems") are a likely candidate for highly salient contextual tags that may most directly augment fast-moving joint exercises.

Successful collaboration also often depends on a compatible pseudo-causal understanding of the world (e.g. through sensemaking, as reviewed in Section 2.6.2, or the development of qualitative models such as Forbus (2011)), particularly as these relate to the likely consequences of actions. To skirt the challenges of true causality, perhaps agreement on approximate conditional dependencies would be an effective starting place. However, not all collaborators will be conversant in interpreting (or fore that matter generating) the probabilistic graphical representations which are often used to encode such relations. Perhaps, then, a concept central to much of this work could play a role: counterfactual simulation. Indeed, if individuals were able to convey rich (imagined) sensorimotor trajectories through a state space relevant to the task at hand, such trajectories would necessarily capture a wide variety of conditional dependencies. By sharing such trajectories with a collaborator, inconsistencies in probabilistic understanding would be highlighted and made legible for ongoing joint reasoning. In line with the ongoing and enactive features of sensemaking, recurrent iteration upon these trajectories could constitute an interactive process enabling refinement through mutual feedback.

To make this more concrete, let's imagine a joint task of building a shelter given a set of materials. Effective collaboration may benefit from a shared vision for the final outcome (e.g. the type, size, structure, and other attributes) of the target shelter to be built. However, while a static image or diagram representing the end goal may aid this convergence, assumptions critical to the design's realization may remain implicit. Rather, if a collaborator could share a proposal represented by a sequence of key steps (e.g. an intermediate stage in which each individual holds up a base pole, while canvas is pulled over the top) assumptions such as "each of us can support a pole with a single arm" would be made explicit, and divergent beliefs identified for discussion and perhaps revision of the plan. In addition to verbal reasoning, drawn sketches offer an obvious medium to convey abstract representations of actions and intermediary states composed into sequences, but recent development of generative artificial intelligence models (Saharia et al., 2022; Rombach, Blattmann, Lorenz, Esser, & Ommer, 2022) may soon lead to tools capable of quickly conveying plans, including the probabilistic assumptions they depend upon, to collaborators and other stakeholders.

With technical systems Similarly, the development of tools and software systems designed to aid reasoning processes, particularly complex planning exercises, may benefit from an improved understanding of users' (prospective) representations of a problem or task.

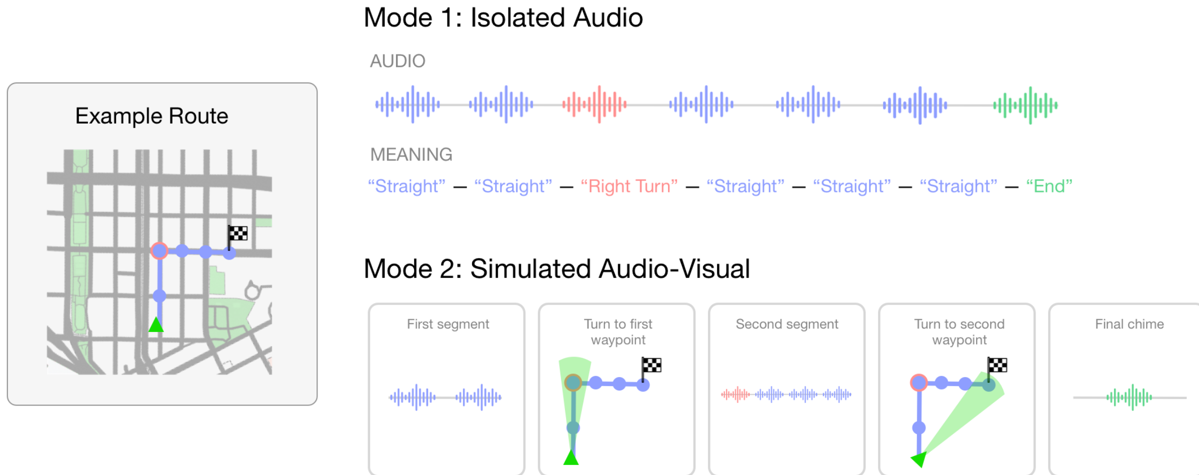


Figure 7.2: Illustration of two spatial audio route communication modes explored in (Gordon, Fiannaca, et al., 2023).

One example of this is recent work I completed with collaborators at Microsoft Research, in which we investigated the use of a novel pedestrian navigation support tool designed with internal representations of navigational foresight in mind. In "Hearing the Way Forward: Exploring Ambient Navigational Awareness with Reduced Cognitive Load through Spatial Audio-AR," we tested the feasibility of a design whereby the next few steps of a navigational route are conveyed to the user via an embodied and interactive spatial-audio driven experience (Gordon, Fiannaca, et al., 2023). A schematic of the two spatial audio modes we designed and tested is shown in Figure 7.2, including both the isolated spatial-audio only condition, as well as the more richly embodied head-direction guided route feedback interaction (which we call simulated audio-visual).

In the final chapter, I will summarize the work presented, revisit a theoretical proposal for embodied prospective simulation and its connections with action, and end by highlighting a few of the many open questions, some broad, and some specific, which remain as invitations for further scientific study.

Chapter 8

Conclusion

8.1 Summary

In this final chapter, I will summarize the studies presented in this dissertation, and a selection of key findings from each. I'll then present a the high level synthesis of the theoretical ideas that have most captured my interest throughout this work, and a more informal description of how I see these ideas coming together in prospective embodied action.

8.1.1 Experimental studies presented

In the preceding chapters, I presented results from three studies leveraging novel experimental paradigms for the study of planning and navigation in environments imbued with various kinds of uncertainty.

In Study A (Chapter 4), visual access to the landscape was highly constrained, requiring narrow visual exploration to identify promising paths and key decision points. Furthermore, the ability to traverse between two nearby holds could not be verified with certainty until the hold was reached, creating an additional level of uncertainty which participants had to contend with during pathfinding.

In Study B (Chapter 5), I explored human behavior within a more immersive and embodied navigation task, which required participants to plan their path through a landscape with unknown goal locations, prior to beginning navigation. This allowed me to ask questions about the sequential, temporal, and spatial properties of visual search during planning, and their relationship with the navigation choices enacted.

In the third and final experiment, Study C (Chapter 6), I explored planning under a fundamentally different kind of uncertainty—where the landscape, and key object locations, were both fully observable, but where a collaboration with another participant produced an uncertainty over *joint actions*, i.e., who is going to do what, and when. Participants managed this type of uncertainty through the development of (a diverse array of) conventions to split work between themselves, thereby decoupling their decisions from the ongoing decisions of their collaborator.

8.1.2 Summary of key findings

In Section 1.4, I outlined research questions in three areas: 1) embodied planning, 2) visual salience and simulation kinematics in navigation planning, and 3) prospection in joint planning. I next summarize the key findings within each area.

Embodied planning Questions in this area aimed to explore limits of classical approaches, and suggest avenues and opportunities to develop a more nuanced and naturalistic theory of planning in the real world. I argued in Chapter 3 that this need to accommodate embodied dynamics—in decision-making as well as less explicitly choice-centered natural activities—requires us, as a scientific community, to develop new models better able to capture the richness of the mental processes that anticipate and drive action in the world. The reviewed work displays a range of approaches to bring embodied intelligence into communication with the classical literature on decision-making and planning. In particular, the concept of the affordance landscape is developed, with example studies in various naturalistic (and fundamentally embodied) tasks such as playing team sports, climbing, and crossing a river. These theoretical arguments and methodologies set a foundation for the next research area, in which a mode of embodiment particularly relevant to visual creatures like ourselves is explored.

Visual salience and simulation kinematics in navigation planning This area captures the primary empirical questions of this dissertation, and motivates the use of an agent model designed to iteratively learn the sort of affordance landscapes discussed in Section 3.2.3. Questions about attractors of visual attention, and temporal sampling dynamics, were tackled through a variety of behavioral analyses. Key findings in Chapters 4 and 5 include: the predictive relationship between planning-time movements of eyes and arms and navigation-time decisions, geometric and map-geometry attributes most correlated with visual attention, as well as various hierarchical aspects of the sequential gaze patterns seen during planning.

Prospection in joint planning In the third area of inquiry I looked beyond the individual spatial navigation paradigm to explore the case of multiple collaborators who must, through a variety of techniques (of which some necessitate planning and simulation), converge on a shared strategy to jointly solve the task at hand. I reported findings suggesting a history-dependence of strategy selection when tasks dynamics changed. This implies an asymmetry in either the expected cost of establishing a joint convention cost, the expected cost of its implementation, or both, dependent on which task context was presented first. Balance of work (i.e. fairness) was found to be lower in imbalanced environments, as well as during the more difficult blocks requiring greater partner monitoring. Finally, I found a persistent bias towards spatial separation (likely to avoid interaction and enhance decoupling between collaborators), even among dyads using strict color-based strategies, highlighting that non-trivial hybrid approaches which mix conventions can be productively adopted even in non-verbal collaborative settings.

8.2 Embodying the future

I introduced this work with a grounded example highlighting that even the acts of reading and writing (which we typically consider to be primarily cognitive endeavors), are active, embodied, and necessarily forward-looking experiences. This set the tone for a series of arguments about why cognition can be thought of as fundamentally future oriented, as well as the importance of considering the role of the whole organism, the dynamic array of sensors that are themselves active and adaptive samplers of a world far too complex to consume all at once. To end, I'd like to take the opportunity to synthesize a number of ideas explored throughout my research, and, with a more speculative lens, extend the story of the embodied prospective world model back to what, as William James claimed, cognition was always for—movement (James et al., 1890).

I take as a starting point, the sort of spatiotemporal generative model common to the active inference approach (as introduced in Section 1.2.3, and further reviewed in Section 2.4.3). Picking up here, I start with the question: how do the covert simulations inherent to our conception of the SGM during planning give rise to the overt actions necessary for interaction in the world? The answer to this, I believe, is a recurrent information-sharing dance between two tightly coupled levels of cognition.

At the lower level is the *actor*, an (evolutionarily older) motor-centric system responsive to immediate rewarding consequences. Then, operating at a higher level, is the spatiotemporal generative model (SGM) that has been the focus of much of our attention; for brevity, we can call this the *simulator*. The interface between the two is best explained using Pearl's *do*-calculus—through counterfactual experiments that allow a causal reasoner to hold the world constant while "wiggling" a variable of interest to identify its causal influence(Pearl & Mackenzie, 2018). Such wiggles operate at both levels. Within the actor, motor preparations for action (whether overt or covert via the corollary discharge mechanism described in Section 1.2.3), signal intention to the simulator, and prompt adaptation around a (now) more likely sequence of actions. The result is a concentration of simulations originating in newly probable (lower energy) regions of the landscape, each rollout a counterfactual experiment, and each generating sensory consequences as it passes through states of ambiguity, reward, and risk.

Similarly, counterfactual wiggles flow top-down, resulting both from stochasticity in the rollout dynamics that guide these sequences, as well as affordance competition (Cisek, 2007) in which multiple sensory causes or sensorimotor contingencies (some compatible and some mutually exclusive) generate ambiguity via their rivalry. Such time-varying changes in the simulator may alternately inhibit or reinforce the actor's experiments, prompting changes to the affordance landscape, and therefore the focus of simulations, and so on.

As such, a recurrent dynamic appears, with the simulator continuously predicting the state of the actor (even the slightest change to activation in the motor system could be a sign of an emerging action commitment that fundamentally alters the anticipated sensory future), and the actor responding to ever-changing forecasts (future affordances) of rewards and action sequences, and potentiating new motor programs based on the simulator's rehearsals. This dance continues until a winner emerges—a nearby density within the simulator offering low enough risk, and attractive enough reward expectations, and parsimonious enough to fit in working memory. In other words, a *plan*, composed possibly of multiple action sequences conditioned on not-yet-observed aspects of the future.

This framing, however, brings to mind the discrete and reductive "plan-then-act" paradigm I argued against in Chapter 3. As a potential remedy, let's consider the Lewinian "field of safe travel" used as a conceptual illustration in Figure 1.2a. In this conception, driving entails the generation and protection of a spatially extended proxy for the self and its homeostatic imperative. Threats to this extended and, as I have argued, embodied, representation such as the detection of a vehicle entering its boundary, necessitate the recruitment of motor responses (the impulse to force the steering wheel left) which modifies its shape and extent, and unleashes a flood of uncertainty in newly probable regions (the road to the left) necessitating new, retargeted simulations.

It is useful, then, to push the idea of perception as unconscious inference attributed to Hermann von Helmholtz (Boring, 2008) beyond questions of "what are the most likely causes of this observation?" to an inferential process much more interested in the nature of an individual's temporal trajectory, i.e. one that asks: "which timeline are we in?" The sensory evidence necessary to answer this question is continuous and distributed across a number of sources: current internal motor activity, as well as exteroceptive information from ongoing sampling in areas of the environment most likely to reduce the number of possible answers to this question.

Such a conception of a coupled interplay between subunits of a multilevel cognitive system is, of course, not novel. Kahneman proposed system one and system two (Kahneman, 2011) to reason about the complementary nature of similar functions[1]. Actor-Critic algorithms and subsequent architectures involving a behavior-inducing policy unit and a learner of state-to-value maps reflects similar dynamics, and goes back at least to Barto and Sutton (Barto, Sutton, & Anderson, 1983). Further, Schmidhuber and Ha's conceptions of a policy unit afforded a query interface with a world model suggest yet another adjacent configuration of these ideas (Ha & Schmidhuber, 2018; Schmidhuber, 2015).

All these related conceptions have been foundational to my thinking, and to the research program undertaken here, so it may be helpful to summarize the key novel contributions I hope to offer through this work.

Main contributions First is a conceptual extension made by projecting the body outwards into its likely future, and seeing this projection as a first class representation of the self within which simulations are continuously run. It is, therefore, in some important sense, inseparable from the individual. Furthermore, the sensorimotor patterns generated in these simulations are not fundamentally different from the inferential (and simulative) process of perception in the present, and therefore are equally constitutive of our moment-to-moment experience.

Second, I argue that models, such as those explored in this work, can help to achieve a unified conceptual framing of future-oriented dynamics that links both low-level predictive processing (which may operate over short temporal durations), with the kinds of dramatically extended episodic simulations typical when studying mental time travel. While a diversity of

[1] Though in an interview with Scott Kaufman, Kahneman also warns against the temptation to excessively agentize these subunits, which he notes are distributed and difficult to empirically decouple (S. B. Kaufman, 2021).

mechanisms are entailed in such a unified framework, I believe seeing these as two ends of a continuum with shared function (i.e. at David Marr's top, computational, level of analysis), and more than likely, algorithmic and process-level correspondence as well.

Third, is to offer a series of empirical findings, some of them consistent with the more speculative ideas raised earlier in this section, and others highlighting the remaining gaps in our understanding of the tremendously complex processes underlying visual search, prospection, and planning.

8.3 Open questions

Given the size of the project undertaken here, it is maybe not surprising that many of the questions that attracted my interest in these areas remain open. These inquiries have also raised new questions I did not know to ask at the outset. To close, I'll highlight several of these questions according to increasing levels of abstraction.

At the lowest levels, can we characterize the temporal relationship between saccades and (visually guided) simulations? For example, in spatial planning paradigms like that studied in Chapter 5, can physiological or neural data help identify when gaze fixated locations correspond directly to simulated states, when they diverge, and when (as hypothesized by the agent models explored) they play a scaffolding role in which simulations are enriched by present visual stimulate, but operate independently?

The highly parallel nature of neural processing has been well studied in neuroscience, especially in the visual system (Nassi & Callaway, 2009), however, the implications of this neural parallelism for the kinds of dynamic mental imagery underlying simulation is still unclear. Future work must identify the what constitutes unambiguous evidence of parallel simulation. In spatial navigation, perhaps documentation of hippocampal theta sequences (e.g. in results like A. Johnson and Redish (2007) discussed in Section 2.3.2), but instead following divergent maze segments simultaneously might be evidence that meets this objective. In other domains, disambiguating between the parallel and serial (or unitary versus probabilistic) divide, is likely more challenging. However, identifying the limits to parallelism in simulation, the contexts under which multiple incongruous representations can be reasoned about simultaneously (and the thresholds of incongruity), and understanding how these relate to working memory constraints, are all valuable lines of inquiry.

Looking beyond the scope of Study B, for example in contexts where explicit visual representations are not available, or when visual stimuli are inconsistent with the planning representation, an additional open question is whether overt eye movements are still useful to spatial planning. In these cases, are saccade distances attenuated to conserve motor resources? More speculatively, might we find evidence that saccades can operate as well-timed visual occlusion, thereby favoring top-down predictions, and amplifying simulated percepts?

Significant scholarship has explored the dynamics of statistical learning—how individuals encode sensory experiences and develop models of what to expect—but in more continuous domains, what factors affect the stochasticity of simulated trajectories? How does stochasticity adapt to the dynamics of the task at hand, and other considerations like resource constraints, urgency, and risk?

Similarly, at higher levels, what affects the temporal horizon within which simulated futures play out? While economists have long studied temporal discounting of future value, might an embodied perspective offer insights into what specific contextual factors (e.g. physiological factors like energetic considerations, stress, or ongoing cognitive load) allow for longer or shorter time horizons to be explored? Similarly, what factors affect the depth, and level of abstraction, of simulated episodes?

And finally, it remains to be seen how well the conceptual frameworks explored here, which are easily applied to navigation-like settings where the generative model is explicitly spatial in nature, extend to more abstract domains. In the studies presented here, the state space maps roughly one-to-one onto spatial occupancy, thereby allowing visual attention to more plausibly capture consideration of specific future states. But in more abstract planning paradigms, gaze targets have a less direct correspondence with the future. Are we correct to assume that sensorimotor imagery and simulation play an equally important role in these domains, e.g. when counterfactual simulations contend with factors that do not have a representation easily reduced to a spatial location (e.g. economic transitions, changes in knowledge or conceptual understanding, or relational shifts)? How embodied actions (like eye movements and others) get recruited during exploration of this less spatially defined structure remains a fascinating open question for future study.

While some of these questions may appear subtle or even esoteric, I am optimistic that they can help us develop a more nuanced understanding of the dynamics by which individuals generate and test expectations for the future. Working on these questions should help deepen our understanding of the way these prospective cognitive processes are embedded in a body, and a complex motor system, that grounds them in the world around us. If so, we may better appreciate the aspects of human intelligence that are truly unique, and defined by a complexity that is still far out of reach of replication in artificial systems. As such, perhaps our computational abilities will be better leveraged towards the design of tools personalized to individual preferences and idiosyncrasies, ore even self-acknowledged oversights. Such systems may be able to externalize more legible beliefs, risks, and hypothetical futures, and by availing these to us, help enable more effective communication and collaboration. Finally, by better understanding how individuals project themselves into uncertain futures, and reason about long-term consequences, we might prepare ourselves, as a society, to more effectively face the challenges still ahead.

Appendix A

Supplementary materials for Study B

A.1 Supplementary methods

A.1.1 Map generation

Maps for the task were generated in a three step procedure: generate map pool, filter based on criteria, and finally, select an experimental sample. Below, I detail the procedure for each.

Procedurally generate map pool

Map geometry was defined similarly to Zhu et al. (2022) but with boundaries defined by tile types, rather than walls drawn along tile edges. Specifically, a hexagonal grid composed of equilateral triangular tiles, and a side length of 7, formed the base map. Each tile could be set to one of four types: water (blue), muddy water (dark blue), land (gray), or obstacle (black). Land tiles were low altitude permitting vision, but not navigational access, while obstacle tiles were tall, and blocked visual access. To generate unique maps with variation in structure and connectivity, the algorithm below was followed (See Algorithm 1).

Filter pool

To filter the pool of map candidates (defined by an assignment for each tile type, and an array of 6 reward cache locations), I confirmed that each one satisfies the following criteria:

1. No goal locations visible (line of sight) from origin
2. All goal locations reachable from origin
3. No obstacles adjacent to origin

Select experimental sample

Finally, with the filtered pool of eligible map candidates, I selected 30 maps to compose the main experiment. I then randomly assigned half of the maps to each uncertainty group (low vs high). For both groups, I randomly chose 2 rewards among the 6 goal caches to be present

Algorithm 1: Procedural map generation

```
Data: tileTypes[N] = WATER
;                                          /* Init all tiles to water */
Result: tileTypes[N], rewardCaches[6]
n ← Uniform(10, 40) ;                      /* Choose n strokes to add */
i ← 0;
while i ≤ n do
    k ← Uniform(10, 40) ;                  /* Choose k stroke length */
    j ← 0;
    if i ≤ n/2 then
        type ← MWATER;
    else
        type ← choice(LAND, OBSTACLE);
    tileCursor = randomTile() while j ≤ k do
        tileTypes[tileCursor] = type;
        tileCursor = successors(tileCursor) ;   /* Successors is a list of
         adjacent tiles with prior visited tile removed */
        j ← j + 1;
    i ← i + 1;
rewardCaches[] = {};
r ← 0;
while r ≤ 6 do
    rewardCaches[r] = randomTile(WATER);
```

in the map. For the low uncertainty group, I then randomly removed 3 of the unused goal caches. The end result was a list of maps, where each contained 2 present rewards, among either 3 or 6 possible goal cache locations.

In addition to the 30 procedurally generated maps, I hand designed two practice maps with trivial geometric structures, to be used at the start of the experiment during the tutorial, as well as three maps testing other information geometries of interest. After the practice trials, map order was randomized for all participants, while ensuring that these latter three maps were always shown at the end of the experiment, to avoid any influence on the primary, procedurally generated maps.

Four sample maps from the main experiment can be seen in Figure A.9 (lower-right figures show actual map as represented to participants during planning, with navigation trajectories overlayed).

A.1.2 Subgoal analysis

To investigate questions about the time-course by which key map regions are identified, I explored three additional regions constituting plausible subgoal types. These were defined by the following key features: the location of the first divergence when following the shortest path to each goal cache (first fork), the location along any shortest path to goal at which all goal caches are observable (first certainty), and the location of the first collected goal (first collected). Note that the first two regions were specific to each map, whereas the third was specific to a particular participant and trial, since it compared actual navigation choices to the map's goal locations. Time of first fixation on each subgoal were plotted, in addition to computing the duration of fixations falling within each region during planning.

A.2 Supplementary empirical results

Subgoal planning timing analysis Subgoal analysis timeseries reveal timing differences in first visit to each subgoal type (see Figure A.1). Mean first visit time, in both groups, was earliest for any goal cache, followed by first forks, "first certainty" (the location at which all goal cache locations became visible), then first collected goal, and finally the origin. Performance-group differences were not significant, with the exception of duration of fixations near the origin, for which top performers fixated for longer than bottom performers ($T = -3.417$, $p < 0.001$).

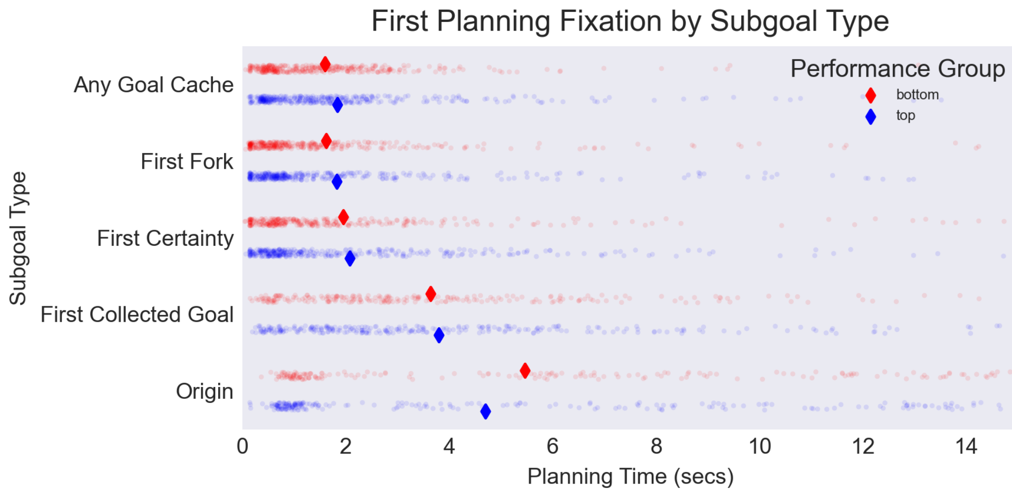


Figure A.1: Distribution of time of first fixation in subgoal regions, segmented by performance group. Diamond markers show median time in group. Median visit times indicate an overall subgoal fixation sequence as follows: any goal cache, first fork, first certainty (fully observed goal cache configuration), first collected goal, origin.

A.3 Monte Carlo Counterfactual Cycling Agent

A.3.1 Introduction

This study builds on an agent model developed in Chapter 4, Active Dynamical Prospection (ADP), which implements exploratory, planning-oriented gaze control inspired by the active inference framework (Friston et al., 2017). Consistent with active inference, a central assumption of the model explored in both Studies A and B, is that mental simulation may be leveraged to simultaneously learn, and run experiments within, a generative model of the agent's environment.

I refer to the updated agent model, which like ADP, leverages simulation for the dual goal of visual search and planning under uncertainty, as "Monte Carlo Counterfactual Cycling" (MC3). This agent constitutes the most recent effort to incorporate the theoretical conceptualization explored in this work (see introductory outline in Section 1.2.3 and returned to in 8.2). Specifically, the proposed model implements four primary theoretical attributes constitutive of cognitive claims about the mechanisms underlying planning:

1. **Landscape learning**. Planning proceeds as the iterative generation and maintenance of an affordance landscape capturing a lower-dimensional representation of the key constraints on future action.
2. **Parallelism**. Simulations run in a non-unitary, as parallel samples, through the learned landscape
3. **Ambiguity heuristic**. Epistemic value may be assessed by simulating futures and

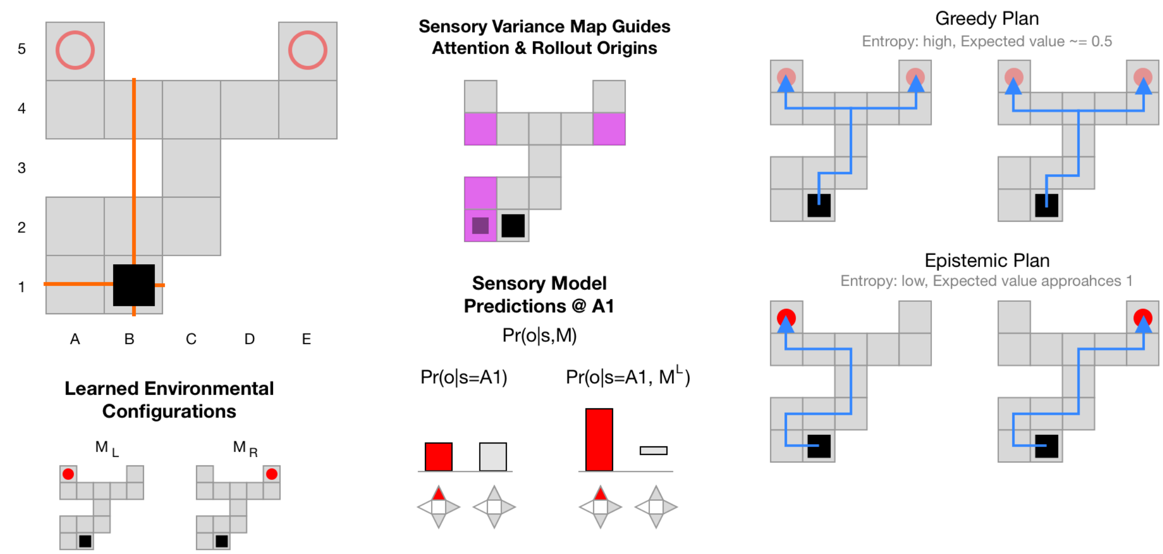


Figure A.2: Model schematic in toy epistemic navigation task.

evaluating ambiguity in multiple contingencies [1].

4. **Embodiment**. Mechanisms in line with corollary discharge recruit the motor systems into simulation both to drive prediction of action consequences, as well as to rehearse likely sensorimotor sequences.

In the next section, I illustrate some of these model dynamics with a toy example in the domain of spatial navigation.

A.3.2 Illustrative example (navigating an epistemic 'T-maze')

Imagine the altered 'T-maze' shown in Figure A.2. Our agent (black square) can move in the four compass directions to adjacent squares in this environment. Observations are 4-dimensional vectors indicating whether a reward is visible in each direction (even if multiple cells away).

Consider an agent with perfect knowledge of the spatial layout of this environment, who has learned the following statistics of reward placement: a uniform-value reward is always located at either A5 (in configuration M_L) or E5 (in M_R). We formalize the belief over these configurations as a categorical distribution $\mathfrak{B}_M = Cat(M)$. If the agent begins at B1, as shown, and must choose an action sequence that maximizes its expected discounted value (which in this case is equivalent to reaching the reward in the minimum number of steps), how would our model behave?

Since M_L and M_R are initially equally likely (under a uniform belief prior), particles rolling out on the greedy path (shown at top right) are equally likely to sample observations

[1]This is in line with previous work predicting that internally generated sequences (IGSs) will appear more frequently during times of uncertainty, and more frequently pass through more informative regions of complex environments (Pezzulo, Van der Meer, et al., 2014)

from M_L and M_R, producing observations of immediate reward experienced at the terminals (A5 and E5) which are equally distributed between 1 and 0. Observational variance, however, is not uniform across all states. Particles passing through cells in columns A and E experience high variance in sampled observations (sometimes sampling reward north, and sometimes sampling no reward north). In contrast, particles passing through states in column B, C and D experience no variance (since both models predict the same sensory observations in these states). We can visualize this as a map of observational variance predictive of possible epistemic value (see map at top center of Figure A.2).

Attention moves to A1 (an adjacent high variance state), thus biasing new particle trajectories to originate at this location. An observation is sampled (e.g. reward is visible north), which triggers a shift to the agent's belief state and a batch of particles is emitted from A1 under configuration M_L. Next, another sample is drawn (e.g. no reward north), the alternative model M_R is induced, and another batch of particles is emitted. By evaluating the aggregate sensorimotor activity constituted by each batch of trajectories (e.g. the blue lines at lower right), our model identifies that when conditioned on either observation (O_{A1} = reward north or O_{A1} = no reward north) an unambiguous policy, taking the reward-location-aware action at C4, can be obtained, and the resulting expected value approaches 1. The result of this simulation step is an update to the generative model making trajectories through A1 more likely.

A.3.3 Methods

The task under study in this chapter differs from that presented in Chapter 4 in a number of ways, summarized in table A.1 below.

Attribute	ADP	MC3 (This Work)
Mode of action	Dragging agent to nearby nodes	Continuous guidance of the agent through heading control[2]
Planning sequencing	Plan-while-acting	Planning phase followed by a navigation phase
Visual access to the environment	Scanning the landscape to uncover its attributes	Full visual access

Table A.1: Agent task comparison (ADP task in Study A versus Study B)

Despite these differences, the core learning process underlying the ADP agent could be applied to the updated task, with a variety of adjustments to overcome limitations identified in Gordon and Chuang (2021). Below I detail the updated agent model, MC3, noting changes from the original agent where relevant.

MC3 details

The agent takes a map geometry, including reward cache locations (but not true reward presence) as input, and produces two stochastic outputs: 1) a trajectory of its artificial "fovea," which can be used to produce a temporal recreation of its visual search, or a static gaze heatmap, and 2) a temporally evolving probability landscape which may be interpreted as a representation of a navigation plan through the provided map.

The agent is equipped with a working memory representation of two energy landscapes analogous to the probability landscape, and utility landscape, described in Chapter 3.

The former can be seen as a score, over tiles, representing the likelihood of occupying each spatial location in the map during navigation. The second provides a trivial prior on the value of occupying each tile based on its proximity to both the origin (higher energy), and any reward cache (lower energy). The agent's generative model avails this utility landscape, which can be conditioned on a particular configuration of known (or hypothetical) true reward locations.

Processing is discrete, and proceeds in steps until the 15-second planning phase is exhausted. The agent performs the following sequence of computations on each step:

1. Update particles (including emission of new particles, and rollouts of existing particles)
2. Learn and update energy landscape based on particle rollouts
3. Select a new fovea location to which to move

I cover each in detail below.

Particle update Particles are maintained in memory from origin, until termination, but (unlike rollout dynamics in ADP) may persist between steps. The first step of the particle update, is to emit new particles equivalent to the count of particles which terminated on the prior step.

The location of new particle emission is determined stochastically, and weighted by fixed parameters of the agent. The emission types selected from a categorical distribution include: a) the origin (p_{emit_origin}), the fovea (p_{emit_fov}), and sampled from the probability landscape (p_{emit_prob}). The latter method is performed by sampling from a softmax-weighted energy landscape (with parameterized temperature). A polarity attribute (positive or negative) is assigned to each particle on emission, and is selected stochastically as $P(polarity = positive) \sim Binomial(utility[tid])$. Polarity determines whether particle dynamics guide the particle towards higher energy (negative polarity) or lower energy (positive polarity) locations in the utility landscape. This key addition to ADP proved effective at handling the multi-goal dynamics unique to Study C, by enabling rollouts to proceed bi-directionally between multiple goal locations and the origin.

All particles additionally maintain attributes for their location, *momentum* (initialized to 1), which is reduced as a function of the present tile's energy on each movement, and a list of reward caches seen during their path to date.

Unterminated particles undergo a rollout akin to sequential Monte Carlo sampling, selecting from candidate successors (adjacent tiles) weighted by their distance from a target defined as the lowest energy tile directly and visually accessible from the particle's present location. Direct visual access is defined as true when a vector can be drawn between the origin and target tile without passing over any obstacle to navigation (gray and black tiles).

Particles terminate when either: a) their momentum is reduced to 0, b) the current tile is impassable, c) the particle has consumed *terminate_on_k_consumed* rewards, or d) the particle reaches the origin (for negative polarity particles).

Learn Learning occurs as an update to the probability landscape under the trajectory of particles which terminated on this step.

$$probability[tid] += \alpha_{learn} \cdot (energy_{target} - probability[tid])$$

Target energy, here, is a function of three values: the best accessible utility at the origin ($p_{energy[0]}$), the best accessible utility at the terminal ($p_{energy[-1]}$), and the number of rewards consumed on the particle's trajectory ($|p_{rewards\_consumed}|$). Formally:

$$energy_{target} = \begin{cases} min(|p_{rewards\_consumed}|, p_{energy[-1]}), & \text{if } p.energy[-1] \geq p.energy[0] \\ 1, & \text{otherwise} \end{cases}$$

Finally, the probability landscape is decayed multiplicatively towards maximum energy on each time step, according to parameter $decay\_rate$.

Move fovea As the final computation for each agent step, the agent selects where to move the fovea. This is performed as a softmax-weighted sample from all probability energy under particle trajectories for the current step, where higher energy (more surprising) tiles are more likely to be selected.

The artificial fovea is moved linearly to the new target location, comsuming simulation duration as per peak saccadic velocity measured empirically.

Baseline (ACO) agent

MC3 results were compared with the uniform distribution (as a naïve baseline), as well as an agent using a hierarchical variant of classical ant colony optimization pathfinding algorithms (ACO).

Hierarchical ACO pathfinding operated as follows. Ants track their origin (a goal cache location), as well as steps during rollout, and leave a "pheromone" trail indicating each of these two attributes. Ants originate at each goal cache location and move stochastically ($s_{i+1} \sim Uniform(adj(i))$) until they meet a trail from an ant from another origin, which terminates their rollout. At this point, the length of a connected (though non-optimal) path between the two caches can be calculated. After all rollouts are complete, the minimum distance between each pair of goal caches is computed. After k-iterations, the final path is chosen by selecting the sequence of at least n goals, with minimum total path distance.

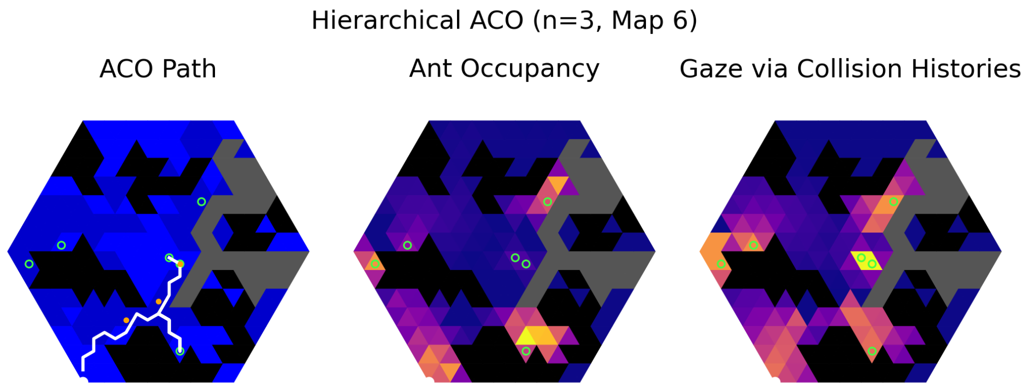


Figure A.3: Sample output from hierarchical ACO agent. Left: path selected, middle: ant occupancy, right: gaze prediction generated from ant collision history.

Gaze predictions are produced by incrementing a counter at each tile when an ant path terminates (through meeting another ant trail). See Figure A.3 for a sample output of this procedure.

Univariate optimization was used to identify parameters consistently producing solutions for all maps with $n = 3$. Final parameters used are summarized in Table A.2.

Parameter	Value
Max iterations	100
Max ant steps	50
Ants per step	30
Goal count (n)	3

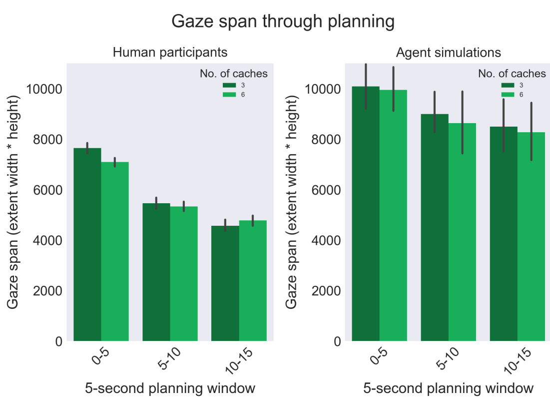
Table A.2: Parameters used for baseline ACO agent.

A.3.4 MC3-Agent results

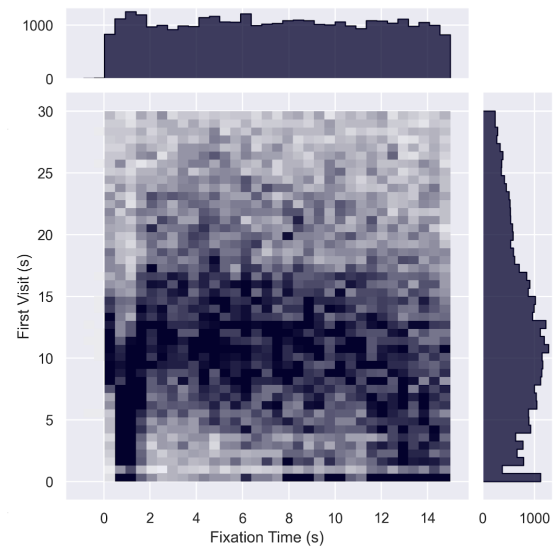
Simulations run using an agent (MC3) combining a gaze controller with a forward simulation mechanism capable of exploring sequential trajectories in parallel, demonstrate similar efficacy in visual salience prediction compared with static feature-map approaches, but with the added benefit of generating plausible spatiotemporal trajectories of visual search.

The sequential Monte Carlo-based planning agent, MC3, achieved similar levels of prediction success, in some cases exhibiting improved performance (see Figure A.5), and has the added benefit of generating spatiotemporal trajectories rather than static heatmaps. Temporal characteristics of these trajectories also show both correspondence with and divergence from human participants' behavior in various ways.

As seen in Figure A.6, the initial phase of planning shows corresponding temporal trends across all three metrics shown: sum of saccade distances and distance to nearest goal reduces, while distance from origin increases. I also find a similar, though less dramatic, decrease in overall gaze span (see Figure A.4a). However, differences between human and agent temporal trends are seen later in the planning phase. For example, I do not find the increasing distance to nearest goal (or corresponding decrease in distance from origin) towards the end of the planning phase in agent results.



(a) Reducing gaze span through planning (across all maps and participants). A corresponding, but shallower, reduction is seen in MC3 results.



(b) 2D distribution of first tile visit time (during navigation) and fixation time (during planning). A general trend is visible in which gaze initially drops to explore areas around the origin (which are, by construction, visited early on), then a broadening to explore tiles not visited until later during the navigation phase, followed by a convergence back to earlier segments on the (ultimately navigated) path.

Figure A.4: Gaze dynamics through planning.

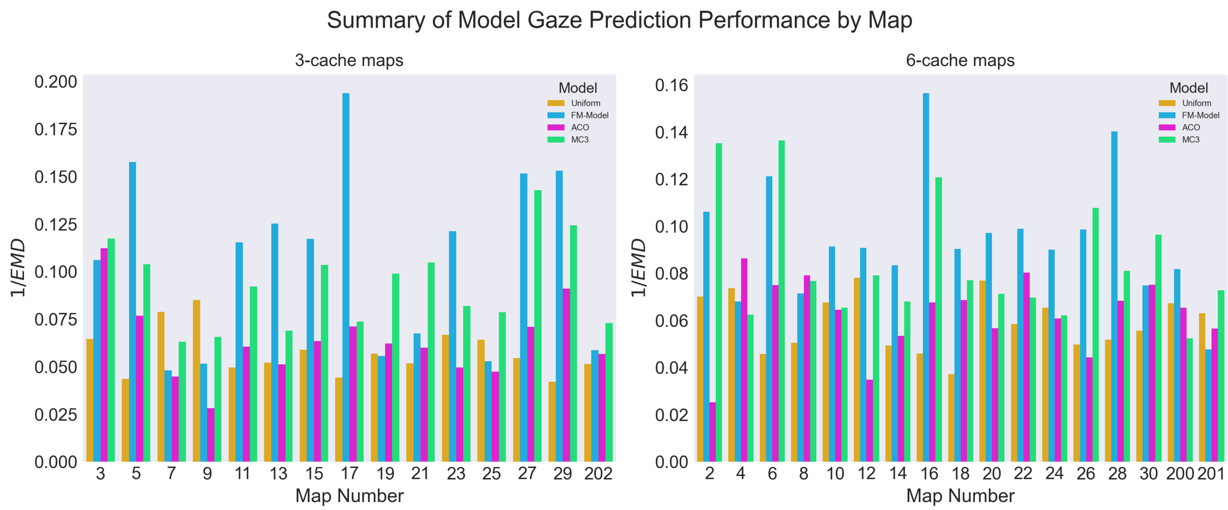


Figure A.5: Summary of models' planning gaze prediction performance by map. For each map, prediction performance (as inverse EMD) between aggregate human gaze and model prediction is plotted. FM-Model is a uniform linear combination of all feature maps. The uniform model predicts equal gaze on all tiles.

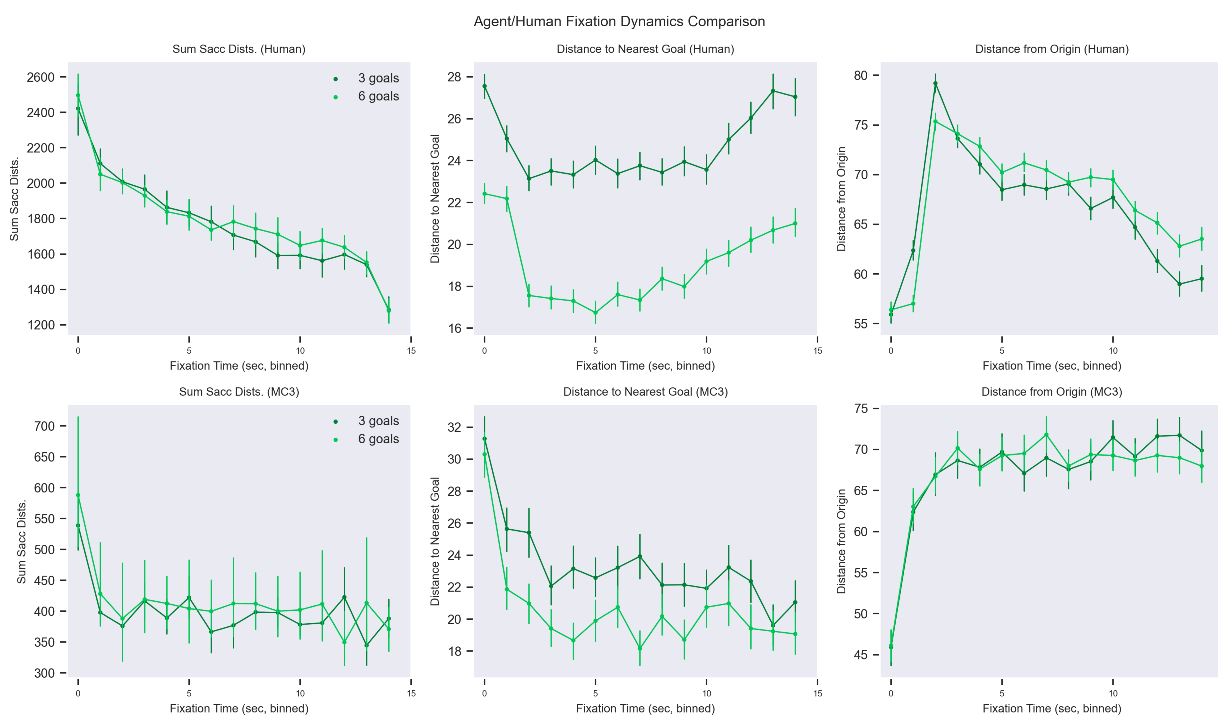


Figure A.6: Comparison of fixation dynamics through planning phase for human participants and agent simulations.

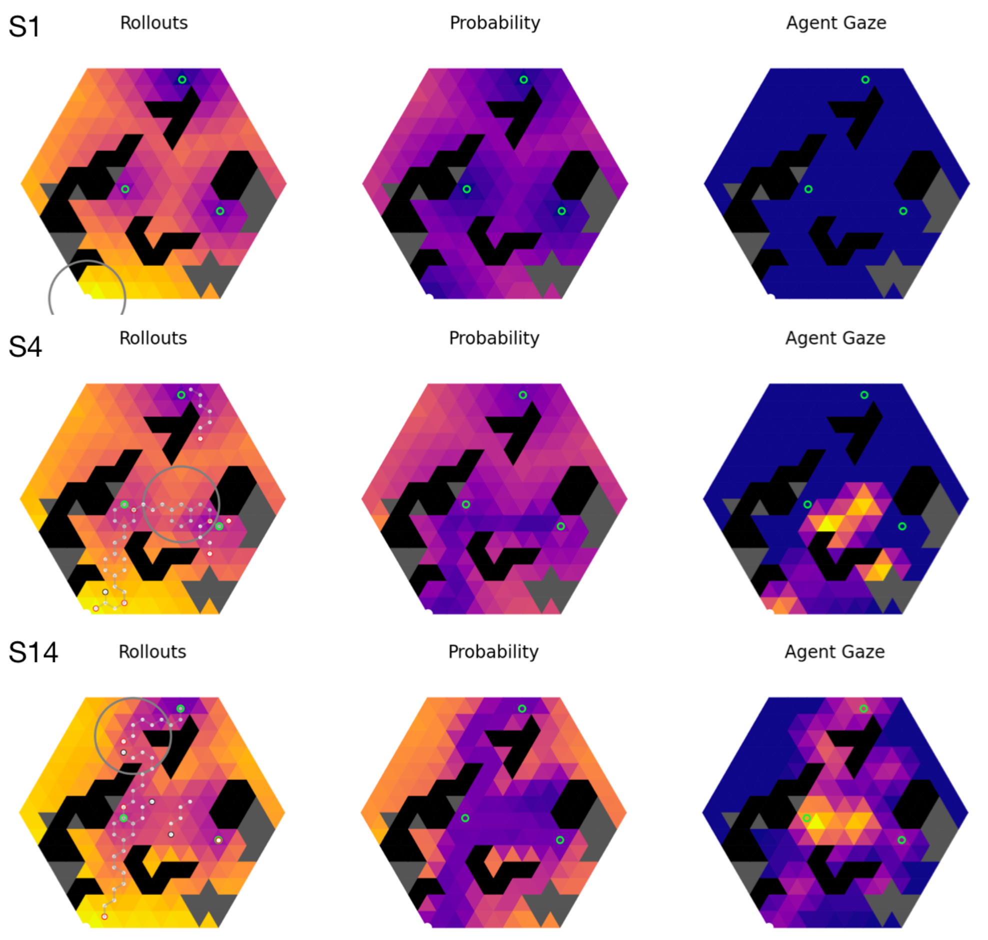


Figure A.7: Learning progression in sample MC3 simulation. White dots show particle rollout trajectories. Large gray ring is fovea location. Each row shows a snapshot of state (rollout dynamics, current learned probability landscape, and aggregate agent gaze) after 1, 4, and 14 steps of the simulation. Videos from sample agent simulations can be downloaded at the OSF project page: `https://osf.io/5tacn/`.

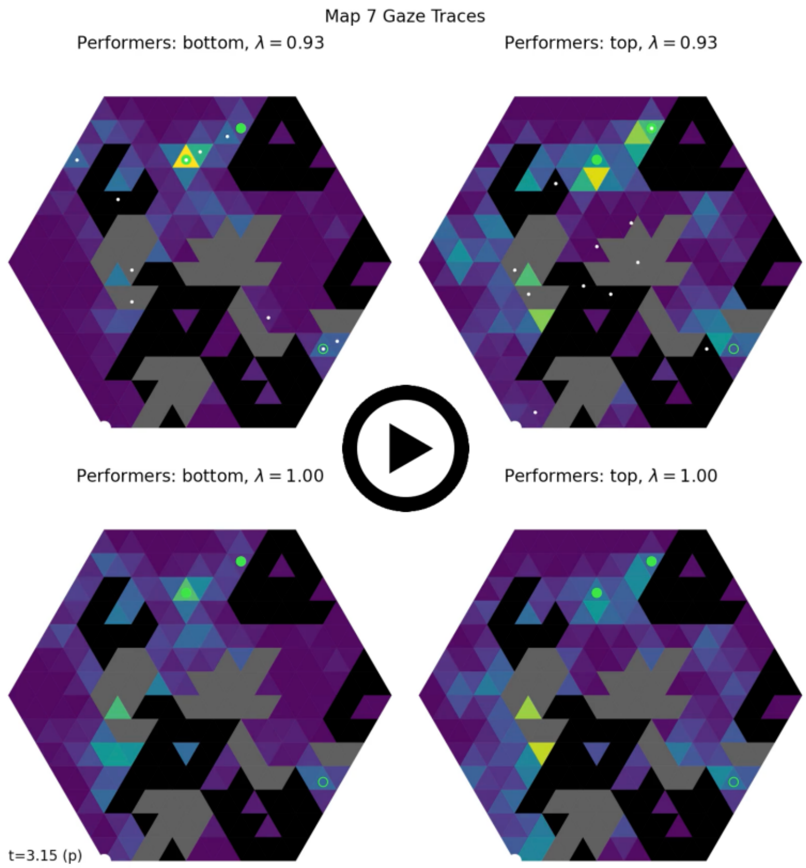


Figure A.8: Sample still frame (map 7) from aggregate human gaze data animations, which can be viewed at `https://osf.io/req2c/`. Figures show, respectively, bottom and top 33% of performers (left and right), and recent ($\lambda = 0.93$) versus aggregate ($\lambda = 1.0$) gaze (top and bottom), where λ sets the visualization decay rate or hysteresis. White dots indicate current individual gaze target during planning and transition phases, and a larger white dot indicates each participant's location during navigation.

A visualization of predicted aggregate gaze from simulated agent runs (as well as output of the feature-map model) for all maps is available at the end of the Appendix; see Figure A.10. For a summary of model performance across maps, see Figure A.5.

A.4 Supplementary figures

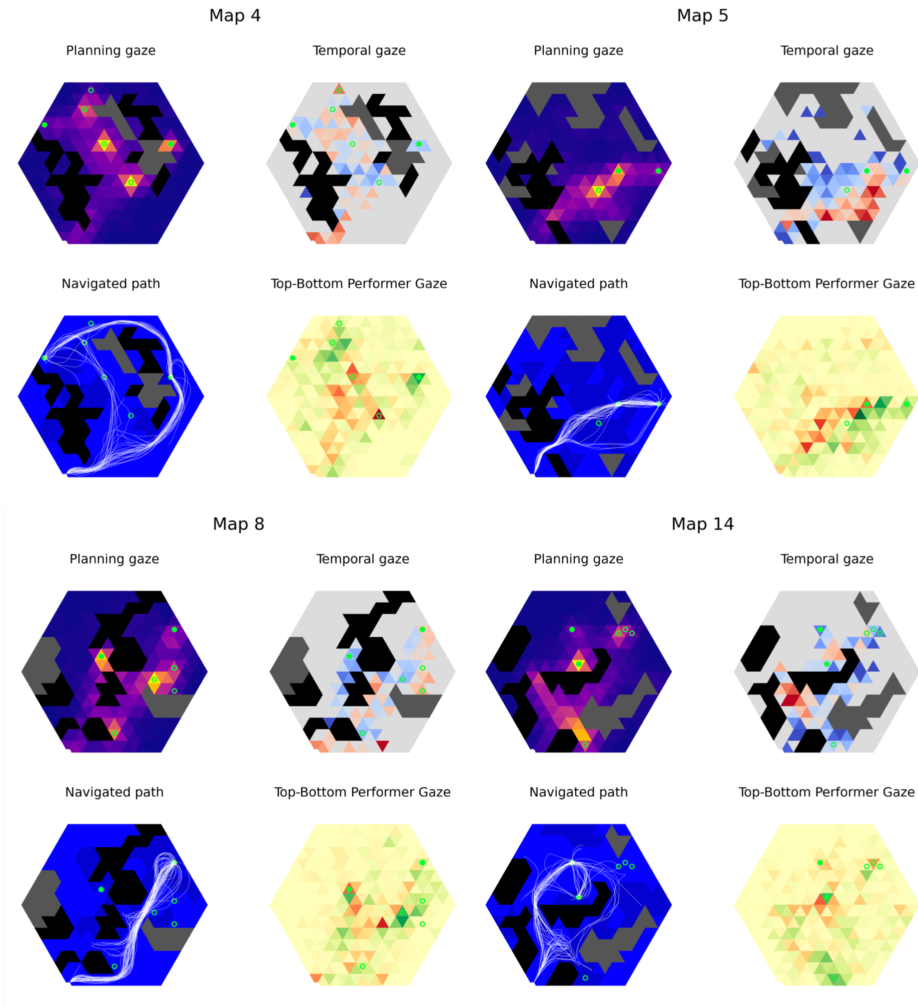


Figure A.9: For each of 4 sample maps, aggregated planning gaze (top left), all navigation trajectories (bottom left), temporal gaze trends (top right, blue tiles visited earlier, red later), and performance-group differences in gaze trends (bottom right, green tiles preferred by top-performers, red by bottom), are shown.

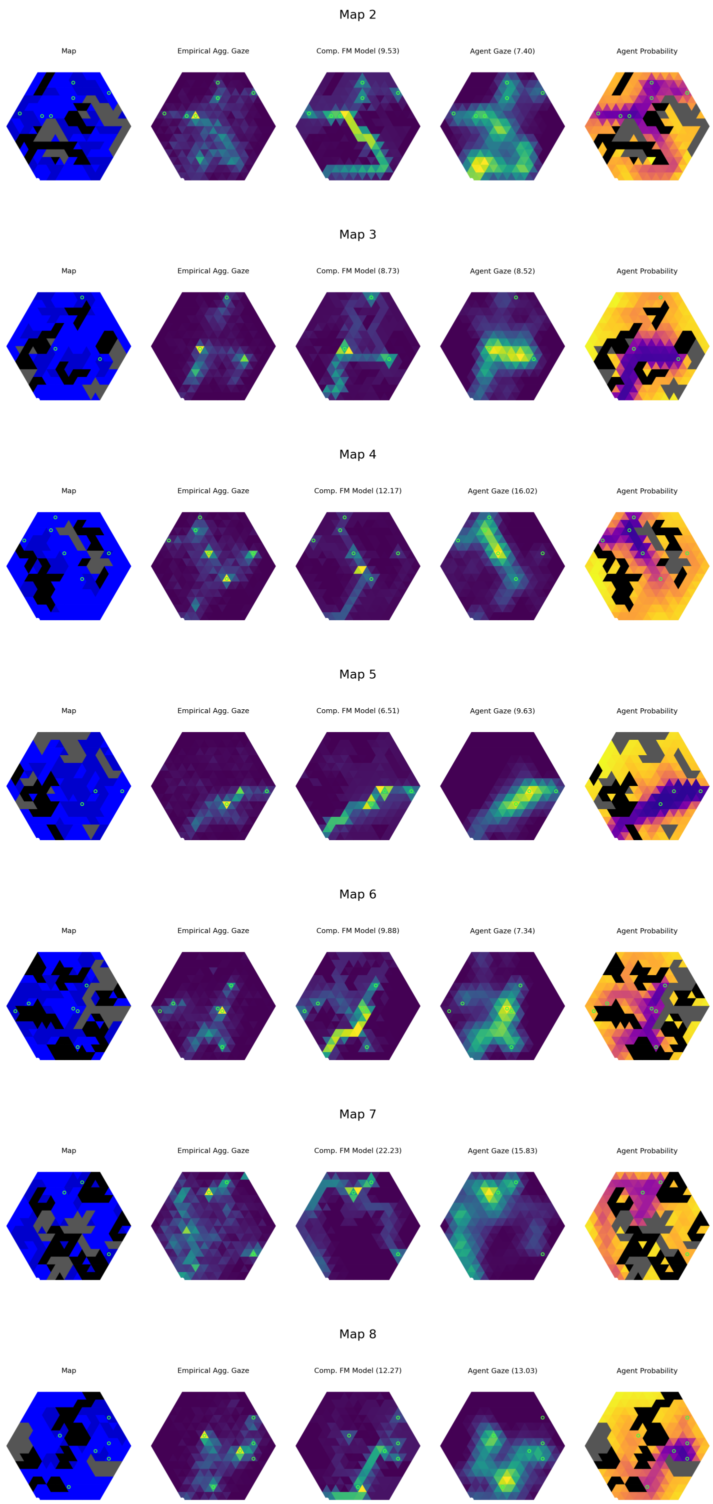


Figure A.10: Map-level gaze summary of: a) map geometry, b) aggregate human gaze, c) uniform linear combination of feature maps, d) mean gaze from agent simulations, and e) mean agent resultant probability landscape. Number in title of FM model and agent gaze reports EMD from human aggregate gaze heatmap. Maps 2-8.

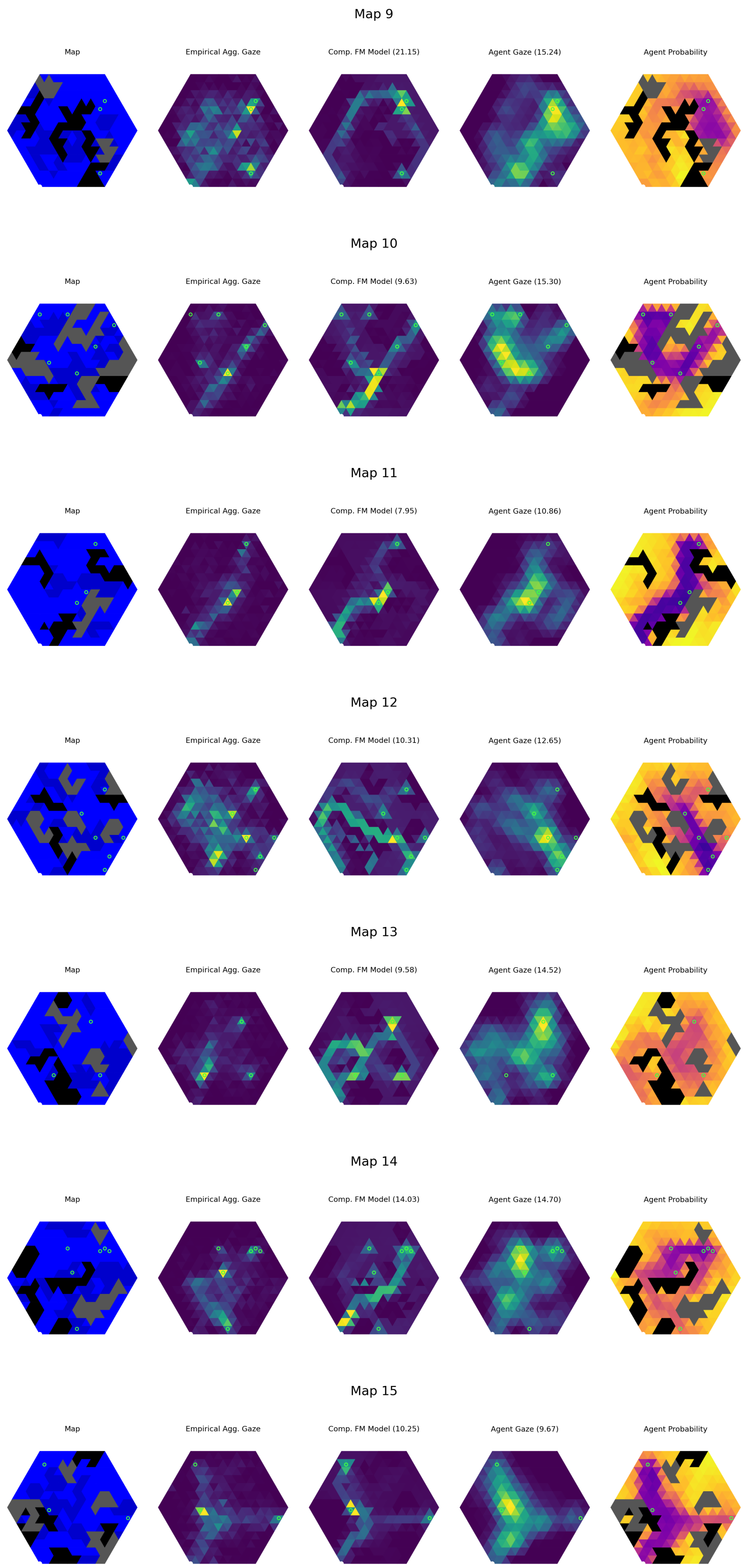


Figure A.11: Continuation of Figure A.10. Maps 9-15.

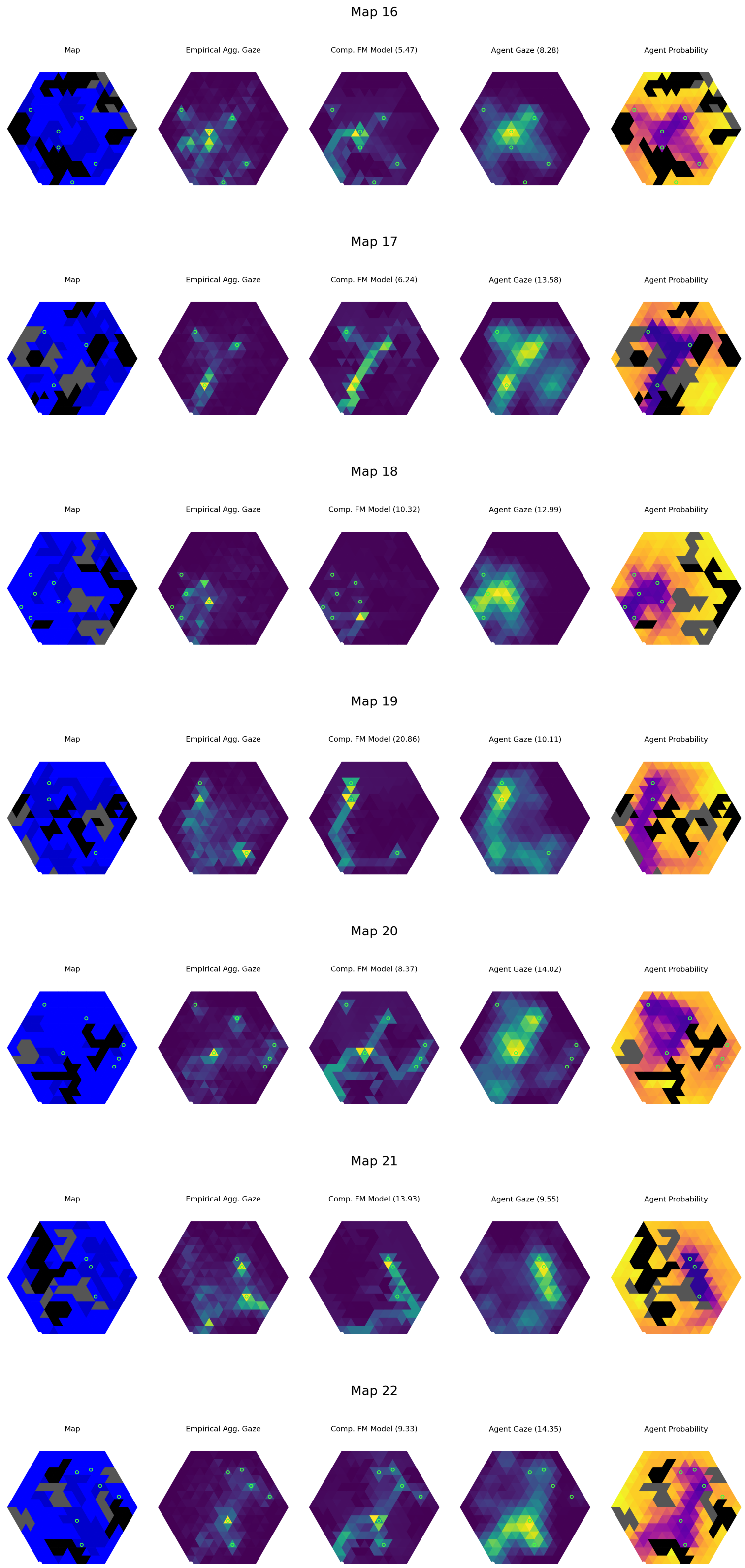


Figure A.12: Continuation of Figure A.10. Maps 16-22.

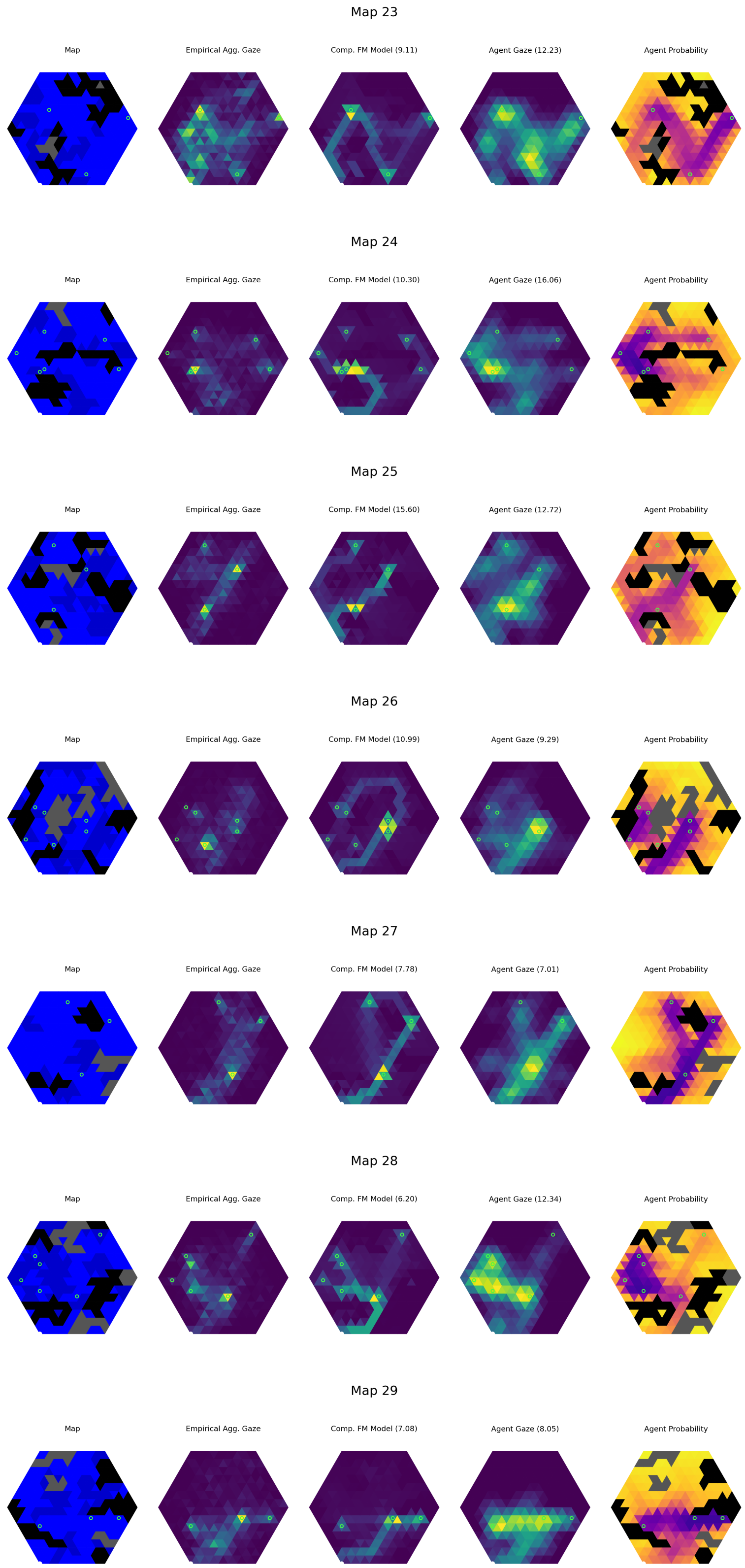


Figure A.13: Continuation of Figure A.10. Maps 23-29.

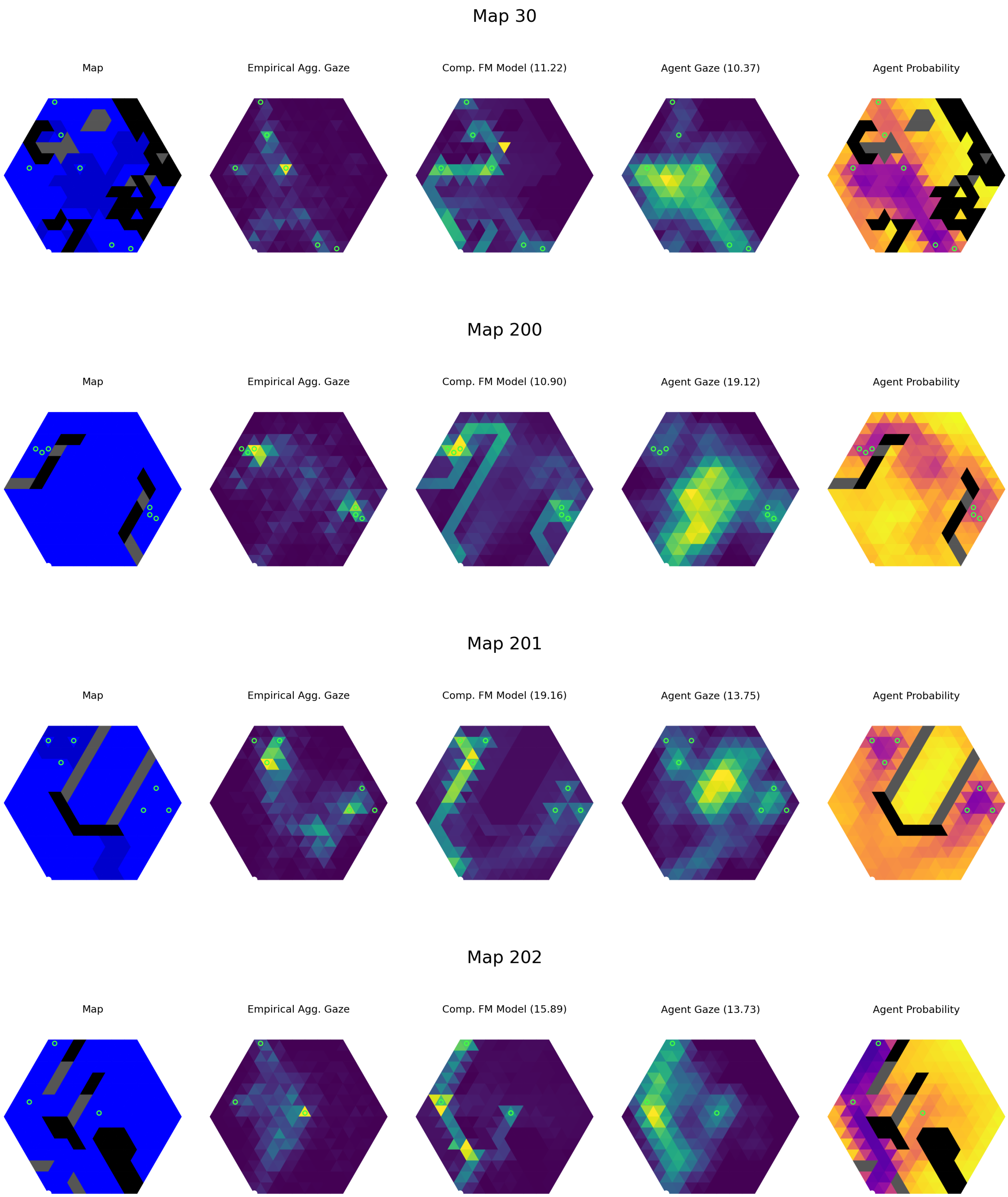


Figure A.14: Continuation of Figure A.10. Map 30 and hand designed maps 200-202.

Bibliography

Abrahamson, D., Lee, R. G., Negrete, A. G., & Gutiérrez, J. F. (2014). Coordinating visualizations of polysemous action: Values added for grounding proportion. *ZDM*, *46*, 79–93.

Amari, S. (1977). Dynamics of pattern formation in lateral inhibition type neural fields. *Biological Cybernetics*, *77–87*, 27.

Ang, J. (2021). *Planning for a joint task distribution: To be fair and timely* (Unpublished doctoral dissertation).

Anwyl-Irvine, A., Armstrong, T., & Dalmaijer, E. (2021). Mouseview. js: Reliable and valid attention tracking in web-based experiments using a cursor-directed aperture.

Apicius, A. (2016). Cookery and dining in imperial rome.

Araújo, D., Davids, K., & Hristovski, R. (2006). The ecological dynamics of decision making in sport. *Psychology of sport and exercise*, *7*(6), 653–676.

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision making*, *3*(5), 396–403.

Arnold, A. E., Iaria, G., & Ekstrom, A. D. (2016). Mental simulation of routes during navigation involves adaptive temporal compression. *Cognition*, *157*, 14–23.

Ashby, W. R. (1952). *Design for a brain* (Vol. ix). Oxford, England: Wiley.

Bahrami, B. (2012). Together, slowly but surely: The role of social interaction and feedback on the build-up of benefit in collective decision-making. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 3. (publisher: US: American Psychological Association) doi: 10.1037/a0025708

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010, 8 27). Optimally interacting minds. *Science (New York, N.Y.)*, *329*(5995), 1081–1085. doi: 10.1126/science.1185718

Bakermans, J. J. W., Warren, J., Whittington, J. C. R., & Behrens, T. E. J. (2023, April). *Constructing future behaviour in the hippocampal formation through composition and replay.* bioRxiv. Retrieved 2023-05-04, from `https://www.biorxiv.org/content/10.1101/2023.04.07.536053v1` (Pages: 2023.04.07.536053 Section: New Results) doi: 10.1101/2023.04.07.536053

Balaguer, J., Spiers, H., Hassabis, D., & Summerfield, C. (2016). Neural mechanisms of hierarchical planning in a virtual subway network. *Neuron*, *90*(4), 893–903.

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017, August). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709–721.e5. doi: 10.1016/j.neuron.2017.06.041

Banich, M. T., De La Vega, A., Andrews-Hanna, J. R., Mackiewicz Seghete, K., Du, Y.,

& Claus, E. D. (2013). Developmental trends and individual differences in brain systems involved in intertemporal choice during adolescence. *Psychology of Addictive Behaviors*, *27*(2), 416.

Bansal, S., Tolani, V., Gupta, S., Malik, J., & Tomlin, C. (2020). Combining optimal control and learning for visual navigation in novel environments. In *Conference on robot learning* (pp. 420–429).

Barca, L., & Pezzulo, G. (2012). Unfolding visual lexical decision in time. *PLoS ONE*.

Barca, L., & Pezzulo, G. (2015, 2 20). Tracking second thoughts: Continuous and discrete revision processes during visual lexical decision. *PLOS ONE*, *10*(2), e0116193. doi: 10.1371/journal.pone.0116193

Barrett, L. F., & Finlay, B. L. (2018, 12). Concepts, goals and the control of survival-related behaviors. *Current Opinion in Behavioral Sciences*, *24*, 172–179. (PMID: 31157289 PMCID: PMC6541420) doi: 10.1016/j.cobeha.2018.10.001

Barto, A. G., Sutton, R. S., & Anderson, C. W. (1983). Neuronlike adaptive elements that can solve difficult learning control problems. *IEEE transactions on systems, man, and cybernetics*(5), 834–846.

Basso, M. A., & Wurtz, R. H. (1998, 9 15). Modulation of neuronal activity in superior colliculus by changes in target probability. *Journal of Neuroscience*, *18*(18), 7519–7534.

Baumann, M. A., Fluet, M.-C., & Scherberger, H. (2009, 5 20). Context-specific grasp movement representation in the macaque anterior intraparietal area. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(20), 6436–6448. (PMID: 19458215 PMCID: PMC6665886) doi: 10.1523/JNEUROSCI.5479-08.2009

Baumeister, R. F., & Vohs, K. D. (2016). Introduction to the special issue: The science of prospection. *Review of General Psychology*, *20*(1), 1–2. doi: 10.1037/gpr0000072

Becchio, C., Sartori, L., & Castiello, U. (2010, 6 1). Toward you: The social side of actions. *Current Directions in Psychological Science*, *19*(3), 183–188. (publisher: SAGE Publications Inc) doi: 10.1177/0963721410370131

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: the hopelessness scale. *Journal of consulting and clinical psychology*, *42*(6), 861.

Beer, R. D. (2003, 12 1). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, *11*(4), 209–243. (publisher: SAGE Publications Ltd STM) doi: 10.1177/1059712303114001

Belhassein, K., Fernández-Castro, V., Mayima, A., Clodic, A., Pacherie, E., Guidetti, M., … Cochet, H. (2022, 2 1). Addressing joint action challenges in hri: Insights from psychology and philosophy. *Acta Psychologica*, *222*, 103476. doi: 10.1016/j.actpsy.2021.103476

Benoit, R. G., Szpunar, K. K., & Schacter, D. L. (2014, November). Ventromedial prefrontal cortex supports affective future simulation by integrating distributed knowledge. *PNAS*, *111*(46), 16550–16555. doi: 10.1073/pnas.1419274111

Boesch, C., & Boesch, H. (1984). Mental map in wild chimpanzees: an analysis of hammer transports for nut cracking. *Primates*, *25*, 160–170.

Bonet, B., & Geffner, H. (2013). Arguing for decisions: A qualitative model of decision making. *arXiv preprint arXiv:1302.3560*.

Boring, E. G. (2008). *History of experimental psychology.* Genesis Publishing Pvt Ltd.

Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S. S., ... Brookes, M. J. (2018, 3). Moving magnetoencephalography towards real-world applications with a wearable system. *Nature*, *555*(7698), 657–661. doi: 10.1038/nature26147

Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in cognitive sciences*, *16*(10), 485–488.

Boyer, P. (2008). Evolutionary economics of mental time travel? *Trends in cognitive sciences*, *12*(6), 219–224.

Braun, D. A., Ortega, P. A., & Wolpert, D. M. (2009, 8 14). Nash equilibria in multi-agent motor interactions. *PLOS Computational Biology*, *5*(8), e1000468. (publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1000468

Brea, J., Gaál, A. T., Urbanczik, R., & Senn, W. (2016). Prospective coding by spiking neurons. *PLoS computational biology*, *12*(6), e1005003.

Browne, C. B., Powley, E., Whitehouse, D., Lucas, S. M., Cowling, P. I., Rohlfshagen, P., ... Colton, S. (2012). A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in games*, *4*(1), 1–43.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in cognitive sciences*, *11*(2), 49–57.

Bulley, A., Henry, J., & Suddendorf, T. (2016). Prospection and the present moment: The role of episodic foresight in intertemporal choices between immediate and delayed rewards. *Review of General Psychology*, *20*(1), 29–47. doi: 10.1037/gpr0000061

Buonomano, D. (2017). *Your brain is a time machine: The neuroscience and physics of time*. WW Norton & Company.

Burk, D., Ingram, J. N., Franklin, D. W., Shadlen, M. N., & Wolpert, D. M. (2014, 3 20). Motor effort alters changes of mind in sensorimotor decision making. *PLOS ONE*, *9*(3), e92681. doi: 10.1371/journal.pone.0092681

Buzsaki, G. (2019). *The brain from inside out*. Oxford University Press, USA.

Byrne, R. M., & Girotto, V. (2012). Cognitive processes in counterfactual thinking. In *Handbook of imagination and mental simulation* (pp. 151–160). Psychology Press.

Byrne, R. M. J. (2007). *The Rational Imagination: How People Create Alternatives to Reality*. MIT Press.

Callaway, F., Lieder, F., Das, P., Gul, S., Krueger, P. M., & Griffiths, T. (2018). A resource-rational analysis of human planning. In *Cogsci.*

Candidi, M., Curioni, A., Donnarumma, F., Sacheli, L. M., & Pezzulo, G. (2015, 9 6). Interactional leader–follower sensorimotor communication strategies during repetitive joint actions. *Journal of The Royal Society Interface*, *12*(110), 20150644. (PMID: 26333815) doi: 10.1098/rsif.2015.0644

Castelfranchi, C., & Falcone, R. (2010). *Trust theory: A socio-cognitive and computational model*. Wiley.

Çatal, O., Verbelen, T., Van de Maele, T., Dhoedt, B., & Safron, A. (2021). Robot navigation as hierarchical active inference. *Neural Networks*.

Centola, D., & Baronchelli, A. (2015, 2 17). The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, *112*(7), 1989–1994. (publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.1418838112

Cernadas Curotto, P., Sander, D., d'Argembeau, A., & Klimecki, O. (2022). Back to the future: A way to increase prosocial behavior. *PloS one*, *17*(8), e0272340.

Cervone, D., D'Amour, A., Bornn, L., & Goldsberry, K. (2016, 4 2). A multiresolution stochastic process model for predicting basketball possession outcomes. *Journal of the American Statistical Association*, *111*(514), 585–599. doi: 10.1080/01621459.2016.1141685

Chackochan, V. T., & Sanguineti, V. (2019, 12 12). Incomplete information about the partner affects the development of collaborative strategies in joint action. *PLOS Computational Biology*, *15*(12), e1006385. (publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1006385

Chadalavada, R. T., Andreasson, H., Schindler, M., Palm, R., & Lilienthal, A. J. (2018). Accessing your navigation plans! human-robot intention transfer using eye-tracking glasses. In *Advances in manufacturing technology xxxii* (pp. 253–258). IOS Press.

Charnov, E. L. (1976, 4). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. doi: 10.1016/0040-5809(76)90040-X

Chaslot, G., Bakkes, S., Szita, I., & Spronck, P. (2008). Monte-carlo tree search: A new framework for game ai. In *Aiide.*

Chersi, F., Donnarumma, F., & Pezzulo, G. (2013). Mental imagery in the navigation domain: a computational model of sensory-motor simulation mechanisms. *Adaptive Behavior*, *21*(4), 251–262.

Chestek, C. A., Gilja, V., Nuyujukian, P., Kier, R. J., Solzbacher, F., Ryu, S. I., ... Shenoy, K. V. (2009, 8). Hermesc: low-power wireless neural recording system for freely moving primates. *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, *17*(4), 330–338. (PMID: 19497829) doi: 10.1109/TNSRE.2009.2023293

Chetwood, A. S., Kwok, K.-W., Sun, L.-W., Mylonas, G. P., Clark, J., Darzi, A., & Yang, G.-Z. (2012). Collaborative eye tracking: a potential training tool in laparoscopic surgery. *Surgical endoscopy*, *26*, 2003–2009.

Chouhan, S. S., & Niyogi, R. (2017, 12 1). Mapja: Multi-agent planning with joint actions. *Applied Intelligence*, *47*(4), 1044–1058. doi: 10.1007/s10489-017-0938-8

Cisek, P. (1999). Beyond the computer metaphor: Behavior as interaction. *Journal of Consciousness Studies*, *6*(12), 125–142.

Cisek, P. (2006, 9). Integrated neural processes for defining potential actions and deciding between them: a computational model. *J Neurosci*, *26*(38), 9761–9770. doi: 10.1523/JNEUROSCI.5605-05.2006

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Phil. Trans. R. Soc. B.*, *362*, 1585–1599.

Cisek, P. (2012, 12). Making decisions through a distributed consensus. *Current Opinion in Neurobiology*, *22*(6), 927–936. doi: 10.1016/j.conb.2012.05.007

Cisek, P. (2019, 6 3). Resynthesizing behavior through phylogenetic refinement. *Attention, Perception I& Psychophysics*. (PMID: 31161495) doi: 10.3758/s13414-019-01760-1

Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron*, *45*(5), 801–814.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annu Rev Neurosci*, *33*, 269–298. doi: 10.1146/annurev.neuro.051508.135409

Cisek, P., & Pastor-Bernier, A. (2014). On the challenges and mechanisms of embodied decisions. *Phil. Trans. Roy. Soc. B.*.

Cisek, P., Puskas, G. A., & El-Murr, S. (2009, 9). Decisions in changing conditions: the urgency-gating model. *J Neurosci*, *29*(37), 11560–11571. doi: 10.1523/JNEUROSCI.1844-09.2009

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *The Behavioral and brain sciences*, *36*(3), 181–204. doi: 10.1017/S0140525X12000477

Clark, A. (2015). *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. Oxford University Press. (Pages: 12-15)

Clark, A. (2018). Beyond the'bayesian blur': predictive processing and the nature of subjective experience. *Journal of Consciousness Studies*, *25*(3-4), 71–87.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clayton, N. S., Bussey, T. J., & Dickinson, A. (2003). Can animals recall the past and plan for the future? *Nature Reviews Neuroscience*, *4*(8), 685–691.

Clodic, A., Pacherie, E., Alami, R., & Chatila, R. (2017). Key elements for human robot joint action. , 159. (publisher: Springer) doi: 10.1007/978-3-319-53133-5_8

Cohen, P. R., & Levesque, H. J. (1991). Teamwork. *Nous*, *25*(4), 487–512. (publisher: JSTOR)

Cohen, R. G., & Sternad, D. (2009, 2). Variability in motor learning: relocating, channeling and reducing noise. *Experimental Brain Research*, *193*(1), 69–83. (PMID: 18953531 PMCID: PMC2756422) doi: 10.1007/s00221-008-1596-1

Collins, A. G., & Frank, M. J. (2013). Cognitive control over learning: creating, clustering, and generalizing task-set structure. *Psychological review*, *120*(1), 190.

Colorni, A., Dorigo, M., Maniezzo, V., et al. (1991). Distributed optimization by ant colonies. In *Proceedings of the first european conference on artificial life* (Vol. 142, pp. 134–142).

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468.

Cos, I., Pezzulo, G., & Cisek, P. (2021, 2 16). Changes of mind after movement onset depend on the state of the motor system. *bioRxiv*, 2021.02.15.431196. (publisher: Cold Spring Harbor Laboratory section: New Results) doi: 10.1101/2021.02.15.431196

Coutrot, A., Manley, E., Goodroe, S., Gahnstrom, C., Filomena, G., Yesiltepe, D., ... others (2022). Entropy of city street networks linked to future spatial navigation ability. *Nature*, *604*(7904), 104–110.

Cruz, F., Parisi, G. I., & Wermter, S. (2018, 7). Multi-modal feedback for affordance-driven interactive reinforcement learning. In (pp. 1–8). (ISSN: 2161-4407) doi: 10.1109/IJCNN.2018.8489237

Curioni, A., Vesper, C., Knoblich, G., & Sebanz, N. (2019, 6 1). Reciprocal information flow and role distribution support joint action coordination. *Cognition*, *187*, 21–31. doi: 10.1016/j.cognition.2019.02.006

Curioni, A., Voinov, P., Allritz, M., Wolf, T., Call, J., & Knoblich, G. (2022, 4 27). Human adults prefer to cooperate even when it is costly. *Proceedings of the Royal Society B:*

Biological Sciences, *289*(1973), 20220128. (publisher: Royal Society) doi: 10.1098/rspb.2022.0128

Cusumano, J. P., & Dingwell, J. B. (2013, 10 1). Movement variability near goal equivalent manifolds: Fluctuations, control, and model-based analysis. *Human Movement Science*, *32*(5), 899–923. doi: 10.1016/j.humov.2013.07.019

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature neuroscience*, *3*(10), 1049–1056.

Daniel, T. O., Sawyer, A., Dong, Y., Bickel, W. K., & Epstein, L. H. (2016). Remembering versus imagining: When does episodic retrospection and episodic prospection aid decision making? *Journal of Applied Research in Memory and Cognition*, *5*(3), 352–358.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007, 4). Egalitarian motives in humans. *Nature*, *446*(7137), 794–796. (number: 7137 publisher: Nature Publishing Group) doi: 10.1038/nature05651

Dayan, P. (1993, 7 1). Improving generalization for temporal difference learning: The successor representation. *Neural Computation*, *5*(4), 613–624. doi: 10.1162/neco.1993.5.4.613

Del Moral, P. (1996, 03). Non linear filtering: Interacting particle solution. *Markov Processes and Related Fields*, *2*, 555–580.

Ding, L., & Gold, J. I. (2010, 11 24). Caudate encodes multiple computations for perceptual decisions. *Journal of Neuroscience*, *30*(47), 15747–15759.

Djebbara, Z., Fich, L. B., & Gramann, K. (2021, 2 2). The brain dynamics of architectural affordances during transition. *Scientific Reports*, *11*(1), 2796. doi: 10.1038/s41598-021-82504-w

Djebbara, Z., Fich, L. B., Petrini, L., & Gramann, K. (2019, 7 16). Sensorimotor brain dynamics reflect architectural affordances. *Proceedings of the National Academy of Sciences*, *116*(29), 14769–14778. (publisher: National Academy of Sciences section: PNAS Plus PMID: 31189596) doi: 10.1073/pnas.1900648116

Donnarumma, F., Dindo, H., & Pezzulo, G. (2017). Sensorimotor communication for humans and robots: improving interactive skills by sending coordination signals. *IEEE Transactions on Cognitive and Developmental Systems*, *PP*(99), 1–1. (00000) doi: 10.1109/TCDS.2017.2756107

Donnarumma, F., Maisto, D., & Pezzulo, G. (2016). Problem solving as probabilistic inference with subgoaling: explaining human successes and pitfalls in the tower of hanoi. *PLoS computational biology*, *12*(4), e1004864.

Eluchans, M., Lancia, G. L., Maselli, A., D'Alessandro, M., Gordon, J., & Pezzulo, G. (2023). Adaptive planning depth in human problem solving. *bioRxiv*, 2023–05.

Emery, N. J., & Clayton, N. S. (2001, November). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, *414*(6862), 443–446. Retrieved 2023-07-04, from `https://www.nature.com/articles/35106560` (Number: 6862 Publisher: Nature Publishing Group) doi: 10.1038/35106560

Ersner-Hershfield, H., Wimmer, G. E., & Knutson, B. (2009). Saving for the future self: Neural measures of future self-continuity predict temporal discounting. *Social cognitive and affective neuroscience*, *4*(1), 85–92.

Etienne, S., Guthrie, M., Goillandeau, M., Nguyen, T. H., Orignac, H., Gross, C., & Boraud, T. (2014, 5 15). Easy rider: Monkeys learn to drive a wheelchair to navigate through a complex maze. *PLOS ONE*, *9*(5), e96275. (publisher: Public Library of Science) doi: 10.1371/journal.pone.0096275

Faumont, S., Lindsay, T., & Lockery, S. (2012). Neuronal microcircuits for decision making in c. elegans. *Current opinion in neurobiology*, *22*(4), 580–591.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics*, *114*(3), 817–868. (publisher: MIT press)

Fernández, J., Bornn, L., & Cervone, D. (2019). Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. In *13th mit sloan sports analytics conference.*

Flamary, R., Courty, N., Gramfort, A., Alaya, M. Z., Boisbunon, A., Chambon, S., . . . Vayer, T. (2021). Pot: Python optimal transport. *Journal of Machine Learning Research*, *22*(78), 1-8. Retrieved from http://jmlr.org/papers/v22/20-451.html

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology.* Cambridge, Mass.: MIT Press.

Forbus, K. D. (2011). Qualitative modeling. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(4), 374–391.

Friston, K., Adams, R. A., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology*, *3*(MAY), 1–20. doi: 10.3389/fpsyg.2012.00151

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2016, November). Active Inference: A Process Theory. *Neural Computation*, *29*(1), 1–49. doi: 10.1162/NECO_a_00912

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: a process theory. *Neural computation*, *29*(1), 1–49.

Furman, M., & Wang, X.-J. (2008, 12 26). Similarity effect and optimal control of multiple-choice decision making. *Neuron*, *60*(6), 1153–1168. (PMID: 19109918 PMCID: PMC2633638) doi: 10.1016/j.neuron.2008.12.003

Fyfe, S., Williams, C., Mason, O. J., & Pickup, G. J. (2008). Apophenia, theory of mind and schizotypy: perceiving meaning and intentionality in randomness. *Cortex*, *44*(10), 1316–1325.

Galantucci, B., & Garrod, S. (2011). Experimental semiotics: a review. *Front Hum Neurosci*, *5*, 11. doi: 10.3389/fnhum.2011.00011

Gallagher, S. (2007). Simulation trouble. *Social neuroscience*, *2*(3-4), 353–365.

Gallese, V. (2000). The inner sense of action. agency and motor representations. *Journal of consciousness studies*, *7*(10), 23–40.

Gallese, V. (2005). Embodied simulation: From neurons to phenomenal experience. *Phenomenology and the cognitive sciences*, *4*(1), 23–48.

Gallup Jr, G. G. (1970). Chimpanzees: self-recognition. *Science*, *167*(3914), 86–87.

Ganesh, G., & Burdet, E. (2013, 4 1). Motor planning explains human behaviour in tasks with multiple solutions. *Robotics and Autonomous Systems*, *61*(4), 362–368. doi: 10.1016/j.robot.2012.09.024

Garvert, M. M., Dolan, R. J., & Behrens, T. E. (2017). A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *Elife*, *6*, e17086.

Gerrans, P., & Sander, D. (2014). Feeling the future: Prospects for a theory of implicit prospection. *Biology & Philosophy*, *29*(5), 699–710.

Gershman, S. J. (2018, 8 15). The successor representation: Its computational logic and neural substrates. *Journal of Neuroscience*, *38*(33), 7193–7200. (PMID: 30006364) doi: 10.1523/JNEUROSCI.0151-18.2018

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological review*, *128*(5), 936.

Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2017). Eye-tracking causality. *Psychological science*, *28*(12), 1731–1744.

Gibson, E. J. (1997). An ecological psychologist's prolegomena for perceptual development: A functional approach. In *Evolving explanations of development: Ecological approaches to organism–environment systems* (pp. 23–45). Washington, DC, US: American Psychological Association. (DOI: 10.1037/10265-001)

Gibson, J. (1979). *The ecological approach to visual perception.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gibson, J. J. (1977). The theory of affordances. *Hilldale, USA*, *1*(2), 67–82.

Gibson, J. J., & Crooks, L. E. (1938). A theoretical field-analysis of automobile-driving. *The American journal of psychology*, *51*(3), 453–471.

Glimcher, P. W., & Rustichini, A. (2004, 10). Neuroeconomics: the consilience of brain and decision. *Science*, *306*(5695), 447–452. doi: 10.1126/science.1102566

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, *19*(1), 121–127.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu Rev Neurosci*, *30*, 535–574. doi: 10.1146/annurev.neuro.29.051605.113038

Gordon, J., & Chuang, J. (2021). Active dynamical prospection: Modeling mental simulation as particle filtering for sensorimotor control during pathfinding. In *Alife 2022: The 2022 conference on artificial life.*

Gordon, J., Curran, M., Chuang, J., & Cheshire, C. (2021). Covert embodied choice: Decision-making and the limits of privacy under biometric surveillance. In *Proceedings of the 2021 chi conference on human factors in computing systems.*

Gordon, J., Fiannaca, A., Kneisel, M., Cutrell, E., Miller, A., & Gonzalez-Franco, M. (2023). Hearing the way forward: Exploring ambient navigational awareness with reduced cognitive load through spatial audio-ar. In *Proceedings of the 2023 chi conference on human factors in computing systems.*

Gordon, J., Knoblich, G., & Pezzulo, G. (2023). Strategic task decomposition in joint action. *Cognitive Science*, *47*(7), e13316. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.13316` doi: https://doi.org/10.1111/cogs.13316

Gordon, J., Maselli, A., Lancia, G. L., Thiery, T., Cisek, P., & Pezzulo, G. (2021). The road towards understanding embodied decisions. *Neuroscience & Biobehavioral Reviews*, *131*, 722–736.

Grießbach, E., Incagli, F., Herbort, O., & Cañal-Bruland, R. (2021, 6 4). Body dynamics of gait affect value-based decisions. *Scientific Reports*, *11*(1), 11894. (number: 1 publisher: Nature Publishing Group) doi: 10.1038/s41598-021-91285-1

Grison, E., Gyselinck, V., Burkhardt, J.-M., & Wiener, J. M. (2017). Route planning with transportation network maps: an eye-tracking study. *Psychological research*, *81*, 1020–1034.

Grossberg, S. (1973). Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, *52*(3), 213–257. doi: 10.1002/sapm1973523213

Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science*, *16*(4), 789–802.

Guillot, A., Lebon, F., Rouffet, D., Champely, S., Doyon, J., & Collet, C. (2007). Muscular responses during motor imagery as a function of muscle contraction types. *International Journal of Psychophysiology*, *66*(1), 18–27.

Gurtner, L. M., Bischof, W. F., & Mast, F. W. (2019). Recurrence quantification analysis of eye movements during mental imagery. *Journal of vision*, *19*(1), 17–17.

Gutnick, T., Byrne, R. A., Hochner, B., & Kuba, M. (2011). Octopus vulgaris uses visual information to determine the location of its arm. *Current biology*, *21*(6), 460–462.

Ha, D., & Schmidhuber, J. (2018, March). World Models. *arXiv:1803.10122 [cs, stat]*. doi: 10.5281/zenodo.1207631

Hamrick, J. B., Smith, K. A., Griffiths, T. L., & Vul, E. (2015). Think again? the amount of mental simulation tracks uncertainty in the outcome. In *Cogsci.*

Harezlak, K., & Kasprowski, P. (2018). Application of eye tracking in medicine: A survey, research issues and challenges. *Computerized Medical Imaging and Graphics*, *65*, 176–190.

Harsh, J. A., Campillo, M., Murray, C., Myers, C., Nguyen, J., & Maltese, A. V. (2019). "seeing" data like an expert: An eye-tracking study using graphical data representations. *CBE—Life Sciences Education*, *18*(3), ar32.

Hartley, C. A., & Somerville, L. H. (2015). The neuroscience of adolescent decision-making. *Current opinion in behavioral sciences*, *5*, 108–115.

Hauskrecht, M. (2000). Value-function approximations for partially observable markov decision processes. *Journal of artificial intelligence research*, *13*, 33–94.

Hayden, B. Y., & Moreno-Bote, R. (2018, 1 1). A neuronal theory of sequential economic choice. *Brain and Neuroscience Advances*, *2*, 2398212818766675. (publisher: SAGE Publications Ltd STM) doi: 10.1177/2398212818766675

Hechter, M., & Opp, K.-D. (2001). Social norms.
(publisher: Russell Sage Foundation)

Hendriks-Jansen, H. (1996). *Catching ourselves in the act: Situated activity, interactive emergence, evolution, and human thought.* MIT Press.

Henrich, J. (2017). *The secret of our success. how our collective intelligence has helped us to evolve and prosper.* Princeton University Press. Retrieved from `https://press.princeton.edu/books/paperback/9780691178431/the-secret-of-our-success` ([Online; accessed 2022-05-24])

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., Group, C. S. R., et al. (2015). Exploration versus exploitation in space, mind, and society. *Trends in cognitive sciences*, *19*(1), 46–54.

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022).

People construct simplified mental representations to plan. *Nature*, *606*(7912), 129–136.

Hoppe, D., & Rothkopf, C. A. (2019). Multi-step planning of eye movements in visual search. *Scientific reports*, *9*(1), 1–12.

Huang, J., Velarde, I., Ma, W. J., & Baldassano, C. (2023). Schema-based predictive eye movements support sequential memory encoding. *Elife*, *12*, e82599.

Humphries, M. (2021). *The spike: An epic journey through the brain in 2.1 seconds.* Princeton University Press.

Huys, Q. J., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS computational biology*, *8*(3), e1002410.

Itai, A., Papadimitriou, C. H., & Szwarcfiter, J. L. (1982). Hamilton paths in grid graphs. *SIAM Journal on Computing*, *11*(4), 676–686.

James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). *The principles of psychology* (Vol. 1) (No. 2). Macmillan London.

Jerde, T. E., Soechting, J. F., & Flanders, M. (2003, 3 15). Coarticulation in fluent fingerspelling. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *23*(6), 2383–2393. (PMID: 12657698)

Johnson, A., & Redish, A. D. (2007, November). Neural Ensembles in CA3 Transiently Encode Paths Forward of the Animal at a Decision Point. *Journal of Neuroscience*, *27*(45), 12176–12189. doi: 10.1523/JNEUROSCI.3761-07.2007

Johnson, S. R., Tomlinson, G. A., Hawker, G. A., Granton, J. T., & Feldman, B. M. (2010). Methods to elicit beliefs for bayesian priors: a systematic review. *Journal of clinical epidemiology*, *63*(4), 355–369.

Johnson-Laird, P. N., & Byrne, R. M. (1991). *Deduction.* Lawrence Erlbaum Associates, Inc.

Joseph MacInnes, W., Hunt, A. R., Clarke, A. D., & Dodd, M. D. (2018). A generative model of cognitive state from task and eye movements. *Cognitive computation*, *10*, 703–717.

Kacelnik, A., Vasconcelos, M., Monteiro, T., & Aw, J. (2011). Darwin's "tug-of-war" vs. starlings'"horse-racing": how adaptations for sequential encounters drive simultaneous choice. *Behavioral Ecology and Sociobiology*, *65*(3), 547–558. (publisher: Springer)

Kahneman, D. (2011). *Thinking, fast and slow.* Macmillan.

Kahneman, D., & Tversky, A. (1981a, May). *The Simulation Heuristic.* (Tech. Rep. No. TR-5). STANFORD UNIV CA DEPT OF PSYCHOLOGY, STANFORD UNIV CA DEPT OF PSYCHOLOGY.

Kahneman, D., & Tversky, A. (1981b, 5 15). *The simulation heuristic* (Tech. Rep.). Retrieved from `http://www.dtic.mil/docs/citations/ADA099504`

Kakrada, E., & Colombo, M. (2022). Mirror mirror on the wall, it's not the mark i care about at all. *Learning and Motivation*, *77*, 101785.

Kaplan, R., & Friston, K. J. (2018). Planning and navigation as active inference. *Biological cybernetics*, *112*(4), 323–343.

Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2014, 3). Cortical activity in the null space: permitting preparation without movement. *Nature Neuroscience*, *17*(3), 440–448. (number: 3 publisher: Nature Publishing Group) doi: 10.1038/nn.3643

Kaufman, S. B. (2021, April). *Daniel kahneman —— a remarkable life, fast and slow.* Retrieved from `https://the-psychology-podcast-with-scott-barry-kaufman.simplecast.com/episodes/daniel-kahneman-a-remarkable-life-fast-and-slow-a7mSDESb`

Keramati, M., Smittenaar, P., Dolan, R. J., & Dayan, P. (2016). Adaptive integration of habits into depth-limited planning defines a habitual-goal–directed spectrum. *Proceedings of the National Academy of Sciences*, *113*(45), 12868–12873.

Khetarpal, K., Ahmed, Z., Comanici, G., Abel, D., & Precup, D. (2020). What can i do here? a theory of affordances in reinforcement learning. In (p. 5243–5253). PMLR.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive science*, *18*(4), 513–549.

Klein, G., & Crandall, B. (1995). The role of mental simulation in problem solving and decision making. In P. A. Hancock (Ed.), *Local applications of the ecological approach to human-machine systems.* Hillsdale, N.J: L. Erlbaum Associates.

Konvalinka, I., Sebanz, N., & Knoblich, G. (2023, 1 1). The role of reciprocity in dynamic interpersonal coordination of physiological rhythms. *Cognition*, *230*, 105307. doi: 10.1016/j.cognition.2022.105307

Kourtis, D., Sebanz, N., & Knoblich, G. (2013, 1 1). Predictive representation of other people's actions in joint action planning: An eeg study. *Social Neuroscience*, *8*(1), 31–42. doi: 10.1080/17470919.2012.694823

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, *13*(10), 1292–1298.

Krichmar, J. L., & He, C. (2023). Importance of path planning variability: A simulation study. *Topics in Cognitive Science*, *15*(1), 139–162.

Krist, H., Fieberg, E. L., & Wilkening, F. (1993). Intuitive physics in action and judgment: The development of knowledge about projectile motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 952.

Krumin, M., Lee, J. J., Harris, K. D., & Carandini, M. (2018, 11 23). Decision and navigation in mouse parietal cortex. *eLife*, *7*, e42583. (publisher: eLife Sciences Publications, Ltd) doi: 10.7554/eLife.42583

Kushnir, T., & Gopnik, A. (2005). Young children infer causal strength from probabilities and interventions. *Psychological science*, *16*(9), 678–683.

Lakshminarasimhan, K. J., Avila, E., Neyhart, E., DeAngelis, G. C., Pitkow, X., & Angelaki, D. E. (2020). Tracking the mind's eye: Primate gaze behavior during virtual visuomotor navigation reflects belief dynamics. *Neuron*, *106*(4), 662–674.

Lancia, G. L., Eluchans, M., D'Alessandro, M., Spiers, H. J., & Pezzulo, G. (2023). Humans account for cognitive costs when finding shortcuts: An information-theoretic analysis of navigation. *PLOS Computational Biology*, *19*(1), e1010829.

Lara, B., Gaona, W., Escobar, E., Pardo, J. M., & Hermosillo-Valadez, J. (2021). Development of body-based spatial knowledge through mental imagery in an artificial agent. *Adaptive Behavior*, *29*(4), 349–368.

Lee, M. D., & Cummins, T. D. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic bulletin & review*, *11*(2), 343–352.

Lepora, N. F., & Pezzulo, G. (2015, 4 7). Embodied choice: How action influences perceptual decision making. *PLoS Comput Biol*, *11*(4), e1004110. doi: 10.1371/journal.pcbi.1004110

Lewis, M., Purdy, S., Ahmad, S., & Hawkins, J. (2019). Locations in the neocortex: a theory of sensorimotor object recognition using cortical grid cells. *Frontiers in neural circuits*, *13*, 22.

Marcos, E., Cos, I., Girard, B., & Verschure, P. F. M. J. (2015, 12 16). Motor cost influences perceptual decisions. *PLoS ONE*, *10*(12), e0144841. doi: 10.1371/journal.pone.0144841

Maselli, A., Gordon, J., & Pezzulo, G. (n.d.). *Expertise and the benefit of externalized affordance landscapes in climbing training.* (unpublished)

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: the realization of living.* Dordrecht, Holland: D. Reidel Pub.

McCloskey, D. (2011). Corollary discharges: motor commands and perception. *Comprehensive physiology*, 1415–1447.

McNamee, D., Wolpert, D. M., & Lengyel, M. (2016). Efficient state-space modularization for planning: theory, behavioral and neural signatures. *Advances in neural information processing systems*, *29*.

Meyer, M., van der Wel, R. P. R. D., & Hunnius, S. (2013, 4). Higher-order action planning for individual and joint object manipulations. *Experimental Brain Research*, *225*(4), 579–588. (PMID: 23361302) doi: 10.1007/s00221-012-3398-8

Michalski, J., Green, A. M., & Cisek, P. (2020, 2 12). Reaching decisions during ongoing movements. *Journal of Neurophysiology*, *123*(3), 1090–1102. (publisher: American Physiological Society) doi: 10.1152/jn.00613.2019

Miloyan, B., Pachana, N. A., & Suddendorf, T. (2014). The future is here: A review of foresight systems in anxiety and depression. *Cognition & emotion*, *28*(5), 795–810.

Mink, J. W. (1996, 11). The basal ganglia: focused selection and inhibition of competing motor programs. *Progress in Neurobiology*, *50*(4), 381–425. (PMID: 9004351)

Momennejad, I., & Howard, M. W. (2018). Predicting the future with multi-scale successor representations. *BioRxiv*, 449470.

Momennejad, I., Russek, E. M., Cheong, J. H., Botvinick, M. M., Daw, N. D., & Gershman, S. J. (2017, 9). The successor representation in human reinforcement learning. *Nature Human Behaviour*, *1*(9), 680–692. (number: 9 publisher: Nature Publishing Group) doi: 10.1038/s41562-017-0180-8

Montello, D. R. (2005). *Navigation.* Cambridge University Press.

Moreno-Bote, R., Ramírez-Ruiz, J., Drugowitsch, J., & Hayden, B. Y. (2020). Heuristics and optimal solutions to the breadth–depth dilemma. *Proceedings of the National Academy of Sciences*, *117*(33), 19799–19808.

Moser, E. I., Kropff, E., & Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, *31*, 69–89.

Müller, H., & Sternad, D. (2004, 2). Decomposition of variability in the execution of goal-oriented tasks: three components of skill improvement. *Journal of Experimental*

Psychology. Human Perception and Performance, *30*(1), 212–233. (PMID: 14769078) doi: 10.1037/0096-1523.30.1.212

Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature reviews neuroscience*, *10*(5), 360–372.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, N.J: Prentice-Hall.

Newell, A., Simon, H. A., et al. (1972). *Human problem solving* (Vol. 104) (No. 9). Prentice-hall Englewood Cliffs, NJ.

Nilsson, D.-E. (2009). The evolution of eyes and visually guided behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1531), 2833–2847.

Niv, Y. (2020). The primacy of behavioral research for understanding the brain.

Noble, S. U. (2018). Algorithms of oppression. In *Algorithms of oppression.* New York university press.

Nolfi, S. (2009). Behavior and cognition as a complex adaptive system: Insights from robotic experiments. *Philosophy of Complex Systems, Handbook on Foundational/Philosophical Issues for Complex Systems in Science. Elsevier*. Retrieved from `http://www.tech.plym.ac.uk/SoCCE/ITALK/publications/Nolfi_behaviour_as_cas_elsevier_handbook.pdf` ([Online; accessed 2014-09-16])

Nolfi, S., & Floreano, D. (2001). *Evolutionary robotics. the biology, intelligence, and technology of self-organizing machines* (Tech. Rep.). Retrieved from `http://infoscience.epfl.ch/record/63929` ([Online; accessed 2014-09-11])

Ortega, P. A., & Braun, D. A. (2013, 8 5). Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, *469*(2153). Retrieved from `http://rspa.royalsocietypublishing.org/content/469/2153/20120683` ([Online; accessed 2013-07-11]) doi: 10.1098/rspa.2012.0683

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107*, 179–217.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: A good-based model. *Annual Review of Neuroscience*, *34*, 333–359. doi: 10.1146/annurev-neuro-061010-113648

Padoa-Schioppa, C., & Assad, J. A. (2006, 5). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. doi: 10.1038/nature04676

Passingham, R. E., & Wise, S. P. (2012). *The neurobiology of the prefrontal cortex: Anatomy, evolution, and the origin of insight* (1st ed.). Oxford University Press, USA.

Pastor-Bernier, A., & Cisek, P. (2011, 5 11). Neural correlates of biased competition in premotor cortex. *Journal of Neuroscience*, *31*(19), 7083–7088.

Pastor-Bernier, A., Tremblay, E., & Cisek, P. (2012). Dorsal premotor cortex is involved in switching motor plans. *Frontiers in Neuroengineering*, *5*, 5. (PMID: 22493577 PMCID: PMC3318308) doi: 10.3389/fneng.2012.00005

Pearl, J., & Mackenzie, D. (2018). *The book of why: the new science of cause and effect.* Basic books.

Perrone-Bertolotti, M., Kujala, J., Vidal, J. R., Hamame, C. M., Ossandon, T., Bertrand, O., . . . Lachaux, J.-P. (2012). How silent is silent reading? intracerebral evidence for top-down activation of temporal voice areas during reading. *Journal of Neuroscience*, *32*(49), 17554–17562.

Pezzulo, G., Barca, L., Bocconi, A. L., & Borghi, A. M. (2010). When affordances climb into your mind: Advantages of motor simulation in a memory task performed by novice and expert rock climbers. *Brain and Cognition*, *73*(1), 68–73.

Pezzulo, G., Candidi, M., Dindo, H., & Barca, L. (2013). Action simulation in the human brain: twelve questions. *New Ideas in Psychology*, *31*(3), 270–290.

Pezzulo, G., & Cisek, P. (2016). Navigating the affordance landscape: Feedback control as a process model of behavior and cognition. *Trends in Cognitive Sciences*. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1364661316300067` ([Online; accessed 2016-04-26]) doi: 10.1016/j.tics.2016.03.013

Pezzulo, G., & Dindo, H. (2011). What should i do next? using shared representations to solve interaction problems. *Experimental Brain Research*, *211*(3), 613–630.

Pezzulo, G., Donnarumma, F., & Dindo, H. (2013, 11 20). Human sensorimotor communication: A theory of signaling in online social interactions. *PLoS ONE*, *8*(11), e79876. doi: 10.1371/journal.pone.0079876

Pezzulo, G., Donnarumma, F., Dindo, H., D'Ausilio, A., Konvalinka, I., & Castelfranchi, C. (2018, 6 30). The body talks: Sensorimotor communication and its brain and kinematic signatures. *Physics of Life Reviews*. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1571064518300770` ([Online; accessed 2018-12-13]) doi: 10.1016/j.plrev.2018.06.014

Pezzulo, G., Donnarumma, F., Maisto, D., & Stoianov, I. (2019). Planning at decision time and in the background during spatial navigation. *Current opinion in behavioral sciences*, *29*, 69–76.

Pezzulo, G., Kemere, C., & Van Der Meer, M. A. (2017). Internally generated hippocampal sequences as a vantage point to probe future-oriented cognition. *Annals of the New York Academy of Sciences*, *1396*(1), 144–165.

Pezzulo, G., & Ognibene, D. (2011). Proactive action preparation: Seeing action preparation as a continuous and proactive process. *Motor Control*.

Pezzulo, G., Parr, T., & Friston, K. J. (2021). The evolution of brain architectures for predictive coding and active inference. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Progress in neurobiology*, *134*, 17–35.

Pezzulo, G., Roche, L., & Saint-Bauzel, L. (2021, 1 13). Haptic communication optimises joint decisions and affords implicit confidence sharing. *Scientific Reports*, *11*(1), 1051. doi: 10.1038/s41598-020-80041-6

Pezzulo, G., van der Meer, M. A. A., Lansink, C. S., & Pennartz, C. M. A. (2014, December). Internally generated sequences in learning and executing goal-directed behavior. *Trends in Cognitive Sciences*, *18*(12), 647–657. doi: 10.1016/j.tics.2014.06.011

Pezzulo, G., Van der Meer, M. A., Lansink, C. S., & Pennartz, C. M. (2014). Internally generated sequences in learning and executing goal-directed behavior. *Trends in cognitive sciences*, *18*(12), 647–657.

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74–79.

Piaget, J. (1952). *The origins of intelligence in children.* New York: International Universities Press. (Originally published 1936)

Platt, M. L., & Glimcher, P. W. (1999, 7). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233–238. doi: 10.1038/22268

Pohl, I. (1971). Bi-directional search. *Machine intelligence*, *6*, 127–140.

Pomplun, M., Reingold, E. M., & Shen, J. (2001). Investigating the visual span in comparative search: The effects of task difficulty and divided attention. *Cognition*, *81*(2), B57–B67.

Powers, W. T. (1973). *Behavior: The control of perception.* Hawthorne, NY: Aldine.

Proffitt, D. R. (2006, 6 1). Embodied perception and the economy of action. *Perspectives on Psychological Science*, *1*(2), 110–122. doi: 10.1111/j.1745-6916.2006.00008.x

Pylyshyn, Z. (1984). *Computation and cognition: Toward a foundation for cognitive science.* MIT Press.

Quelhas, A. C., & Byrne, R. (2003). Reasoning with deontic and counterfactual conditionals. *Thinking & Reasoning*, *9*(1), 43–65.

Raby, C., & Clayton, N. S. (2009). Prospective cognition in animals. *Behavioural processes*, *80*(3), 314–324.

Ramirez, M., & Geffner, H. (2010). Probabilistic plan recognition using off-the-shelf classical planners. Atlanta, USA..

Rand, D. G., & Nowak, M. A. (2013, 8 1). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413–425. doi: 10.1016/j.tics.2013.06.003

Ranney, M. A., Munnich, E. L., & Lamprey, L. N. (2016). Increased wisdom from the ashes of ignorance and surprise: Numerically-driven inferencing, global warming, and other exemplar realms. In *Psychology of learning and motivation* (Vol. 65, pp. 129–182). Elsevier.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016, 4 1). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260–281. doi: 10.1016/j.tics.2016.01.007

Rens, N., Lancia, G. L., Eluchans, M., Schwartenbeck, P., Cunnington, R., & Pezzulo, G. (2023). Evidence for entropy maximisation in human free choice behaviour. *Cognition*, *232*, 105328.

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009, 8). Changes of mind in decision-making. *Nature*, *461*(7261), 263–266. doi: 10.1038/nature08275

Rietveld, E., & Kiverstein, J. (2014, 10 2). A rich landscape of affordances. *Ecological Psychology*, *26*(4), 325–352. (00104) doi: 10.1080/10407413.2014.958035

Roberts, S. F., Koditschek, D. E., & Miracchi, L. J. (2020). Examples of gibsonian affordances in legged robotics research using an empirical, generative framework. *Frontiers in Neurorobotics*, *14*. Retrieved from `https://www.frontiersin.org/articles/10.3389/fnbot.2020.00012/full` (publisher: Frontiers) doi: 10.3389/fnbot.2020.00012

Rodríguez, F., López, J. C., Vargas, J. P., Gómez, Y., Broglio, C., & Salas, C. (2002, 4 1). Conservation of spatial memory function in the pallial forebrain of reptiles and ray-finned fishes. *Journal of Neuroscience*, *22*(7), 2894–2903.

Roepke, A. M., & Seligman, M. E. (2016). Depression and prospection. *British Journal of Clinical Psychology*, *55*(1), 23–48.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 10684–10695).

Rosenbaum, D. A., Chapman, K. M., Weigelt, M., Weiss, D. J., & van der Wel, R. (2012, 9). Cognition, action, and object manipulation. *Psychological bulletin*, *138*(5), 924–946. (PMID: 22448912 PMCID: PMC3389205) doi: 10.1037/a0027839

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS computational biology*, *13*(9), e1005768.

Russell, S. (2023). *How not to destroy the world with ai.* Retrieved from https://www.youtube.com/watch?v=ISkAkiAkK7A (UC Berkeley CITRIS Research Exchange)

Russell, S., & Norvig, P. (1995). A modern, agent-oriented approach to introductory artificial intelligence. *Acm Sigart Bulletin*, *6*(2), 24–26.

Sacheli, L. M., Arcangeli, E., & Paulesu, E. (2018, 3 22). Evidence for a dyadic motor plan in joint action. *Scientific Reports*, *8*(1), 5027. (number: 1 publisher: Nature Publishing Group) doi: 10.1038/s41598-018-23275-9

Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E. L., ... others (2022). Photorealistic text-to-image diffusion models with deep language understanding. *Advances in Neural Information Processing Systems*, *35*, 36479–36494.

Saitoh, K., Ménard, A., & Grillner, S. (2007, 4). Tectal control of locomotion, steering, and eye movements in lamprey. *Journal of Neurophysiology*, *97*(4), 3093–3108. (PMID: 17303814) doi: 10.1152/jn.00639.2006

Sandamirskaya, Y., Zibner, S. K., Schneegans, S., & Schöner, G. (2013). Using dynamic field theory to extend the embodiment stance toward higher cognition. *New Ideas in Psychology*, *31*(3), 322–339.

Santos-Pata, D., & Verschure, P. F. M. J. (2018). Human Vicarious Trial and Error Is Predictive of Spatial Navigation Performance. *Frontiers in Behavioral Neuroscience*, *12*. Retrieved 2019-11-12, from https://www.frontiersin.org/articles/10.3389/fnbeh.2018.00237/full doi: 10.3389/fnbeh.2018.00237

Sartori, L., Straulino, E., & Castiello, U. (2011, 9 28). How objects are grasped: The interplay between affordances and end-goals. *PLOS ONE*, *6*(9), e25203. (publisher: Public Library of Science) doi: 10.1371/journal.pone.0025203

Schaal, S., Atkeson, C. G., & Sternad, D. (1996). One-handed juggling: A dynamical approach to a rhythmic movement task. *Journal of motor behavior*, *28*(2), 165–183.

Schaal, S., Mohajerian, P., & Ijspeert, A. (2007). Dynamics systems vs. optimal control—a unifying view. *Progress in brain research*, *165*, 425–445.

Schelling, T. C. (1980). *The strategy of conflict.* Harvard University Press.

Schmidhuber, J. (2015). On learning to think: Algorithmic information theory for novel combinations of reinforcement learning controllers and recurrent neural world models. *arXiv preprint arXiv:1511.09249*.

Schneegans, S., & Schöner, G. (2008, 1 1). 13 - dynamic field theory as a framework for understanding embodied cognition. In P. Calvo & A. Gomila (Eds.), *Handbook of cognitive science* (pp. 241–271). San Diego: Elsevier. Retrieved from https://www.sciencedirect.com/science/article/pii/B978008046616300013X (DOI: 10.1016/B978-0-08-046616-3.00013-X)

Scholz, J. P., & Schöner, G. (1999, 5 1). The uncontrolled manifold concept: identifying control variables for a functional task. *Experimental Brain Research*, *126*(3), 289–306. doi: 10.1007/s002210050738

Schwarz, D. A., Lebedev, M. A., Hanson, T. L., Dimitrov, D. F., Lehew, G., Meloy, J., ... Nicolelis, M. A. L. (2014, 6). Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. *Nature Methods*, *11*(6), 670–676. doi: 10.1038/nmeth.2936

Sebanz, N., Bekkering, H., & Knoblich, G. (2006, 2). Joint action: bodies and minds moving together. *Trends Cogn Sci*, *10*(2), 70–76. doi: 10.1016/j.tics.2005.12.009

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*, 353–367.

Seligman, M. E. P., Railton, P., Baumeister, R. F., & Sripada, C. (2013, March). Navigating Into the Future or Driven by the Past. *Perspect Psychol Sci*, *8*(2), 119–141. doi: 10.1177/1745691612474317

Sennesh, E., Theriault, J., Brooks, D., van de Meent, J.-W., Barrett, L. F., & Quigley, K. S. (2022). Interoception as modeling, allostasis as control. *Biological Psychology*, *167*, 108242.

Seth, A. K. (2014). *The cybernetic bayesian brain.* Open MIND. Frankfurt am Main: MIND Group.

Shadlen, M., Kiani, R., Hanks, T., & Churchland, A. (2008). Neurobiology of decision making: An intentional framework. In C. Engel & W. Singer (Eds.), *Better than conscious?: decision making, the human mind, and implications for institutions.* The MIT Press.

Shadlen, M. N., & Newsome, W. T. (2001, 10). Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *J Neurophysiol*, *86*(4), 1916–1936.

Shadmehr, R., Huang, H. J., & Ahmed, A. A. (2016, 7 25). A representation of effort in decision-making and motor control. *Current biology: CB*, *26*(14), 1929–1934. (PMID: 27374338) doi: 10.1016/j.cub.2016.05.065

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703.

Shimojima, A., & Katagiri, Y. (2013). An eye-tracking study of exploitations of spatial constraints in diagrammatic reasoning. *Cognitive science*, *37*(2), 211–254.

Shum, M., Kleiman-Weiner, M., Littman, M. L., & Tenenbaum, J. B. (2019, 1 27). Theory of minds: understanding behavior in groups through inverse planning. In (p. 6163–6370). Honolulu, Hawaii, USA: AAAI Press. Retrieved from `https://doi.org/10.1609/aaai.v33i01.33016163` ([Online; accessed 2022-05-24]) doi: 10.1609/aaai.v33i01.33016163

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... others (2016). Mastering the game of go with deep neural networks and tree search. *nature*, *529*(7587), 484–489.

Silver, D., & Veness, J. (2010). Monte-carlo planning in large pomdps..

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review*, *63*(2), 129.

Sims, T., Raposo, S., Bailenson, J. N., & Carstensen, L. L. (2020). The future is now: Age-progressed images motivate community college students to prepare for their financial futures. *Journal of Experimental Psychology: Applied*, *26*(4), 593.

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *Neuroimage*, *25*(1), 76–89.

Sodagar, A. M., Wise, K. D., & Najafi, K. (2007, 6). A fully integrated mixed-signal neural processor for implantable multichannel cortical recording. *IEEE Transactions on Biomedical Engineering*, *54*(6), 1075–1088. (event: IEEE Transactions on Biomedical Engineering) doi: 10.1109/TBME.2007.894986

Solway, A., Diuk, C., Córdova, N., Yee, D., Barto, A. G., Niv, Y., & Botvinick, M. M. (2014). Optimal behavioral hierarchy. *PLoS computational biology*, *10*(8), e1003779.

Song, J.-H., & Nakayama, K. (2009, 8). Hidden cognitive states revealed in choice reaching tasks. *Trends Cogn Sci*, *13*(8), 360–366. doi: 10.1016/j.tics.2009.04.009

Sperber, D., & Hirschfeld, L. A. (2004, 1). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences*, *8*(1), 40–46. (PMID: 14697402) doi: 10.1016/j.tics.2003.11.002

Spivey, M. (2007). *The continuity of mind.* Oxford University Press, USA.

Spivey, M., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences of the USA*, *102*, 10393–10398.

Spivey, M. J., & Dale, R. (2006, 10 1). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science*, *15*(5), 207–211. doi: 10.1111/j.1467-8721.2006.00437.x

Spivey, M. J., & Geng, J. J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological research*, *65*(4), 235–241.

Standage, D., You, H., Wang, D.-H., & Dorris, M. C. (2011). Gain modulation by an urgency signal controls the speed-accuracy trade-off in a network model of a cortical decision circuit. *Frontiers in Computational Neuroscience*, *5*, 7. (PMID: 21415911 PMCID: PMC3042674) doi: 10.3389/fncom.2011.00007

Steinberg, L., Graham, S., O'brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child development*, *80*(1), 28–44.

Sternad, D. (2018, 4). It's not (only) the mean that matters: Variability, noise and exploration in skill learning. *Current opinion in behavioral sciences*, *20*, 183–195. (PMID: 30035207 PMCID: PMC6051545) doi: 10.1016/j.cobeha.2018.01.004

Stoianov, I., Maisto, D., & Pezzulo, G. (2020). The hippocampal formation as a hierarchical generative model supporting generative replay and continual learning. *bioRxiv*.

Strachan, J. W., & Török, G. (2020). Efficiency is prioritised over fairness when distributing joint actions. *Acta Psychologica*, *210*, 103158. (publisher: Elsevier)

Striedter, G. F., & Northcutt, R. G. (2019). *Brains through time: a natural history of vertebrates.* Oxford University Press.

Sutton, R., & Barto, A. (1998a). *Reinforcement learning: An introduction.* Cambridge MA: MIT Press.

Sutton, R., & Barto, A. (1998b). *Reinforcement Learning: An Introduction.* MIT Press.

Sutton, R. S. (1991). Dyna, an integrated architecture for learning, planning, and reacting. *ACM Sigart Bulletin*, *2*(4), 160–163.

Szpunar, K. K., Spreng, R. N., & Schacter, D. L. (2014). A taxonomy of prospection:

Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences*, *111*(52), 18414–18421.

Tambe, M. (1997). Towards flexible teamwork. *Journal of artificial intelligence research*, *7*, 83–124.

Tancredi, S., Abdu, R., Abrahamson, D., & Balasubramaniam, R. (2021). Modeling non-linear dynamics of fluency development in an embodied-design mathematics learning environment with recurrence quantification analysis. *International Journal of Child-Computer Interaction*, *29*, 100297.

Taylor, S. E., Pham, L. B., Rivkin, I. D., & Armor, D. A. (1998). Harnessing the imagination: Mental simulation, self-regulation, and coping. *American psychologist*, *53*(4), 429.

Thompson, E. (2005). Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the cognitive sciences*, *4*(4), 407–427.

Thompson, E. (2010). *Mind in life.* Harvard University Press.

Thura, D., & Cisek, P. (2014, 3 19). Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making. *Neuron*, *81*(6), 1401–1416. (PMID: 24656257) doi: 10.1016/j.neuron.2014.01.031

Thura, D., & Cisek, P. (2017, 8 30). The basal ganglia do not select reach targets but control the urgency of commitment. *Neuron*, *95*(5), 1160–1170.e5. doi: 10.1016/j.neuron.2017.07.039

Tien, T., Pucher, P. H., Sodergren, M. H., Sriskandarajah, K., Yang, G.-Z., & Darzi, A. (2014). Eye tracking for skills assessment and training: a systematic review. *journal of surgical research*, *191*(1), 169–178.

Todorov, E. (2009, 7). Efficient computation of optimal actions. *Proc Natl Acad Sci U S A*, *106*(28), 11478–11483. doi: 10.1073/pnas.0710743106

Tolman, E. C. (1938). The determiners of behavior at a choice point. *Psychological Review*, *45*(1), 1.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, *55*(4), 189.

Tomasello, M. (2014). *A natural history of human thinking.* Harvard University Press. ([Online; accessed 2014-12-09])

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005, 10). Understanding and sharing intentions: the origins of cultural cognition. *Behav Brain Sci*, *28*(5), 675–91; discussion 691–735. doi: 10.1017/S0140525X05000129

Tommasino, P., Maselli, A., Campolo, D., Lacquaniti, F., & d' Avella, A. (2021, 4 8). A hessian-based decomposition characterizes how performance in complex motor skills depends on individual strategy and variability. *bioRxiv*, 645317. (publisher: Cold Spring Harbor Laboratory section: New Results) doi: 10.1101/645317

Tomov, M. S., Yagati, S., Kumar, A., Yang, W., & Gershman, S. J. (2020). Discovery of hierarchical representations for efficient planning. *PLoS computational biology*, *16*(4), e1007594.

Tong, S., Li, Z., Zhu, Y., & Thakor, N. V. (2007). Describing the nonstationarity level of neurological signals based on quantifications of time–frequency representation. *IEEE Transactions on Biomedical Engineering*, *54*(10), 1780–1785.

Topalovic, U., Aghajan, Z. M., Villaroman, D., Hiller, S., Christov-Moore, L., Wishard, T. J., . . . Suthana, N. (2020, 10 28). Wireless programmable recording and stimulation of deep brain activity in freely moving humans. *Neuron*, *108*(2), 322–334.e9. (publisher: Elsevier PMID: 32946744) doi: 10.1016/j.neuron.2020.08.021

Trianni, V., & Tuci, E. (2009). Swarm cognition and artificial life. In *European conference on artificial life* (pp. 270–277).

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2008, 8). Decision making, movement planning and statistical decision theory. *Trends Cogn Sci*, *12*(8), 291–297. doi: 10.1016/j.tics.2008.04.010

Tschantz, A., Baltieri, M., Seth, A. K., & Buckley, C. L. (2020). Scaling active inference. In *2020 international joint conference on neural networks (ijcnn)* (pp. 1–8).

Török, G., Pomiechowska, B., Csibra, G., & Sebanz, N. (2019, 6). Rationality in joint action: Maximizing coefficiency in coordination. *Psychological Science*, *30*(6), 930–941. (PMID: 31088200 PMCID: PMC6560520) doi: 10.1177/0956797619842550

Ueltzhöffer, K. (2018). Deep active inference. *Biological cybernetics*, *112*(6), 547–573.

Usher, M., & McClelland, J. L. (2001). On the time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108(3)*, 550–592.

van der Wel, R. P., Becchio, C., Curioni, A., & Wolf, T. (2021). *Understanding joint action: Current theoretical and empirical approaches* (Vol. 215). Elsevier.

van Dijk, E., & Wilke, H. (1996). Tacit coordination and fairness judgments in social dilemmas. In *Frontiers in social dilemmas research* (p. 117–134). Springer.

Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2013). Are you ready to jump? predictive mechanisms in interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 48–61. doi: 10.1037/a0028066

Vidal, A., Soto-Faraco, S., & Moreno-Bote, R. (2022). Balance between breadth and depth in human many-alternative decisions. *Elife*, *11*, e76985.

Wahn, B., Kingstone, A., & König, P. (2017). Two trackers are better than one: Information about the co-actor's actions and performance scores contribute to the collective benefit in a joint visuospatial task. *Frontiers in Psychology*, *8*. Retrieved from `https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00669` ([Online; accessed 2022-12-04])

Wang, X. J. (2008, 10 23). Decision making in recurrent neuronal circuits. *Neuron*, *60*(2), 215–34. (PMCID: PMC2710297) doi: 10.1016/j.neuron.2008.09.034

Warren, W. H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 683–703. (publisher-place: US publisher: American Psychological Association) doi: 10.1037/0096-1523.10.5.683

Warren, W. H., & Fajen, B. R. (2004, 1). Behavioral dynamics of human locomotion. *Ecological Psychology*, *16*(1), 61–66. doi: 10.1207/s15326969eco1601_8

Weber, T., Racanière, S., Reichert, D. P., Buesing, L., Guez, A., Rezende, D. J., ... Wierstra, D. (2017, July). Imagination-Augmented Agents for Deep Reinforcement Learning. *arXiv:1707.06203 [cs, stat]*.

Weick, K. E. (1995). *Sensemaking in organizations* (Vol. 3). Sage.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization science*, *16*(4), 409–421.

White, P. A. (2018). Is conscious perception a series of discrete temporal frames? *Consciousness and cognition*, *60*, 98–126.

Whittington, J., Muller, T., Mark, S., Barry, C., & Behrens, T. (2018). Generalisation of structural knowledge in the hippocampal-entorhinal system. *Advances in neural information processing systems*, *31*.

Whittington, J. C., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. (2020). The tolman-eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. *Cell*, *183*(5), 1249–1263.

Wispinski, N. J., Gallivan, J. P., & Chapman, C. S. (2018). Models, movements, and minds: bridging the gap between decision making and action. *Annals of the New York Academy of Sciences*, *0*(0). Retrieved from `https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1111/nyas.13973` ([Online; accessed 2018-10-24]) doi: 10.1111/nyas.13973

Wolpert, D. M., & Landy, M. S. (2012, 12). Motor control is decision-making. *Current Opinion in Neurobiology*, *22*(6), 996–1003. (PMID: 22647641 PMCID: PMC3434279) doi: 10.1016/j.conb.2012.05.003

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature human behaviour*, *2*(12), 915–924.

Wu, S. A., Wang, R. E., Evans, J. A., Tenenbaum, J. B., Parkes, D. C., & Kleiman-Weiner, M. (2021). Too many cooks: Bayesian inference for coordinating multi-agent collaboration. *Topics in Cognitive Science*, *13*(2), 414–432. doi: 10.1111/tops.12525

Yamamoto, Y., Yokoyama, K., Okumura, M., Kijima, A., Kadota, K., & Gohara, K. (2013, 9 4). Joint action syntax in japanese martial arts. *PLoS ONE*, *8*(9), e72436. doi: 10.1371/journal.pone.0072436

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye movements and vision* (pp. 171–211). Springer.

Ydenberg, R. C., & Dill, L. M. (1986, January). The Economics of Fleeing from Predators. In J. S. Rosenblatt, C. Beer, M.-C. Busnel, & P. J. B. Slater (Eds.), *Advances in the Study of Behavior* (Vol. 16, pp. 229–249). Academic Press. Retrieved 2023-07-04, from `https://www.sciencedirect.com/science/article/pii/S0065345408601928` doi: 10.1016/S0065-3454(08)60192-8

Yoo, S. B. M., Hayden, B. Y., & Pearson, J. M. (2021, 3 1). Continuous decisions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1819), 20190664. (publisher: Royal Society) doi: 10.1098/rstb.2019.0664

Yoon, T., Geary, R. B., Ahmed, A. A., & Shadmehr, R. (2018). Control of movement vigor and decision making during foraging. *Proceedings of the National Academy of Sciences*, *115*(44), E10476–E10485.

Young, H. P. (1993). The evolution of conventions. *Econometrica*, *61*(1), 57–84. (publisher: [Wiley, Econometric Society]) doi: 10.2307/2951778

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, *71*, 165–191.

Zakharov, A., Crosby, M., & Fountas, Z. (2020, October). *Episodic Memory for Learning Subjective-Timescale Models.* arXiv. Retrieved 2023-05-04, from `http://arxiv.org/abs/2010.01430` (arXiv:2010.01430 [cs, q-bio, stat]) doi: 10.48550/arXiv.2010.01430

Zech, P., Haller, S., Lakani, S. R., Ridge, B., Ugur, E., & Piater, J. (2017, 10 1). Computational models of affordance in robotics: a taxonomy and systematic classification.

Adaptive Behavior, *25*(5), 235–271. (publisher: SAGE Publications Ltd STM) doi: 10.1177/1059712317726357

Zenon, A., Solopchuk, O., & Pezzulo, G. (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia*, *123*, 5–18.

Zhou, L., Smith, K. A., Tenenbaum, J. B., & Gerstenberg, T. (2023). Mental jenga: A counterfactual simulation model of causal judgments about physical support. *Journal of Experimental Psychology: General*.

Zhou, S., Luccioni, A., Cosne, G., Bernstein, M. S., & Bengio, Y. (2020). Establishing an evaluation metric to quantify climate change image realism. *Machine Learning: Science and Technology*, *1*(2), 025005.

Zhu, S., Lakshminarasimhan, K. J., Arfaei, N., & Angelaki, D. E. (2022). Eye movements reveal spatiotemporal dynamics of visually-informed planning in navigation. *Elife*, *11*, e73097.

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of information technology*, *30*(1), 75–89.