Action in Memory and Memory in Action

Experiments in the Encoding and Retrieval of Spatial Representation

Carla Zoe Cremer

A thesis submitted in partial fulfilment of the requirements of the University of Oxford

for the degree of Doctor of Philosophy

Michaelmas Term 2025

University of Oxford The Queen's College

Acknowledgements

I am indebted to countless hours of skilled attention by my academic supervisors and am immensely grateful to the many people who thought it was worth teaching me. I owe these years of study to Chris Summerfield: thank you for opening the doors to science for me. I will always consider it a privilege to have been given so much time to learn from you. My thinking has been fundamentally reshaped by better understanding how you think. I am grateful to Matan Mazor, who grounded my belief in the purity and possibility of science and who became everything anyone could hope for in a supervisor. Wanting to be your student was one of my very best hunches. I also want to thank Asifa Majid for the steady assurance and Jacques Pesnot Lerousseau for first making me see the elegance in cognitive science. Both the Human Information Processing Lab and the Self-Modelling Group have provided me with a rich environment, filled with subtle thinkers, always ready to provide feedback, advice and friendship. I am grateful to you all.

I want to thank the Future of Humanity Institute for funding this doctoral work and for their trust that somehow, my education will eventually translate into benefits for all. I still take this mandate seriously. I also want to extend my gratitude to the Philosophy Faculty and The Queen's College at Oxford, which showed understanding during times of personally difficult circumstances. I drew energy and delight from my access to the libraries of Oxford and specifically the Oxford Botanic Garden and never settled into taking them for granted.

Lastly, my thanks go to those who shaped my memories of writing this thesis: to Noemi for helping me understand the meaning of attention; to Ellie, for sharing these years and making them feel like an illustrated book for children; and to Aeron, for that music, the clouds, and a new self-model. To my brother Kilian, my motivation. To my brother Benjamin, my role-model. And to our parents, for the gift of a true home without which I could never have ventured and none of this would have been possible.

Table of contents

Word count: 59253

List of Figures

Abstract

In this thesis, I posit a fundamental link between action and memory in human cognition. I present results from a series of behavioural experiments which examine this interplay from two distinct angles: I investigate how action representations modulate memory formation (action in memory) and examine how memory representations shape action sequences (memory in action).

I offer a contribution to the cognitive sciences by presenting fourteen studies across five empirical chapters, which demonstrate behavioural markers of cognitive ties between action and memory. This work pertains to theories of cognition and curriculum design in service of learning.

Chapter I lays the theoretical foundations of the investigation through a review of memory research and previously revealed cognitive effects at the junction of action and memory. I situate the focus on action in memory amidst long established inquiries: at the intersection between procedural and declarative memory formation, amidst theories of active learning and epistemic actions, and in relation to the embodied cognition paradigm. I review and rely on theories of memory as reconstructive, which position memory as a substrate for mental simulation or model-based reconstruction of unseen, future or fictional scenes and behaviours.

Chapters II-IV (Action in Memory) report the foundational modulation of memory through a manipulation of the alignment between action selections and supervision signals in an associative, spatial memory task. These chapters empirically demonstrate that memory performance during test phases is superior when action selections during training were aligned with feedback signals, compared to when action selections were misaligned during training. I introduce this action alignment effect and show its generalisations and robustness against a range of controls.

In chapters V, VI (Memory in Action), I reverse and expand the inquiry and consider how memories are used to guide action sequences in a self-simulation task. Here, the experiments ask humans to reconstruct their own prior behaviour from incomplete memory in the context of a spatial rule discovery task. A re-enactment of their own learning dynamics requires an emulation of the behavioural markers associated with an epistemic transition from ignorance to knowledge. I first identify subtle behavioural markers of this epistemic transition, and then, by use of a comparison between behaviour during discovery and behaviour during a simulation of discovery, I quantify distinct biases in self-simulation of counterfactual epistemic states. Chapters V, VI thus explore the role of partial memory in metacognitive access as revealed by action sequences. In a final discussion chapter (VII), I summarise my contributions and their limitations.

Chapter I General Introduction

Introduction

This thesis is my attempt to contribute to a scientific understanding of cognition by exploring the interaction between memory formation and action selection. Chapters II-IV show how action selections affect the strength of a memory. Chapters V & VI query participants' memory for how their own actions supported their learning. Throughout, I focus on *spatial* memory and actions and orient towards making contributions that help learners understand how their own learning occurs.

This thesis supports the proposition that memory is reconstructive and fundamental to a wide range of resource-constrained cognitive processes; and that memory representations are shaped by actions. The first chapter will review literatures that pertains to these claims and will provide the theoretical background which informs the empirical work of this thesis.

In chapter II-IV, I present behavioural results from a spatial rule learning task, in which adults were asked to select actions and memorise a deterministic mapping between visual stimuli and locations on a Cartesian grid, through a supervised learning paradigm. In chapter V and VI, I present data from an unsupervised learning paradigm, in which participants learned to adjust their actions to a spatial rule and were asked to later reconstruct their action sequences based on incomplete memory. The first chapters thus focus on how actions affect memory, while later chapters are concerned with memories of actions and actions guided by memories.

In keeping with this structure, I now first review research domains which study how action affects memory and subsequently review scholarship related to how memories guide actions. The thesis intentionally spans different research areas, yet all chapters share a focus on how action selections aid (spatial) memory and how the learner's understanding of this interplay can support learning. If action can be *for* cognition, one's memory of how actions affected cognition can guide future actions to optimise learning.

This chapter proceeds as follows. I develop a foundational understanding of memory formation with a particular focus on the constructs which bear relevance to the behavioural studies we present in this thesis. I begin by introducing the multiple-systems theory of memory and on the basis of the existing literature on *associative, declarative memory* I showcase known *neurobiological and theoretical mechanisms* of human memory formation. I focus on mechanism of learning from *feedback* and learning *associations between objects and locations*, because these are features of the behavioural studies presented in chapter II-IV. This chapter then offers a description of how humans learn actions, by briefly introducing *procedural memory* and reinforcement learning, before subsequently reviewing studies relevant to the interplay between memory and action in two separate sections:

Actions in Memory

I will review three research domains that study how cognition is shaped by actions: I consider *interactions between declarative and procedural memory*, the field of *active learning* and epistemic actions, as well as the literature on *embodied cognition*. In keeping with this order, I will first review evidence suggesting that memorising actions and memorising facts might not occur through two fully separable processes. This gives way to critical questions about the classically conceived separation between movement and cognition. I consider the literature on *active learning* as a well-evidenced example of how associations, representations and models of the environments, are shaped by action selections. Here I summarise studies that show that *actions aid cognition*, leading to improved performance in cognitive tasks. Afterwards, I will consider the theoretical framework of embodied cognition, which would predict that motor actions influence abstract representations. This framework contends that intelligence is active, and our actions are mediated by and performed through the body we have.

Memory in Action

To know how actions support learning, intelligent agents must memorise which actions have led to better learning in the past. Memories of actions guide future actions. In this section, I will introduce the literature on *metamemory* and *metacognition*, to suggest that a learner who knows about how their own actions have supported memorisation, can select actions that help them learn better and faster, even during *unsupervised learning* without feedback.

The previous section will have shown that memory can be distorted by actions. Here I will expand this claim, by showing research which finds memory to be fundamentally *re-constructive* and error prone. This claim also underlines the importance of memory as the substrate through which humans *construct* plans, reason about counterfactuals and imagine future actions. This section will motivate the empirical investigations of chapter V & VI, which are premised on the idea that metacognition promises to support learning and that self-knowledge partially relies on memory. The behavioural experiments in chapter V & VI elicit the reconstructive capacity of participants' memory to guide their action sequences in an unsupervised spatial rule learning task.

All work presented in this thesis is my own, but I opted to use the editorial 'we' throughout the central chapters of this thesis to describe all efforts with acknowledgment of years of collaboration with my academic superiors and peers. We thus begin this first chapter by explaining classical and foundational constructs in the study of memory.

Foundations of Memory

Associative Memory Formation

Consider an apprentice gardener who is given two sets of seeds, which she knows to belong to the species *Lathyrus odoratus* and *Canna indica*. She looks across the botanical gardens, where beds of flowers are arranged in a grid-like structure and tries to remember the instructions she received about where each set of seeds is to be planted. Whilst this is an ordinary cognitive function, it is also a puzzling feat. How do humans reliably remember an association between a sensory input, such as the visual perception of a coloured seed, and an arbitrary label (e.g. *Canna indica*) or a spatial location (e.g. *the Northern, right-most flowerbed*)?

Psychologists and neuroscientists have long been interested in this, and similar question related to memory (Ebbinghaus, 1885) and have devised numerous experiments to understand how human memorise associations. Take for instance a classic object-in-place association test by Lhermitte & Signoret (1972), in which a set of cards are arranged in a 3 x 3 grid such that each of the nine visually distinct cards are associated with a unique Cartesian position. A human participant is presented with this arrangement and asked to memorise which visual stimulus is associated with which spatial coordinate. After each of the cards has been studied, the participant is presented with a single card at a time and are asked to tell the experimenter which location each card is associated with. Upon making their choice, participants are given feedback: the experimenter will move the card into the correct location or simply confirm it to be correct. Adults without any memory impairments tend to achieve reliable performance in this task within the first two rounds (Collins et al., 2022).

We will now describe selective and relevant cornerstones of the neurobiological and theoretical mechanisms that may underpin the behavioural measures of memory formation that concern us in the experimental part of this thesis.

The Multiple-Systems View

Memory comes in different forms (James, 1901). In 1980, Cohen and Squire reported that amnesic patients could acquire and retain memories of new routines, in this case mirror-reading, but were unable to remember which words they had read. This, and other studies (Squire, 1986), led to the suggestion that there are classes of memory formation that are independent and site-specific. Lesions in the temporal lobe do not necessarily affect the memorisation of new skills and routines and the ability to take advantage of repeated statistical regularity. Cohen and Squire (1980) suggested that memorising *that* something is the case as opposed to memorising *how* something is done, may constitute two separate types of memory and that indeed “such a distinction is honoured by the nervous system” (Cohen & Squire, 1980). Evidence supporting this view compounded (Schacter & Tulving, 1994) and a multiple-systems theory of memory replaced earlier suggestion that all memory is non-localised or made of the same stuff (Lashley, 1950).

Lesion studies have played a pivotal role in our understanding of long-term memory formation. Smith (1988) confirmed results by (Smith & Milner, 1981), in finding that a patient with bilateral lesions of the temporal lobes, was entirely unable to form new memories in object-in-place tasks. The patient was shown 16 individual objects in 16 random and distinct locations on a board and later asked to recall which object had been placed at which location. Performance in these tasks was at chance. A wealth of studies (Eichenbaum, 2004), many of which were conducted in rodents, have shown that the association between an object and a place depends critically on the hippocampus (Chua et al., 2007; Gilbert & Kesner, 2002, 2004; Sziklas et al., 1998). This is by far not the only function of the hippocampus (Bird & Burgess, 2008; Quian Quiroga, 2023), but it is the function that concerns us here. Studies in humans also show increased activation (in PET scans) of the hippocampal formation during the recall of associations between images of people and locations (Henke et al., 1997), or word pairs (Henke et al., 1999). Aggleton & Nelson (2020) studied object-in-place tasks for rodents (which place objects into 3D environments) and they too found the hippocampal formation implicated.

Decades of research on lesions, neuropharmacology and pathologies (Eichenbaum & Cohen, 2001) have found a functional and neuroanatomical separation between what is now called declarative (*that*) (Squire, 1992) and procedural (*how*) memory formation (Squire, 1986). Eichenbaum & Cohen (2004) traces the history (mainly 20th century) of competing theories about the number of types of memory (Tulving, 1985a). They recount early philosophical inquiries into habit formation and even the influence of phrenologists, who proposed that distinct brain areas host distinct functions. Eichenbaum & Cohen (2004) also lists the persuasive evidence offered by famous amnesic case studies of patients such as H.M. (Milner, 1965) or carefully controlled lesion experiments by e.g. Packard & McGaugh (1996) who showed that a rat's memory for locations versus habitual movement in space depended on functions in the hippocampus and caudate nucleus (striatum) respectively. The distinction between declarative and procedural memory is now part of an accepted cannon of understanding (Squire & Zigmond, 2003) and its discovery a scientifically significant historical milestone. We will rely on the distinction between declarative and procedural memory before considering recent results suggesting that the two systems may not be as separable as schematically suggested.

Memorising Declarative Knowledge

Declarative learning, during which inputs are associated with labels, spatial locations or responses, also implicates the cerebral cortex and the parahippocampus (Eichenbaum, 2011a).

The study of the hippocampus in particular has elucidated the main mechanistic components which are thought to underpin the encoding of such associations (Eichenbaum, 2004). One of the most influential, mechanistic accounts of learning is attributed to Hebb (Hebb, 2002; Hilgard & Marquis, 1940), who put forward the proposition is that learning occurs by an increase in the efficiency of synaptic transmission between proximal neurons. When neurons are concurrently active (pre- and postsynaptic firing), the probability that one cell will excite the other is increased by some molecular mechanism(s). Their association is thereby established or strengthened. In an often-used connectionist framework of neural information processing in the hippocampus (McClelland et al., 1995), this plasticity in synapses (the variability in the strength of association between two neurons) is modelled as a weight parameter between individual nodes (implemented as functions that represent cellular bodies) (Rosenblatt, 1958). The amplitude of the measured post-synaptic potential in biological neural networks, is analogous to the magnitude of the weight scalar in artificial neural networks.

A lasting change in the connection between pre- and postsynaptic hippocampal neurons can be molecularly implemented by N-methyl-d-aspartate (NMDA) glutamate receptors (Caya-Bissonnette & Béïque, 2024; Eichenbaum, 2011c; Squire & Zigmond, 2003). This is referred to as long-term potentiation (LTP). Pre-synaptic activity releases glutamate, which binds to NMDA receptors which, under a condition of sufficiently high depolarization of the membrane, are thus no longer blocked by magnesium ions and thereby open to the influx of calcium. An increased concentration of calcium ions can activate signalling pathways inside the postsynaptic neuron, which can lead to permanently increased sensitivity to pre-synaptic activation. This can occur for instance via an increased production of receptor proteins or total spine-volume (Bliss & Collingridge, 1993). NMDA receptors thus act as a voltage-dependent co-incidence detector and are likely implicated in learning (Tsien et al., 1996). It should be noted here that the hippocampus can implement LTP without NMDA-receptors, that plasticity does not only occur via LTP and that not all LTP is Hebbian. Memory formation occurs via a multitude of mechanisms. Reuse of neural codes is a fundamental element of adaptive cognition, which suggests we should expect a variety of mechanism of memory across different brain regions and scales (Eichenbaum & Cohen, 2001).

Learning with Supervision and Error

Since its conception, Hebbian learning has been extended (Gerstner et al., 2018) and revised to account for neurotransmitters (such as e.g. dopamine) that act as a gating signal. Such gates modulate whether coincidental neuronal activity can establish a lasting increase in efficiency of transmission. According to this model of learning, co-activity between two neurons makes them eligible to neuromorphic modulation, but when this third, modulating factor is absent, correlated neural activity does not result in a weight change. Gating is thought to act as a filter for learning the right kind of associations at the right time. While many perceptual events coincide, only some of these events are relevant to the organism or provide novel information about causal structure in the environment. Associative learning can proceed by comparing the assumed association against the true association, thereby relying on a feedback mechanism. This is also termed a *supervision signal* in machine learning (LeCun et al., 2015). The supervision signal is simply the true label the agent is trying to learn. At times, whether or not an association is indeed the correct one, will only be known after a temporal delay, which means that feedback needs to be incorporated into a synaptic event backwards in time. A mechanism that ‘tags’ particular synapses, combined with temporally proximate gating factors that are guided by attention, reward or surprise, can fulfil this role (Gerstner et al., 2018).

Surprise can be formalised as the difference between predicted and actual outcomes. Learning (or the adjustment of weights) occurs when the prediction error is above zero. This idea is at the heart of decades of research into models of learning and memory. Applications of such a 'delta' principle are found in Widrow & Hoff (1960); Bush & Mosteller (1951), and in Rescorla & Wagner (1972), as early accounts of Pavlovian conditioning to explain a predictable behavioural response that is formed by repeated association with a stimulus (Pavlov & Anrep, 1927). The idea of adapting the strength of neural connections based on (aggregate) prediction error has found use across the neural and cognitive sciences (Soto et al., 2023). While the original formalisation cannot account for many empirical findings (Miller et al., 1995) and there exist complicating critiques of this model of learning (e.g. Spicer et al., 2020), many fruitful variants have sprung from this attractively simple mechanism. One modification for instance, shapes the prediction error computation into a Bayesian framework.

Bayesian formulations assume that an agent is always engaged in considering multiple propositions (hypotheses) about the true state of the environment. The goal is to represent the environment, or more precisely, to form a belief over how the environment is structured, with sufficient accuracy to accomplish tasks. To give an example: in our object-in-place task, the hypotheses are all possible combinations of different degrees of associations between each visual input and each Cartesian coordinate. In a Bayesian model of reasoning, degrees of belief are distributed over different hypotheses and adjusted in light of the observations made (evidence). For a given observation, Bayes rule considers the likelihood of making this observation, given a particular hypothesis, by combining it with the prior belief distribution (and normalising) to yield the posterior distribution over the hypotheses given all previously acquired evidence (Chater et al., 2008). Bayesian approaches model the agent's uncertainty over the true state of the environment. Akin to a scientist who evaluates multiple theories and models and who might even favour a particular model of reality due to its evidential support, the agent and the scholar can never be completely sure, since perceptual experience, evidence and data are always finite. Learning thus occurs by updating a probability distribution based on new incoming evidence. This framework has had success in describing human behaviour and for instance accounts for the finding that humans track uncertainty and that they can report their confidence in decisions or the reliability of the information they receive (Fleming & Daw, 2017; Meyniel et al., 2015; Yeung & Summerfield, 2012).

A Kalman filter achieves a probabilistic variant of the usual point-estimates in the Rescola-Wagner model (Gershman, 2015a; Kruschke, 2008). The Kalman filter represents belief confidence as a multivariate normal distribution over all possible association strengths and models a learning rate through the variance parameter. Greater variance (uncertainty) yields larger updates to the prior probability distribution. This Bayesian formulation of associative learning can account for a greater range of empirical results (p.7, Gershman, 2015a) and has the additional advantage that it makes possible to model *active learning* (Kruschke, 2008), a concept which we will return to later.

Memorising Action Sequences

The completion of everyday and expert tasks depends on learning complex, fine-grained motor actions. Consider the head gardener who returns to the scene after lunch and begins, absent-mindedly, to execute a series of smooth and swift, well-practised actions: removing weeds, loosening and mixing the soil with fertiliser or moistening selective spots by cautiously tilting the watering can.

Her routine actions are retrieved from procedural memory, which encode the associations between visual inputs (e.g. watering can, soil, fertiliser) and habitual, fine-grained motor outputs (such as eye movements and hand- and arm-movements). The traditional multiple-systems theory of memory formation contends that action sequences are learned and memorised by an independent process and processing site (Eichenbaum, 2011b; Yang et al., 2024). We will for now consider descriptions of procedural learning as such, before turning towards more recent studies that have looked at interactions and interdependence between procedural and declarative memory.

The assumption of independence between declarative and procedural memory encoding is derived from the observation that certain lesioned patients fail to acquire new declarative memory, even if procedural memory acquisition is intact (Squire, 1986). Procedural memory involves coordinated cell activity across the pre-motor and motor cortex, the cerebellum and the striatum (caudate nucleus and putamen) (Eichenbaum, 2011b; Hardwick et al., 2013). The main model for the study of procedural memory in humans are pathologies which present symptoms of motor impairment. Parkinson's and Huntington's disease for instance affect the basal ganglia and striatum respectively. Memory acquisition for a statistically recurring structure in a sequence of actions (e.g. in a Serial Reaction Time task or a probabilistic classification task) is impaired in Parkinson's patients compared to healthy controls (Clark et al., 2014; Jackson et al., 1995, 1995; Knowlton et al., 1996; Ruitenberg et al., 2015; Vakil et al., 2021).

We consider procedural memory as it relates to our interest in object-in-place learning. Aggleton & Nelson (2020) reviews brain regions involved in rodents doing objects-in-place tasks. A learned sequence of actions in space (saccades or arm movements) is known to coincide with neuronal activation of striatal populations (p.278, Eichenbaum, 2011b). Kermadi & Joseph (1995) trained rhesus monkeys to perform six different spatial pointing sequences in response to visual cues (illuminated location) and found action-specific cell populations in the caudate nucleus, which fired anticipatorily. This suggests that the striatum encodes movements, but also an expectation over spatial movements (Graybiel, 1995). Expectation here plays a similar role as described in the previous section, in that it permits action learning via prediction error. The expected sequence of motor action is compared against the correct sequence of motor actions.

A typical formalisation of action learning in humans is reinforcement learning (RL) (Sutton & Barto, 1998) and the temporal difference (TD) prediction error (Seymour et al., 2004). These formalisations typically depend on reward-based learning. Reward-learning does not merely rely on labelling an outcome as true or correct. It incentivises the learner by providing a subjective experience of a particular action outcome and labels it correct by way of providing an inherently (to the organism) desirable quality in association with the correct action (such as e.g. the taste of sugar in Ottenheimer et al. (2020)). In humans, connections between the striatum and the basal ganglia and prefrontal cortex are thought to enable goal-directed and reward guided action (Eichenbaum, 2011b). Learning via reward is thought to occur relative to the reward prediction s (RPE), i.e. by comparing predicted reward gain against actual reward gain after a particular action (Ergo et al., 2020; Rouhani & Niv, 2021).

In RL, actions are modelled as transitions between states, each of which have an estimated value. In a basic temporal-difference reinforcement learning (TD-RL) paradigm, the value of a state (s_t) is determined by the agent's expectation of how much reward can be gained from all future states ($s_t + n$) that are accessible from the state in question (s_t). TD learning thereby addresses one of the aforementioned shortcomings of the Rescorla-Wagner model, by transcending a trial-wise learning structure and by considering temporally distant outcomes. The expected payoff received in the next state $s_t + 1$ is compared against the estimated value of the current state s_t. Once again, errors guide learning. If, for instance, this difference is large and positive, the agent will update its representation of state s_t to increase its value, since this state is likely to lead to an unexpectedly high payoff in a subsequent state (see e.g. Niv & Montague (2009) for a formal justification for TD learning). Further adaptation with decay factors as an eligibility trace exists too (Sutton & Barto, 1998).

Memory formation of an adaptive action sequence is thought to occur in response to a reward prediction error, that is signalled through dopaminergic projections into the striatum

(Montague et al., 1996; Schultz et al., 1997; Schultz, 2002). Evidence for this has accumulated on multiple levels of inquiry. Brzosko et al. (2015) report evidence in patients Shindou et al. (2019) reported in vitro evidence of dopamine acting on an eligibility trace to establish LTP (Rioult-Pedotti et al., 1998) and Pan et al. (2005) showed that a TD-RL framework best accounts for single-cell recordings of dopaminergic activity during an in vivo cue-reward experiment with rats. A recent study (Qian et al., 2025) confirms that a TD-model still appears to be the best candidate to account for dopaminergic responses (in the ventral striatum of mice) during a contingency degradation paradigm with cued stochastic reward. Here too, a Bayesian generalisation (M. Jones & Love, 2011) of the RPE in RL (Kang et al., 2024) as well as a Kalman filter for TD learning (Geist & Pietquin, 2010; Gershman, 2015a; Kruschke, 2008) has been proposed to account for the relationship between dopamine and uncertainty (Gershman & Uchida, 2019).

Action in Memory

Interactions between Declarative and Procedural Memory

We have so far considered procedural memory as a separate system from declarative memory. But there is now renewed interest in revisiting and complicating the narrative of independence between kinds of memory (M. Freedberg et al., 2020; Gade et al., 2014; Irish & Vatansever, 2020; Kalra, 2025; Sherman et al., 2024; Yang et al., 2024). Without discarding the well-evidenced distinctions between memory types, the field has turned towards studying the heterogeneity of functions within anatomical subregions, as well as the overlap and interactions between memory modules at all levels of inquiry, ranging from neuropharmacological to behavioural. This is part of an effort to incorporate neuroanatomical and behavioural evidence which calls into question a neat mapping of functions to distinct and independent brain regions. Identifying predictors of cooperation between procedural and declarative memory, might help to inform which learning conditions can optimise memory performance.

R. M. Brown & Robertson (2007) found that consolidation of a procedural learning task (a serial reaction time task) was reduced if learners participated in a declarative learning task (word lists) during the consolidation period (but unaffected by a control task in which they counted vowels). The converse was also the case: declarative learning was blocked by a subsequent procedural learning task, suggesting an interference between declarative and procedural learning and that the two systems interact rather than run in parallel. Hund & Plumert (2003) showed that spatial memory for objects in children is not encoded or retrieved in isolation but instead distorted by the identity of the objects and their semantic relations. In a review from 2011, Nadel & Hardt (2011) question the sharp and presumed region-function specificity of multiple-memory systems wholesale, by zooming in on the medial temporal lobe, to show that it is involved in cognitive function far beyond long-term declarative memory. Similarly, Hart et al. (2018) showed that the striatum plays an unexpected role in goal-directed learning (using anatomical tracing and chemogenetic manipulation) and Zeithamova & Bowman (2020) present evidence that the hippocampus also shapes generalisation and inference, thereby revealing that sites of memory formation contribute to cognitive functions beyond memorisation (Kumaran & Maguire, 2009).

In a behavioural study, Asby and Crossley showed that human participants in a category learning task are unable to deploy a hybrid of procedural and declarative learning (Ashby & Crossley, 2010). In their study, participants learned a category boundary in which optimal performance depended on both considering two stimulus dimensions (an information-integration approach), which is thought to rely on procedural learning, as well as considering a single dimension (rule-based approach), which is thought to depend on declarative learning. A model which assumed this required hybrid strategy, never provided the best fit to participants' data. Instead, participants defaulted to using a suboptimal, single strategy. The authors reason that these results suggest an interference between declarative and procedural learning, at least in respect to equal access to responses in a category learning task.

Neuroimaging and behavioural evidence for the interference between declarative and procedural memory was found by Kim (2020), who interleaved an object-location association task with a visiomotor adaptation task in close temporal proximity. In M. Freedberg et al. (2020), the authors take stock of results which show competition (e.g. Poldrack et al., 2001) as well as cooperation (see e.g. Mattfeld & Stark, 2015; Sadeh et al., 2011; Scimeca & Badre, 2012) between the declarative-hippocampal and procedural-striatal systems. They review the structural connectivity between the two regions (both for instance connect to the prefrontal cortex) and attempt to narrow the confluence of factors which appear to determine whether these two major memory systems will cooperate or interfere. M. V. Freedberg et al. (2022) reported further evidence for competitive dynamics. They used non-invasive methods (rTMS) to increase regional connectivity between the hippocampus and caudate, which led to improvements in episodic relative to procedural memory performance.

Procedural learning can shape representations of declarative nature. This provides a first example of how actions can affect cognition.

Active Learning

Active learning refers to learning contexts in which the agent selects the learning material (Nelson, 2008). By way of intervention and targeted inquiry, the learner manipulates the likelihood of what observations are made. These are thus actions which *serve* learning, memory and inference. Imagine for instance, that our apprentice gardener studies the associated nomenclature of a range of flower seeds. In her attempt to try and recall the name of the seed, she might not select them at random but instead study them in a clustered order:

first all zygomorphic flowers, followed by all actinomorphic flowers. Can her active ordering of the study material help her learn faster?

Theories of active learning are underpinned by the assumption that not all information is equally informative at all times. Some data is more conducive to learning than others. The Bayesian variation of associative learning above can demonstrate this idea. Here, evidence is incorporated in respect to a pre-specified range of hypotheses, which means that evidence which pertains to a hypothesis that is *not* considered by the agent, will not be able to alter a prior representation of said hypothesis (e.g. consider that the apprentice gardener is unaware of the existence of the adjacent rose garden and therefore cannot think it likely that the rose seeds should be planted in the annex). Similarly, in the case in which evidence does pertain to the set of known hypotheses, it may be, and in fact most often is, the case that not all hypotheses are thought to be equally likely. The prior distribution over hypotheses is not uniform. There then exist interventions which reduce the number of plausible hypotheses by virtue of logical incoherence between an observation and a theoretical assumption, which is common to a subset of hypotheses, or, in probabilistic terms (Kruschke, 2008), by virtue of the low probability of observing a particular piece of evidence under a given subset of hypotheses. Some, but not all, observation will update the probability distribution over hypotheses such that a subset becomes exceedingly improbable as an explanation for the observed environment. In the language of point estimates, the prediction error can only be made use of if an observation (mis-)matches a previously held expectation in kind.

The optimality of an action thus depends on sequence and abstraction. Firstly, the best intervention is specific to the remaining, likely hypotheses, which means that the best action at step t is not the same as the best action at t+3. Secondly, the learner must often cluster and categorise hypotheses by their dependencies, assumptions and similarities, so as to estimate how many hypotheses may be excluded for any one assumption that is rendered unlikely by a piece of evidence.

Humans do indeed learn actively, not passively. Infants already show control over their own selective attention by eye and head movements, which led Raz & Saxe (2020) to propose that active learning is a key missing component of existing theories of cognitive development. Attention in this context is framed as a covert action. Children pay preferential attention to novelty and surprise and an intermediate range of complexity: not too simple, such that there is nothing to learn, and not too difficult, such that there is nothing to learn (Kidd et al., 2012). Children as young as three years of age adapt their inquiries to the statistical structure of the task (Ruggeri et al., 2019), thereby displaying a sensitivity of action choices to the environmental factors of the learning task. Interventions are known to play a pivotal role in causal structure learning (Bramley et al., 2015) and active learning has been shown to benefit children in schooling environments (Prince, 2004; Stillesjö et al., 2021; Theobald et al., 2020; Yannier et al., 2021).

In a particularly illustrative experiment, Markant & Gureckis (2014) showed that adults benefit from selecting their own training data during a category boundary learning task. Participants were tasked to give one of two labels to a set of two-dimensional stimuli, which varied along the continuous features of angle and size. They were either shown a set of stimuli and their labels *passively* or they were able to *actively* select what feature combination they wanted to see the category label of. In this latter condition, participants thus constructed their own training dataset. The authors set out to test if learning proceeds in a hypothesis dependent manner, such that observations would be tailored to the current, hypothesised category boundary of the observer at the time. The range of hypotheses were constrained by which of two stimulus features was most predictive of the category boundary. The experimenters reasoned that if active learning was superior to passive learning, this tailoring should translate into a performance boost. And they indeed found superior test performance after the active training condition compared to learning passively from highly informative (close to the category boundary) training examples. Furthermore, the distance between self-selected stimuli and the category boundary that was most likely assumed by an individual participant at the time, was anti-correlated with performance. Participants who sampled near the category boundary were better learners. Interestingly, if a participant passively learned from the exact same order and identify of stimuli which were actively selected by another participant, performance also suffered. The benefit from self-selection thus likely derives not from the fact that the selected stimuli are per se more informative, but that they are informative in respect to a particular hypothesis that is considered at that time. A variant of this study shows that active learning studies are influenced by uncertainty in that people do not repeatedly sample unresolvable uncertainty (Patterson & Karuza, 2022).

The expectation about a hypothesis-dependent sampling bias was derived from normative accounts of how agents should prioritise informative observations when they test hypotheses (Oaksford & Chater, 2007). The framework of optimal experimental design (OED), originally developed to help select interventionist experiments by how much work they could do in buttressing scientific theories, is now often applied to describe the actions of learning agents in both machine learning and cognitive science (Coenen et al., 2019). OED provides a measure of the *epistemic utility* of an action. Formalisation capture how helpful an intervention would be, often by either how much performance in the task at hand might improve (probability gain) after new knowledge is gained, or by how much uncertainty would be reduced though the intervention (information gain) (Lindley, 1956; Nelson, 2005). Information gain is assessed by comparing the Shannon entropy before and after an intervention. This depends on a pre-posterior analysis, which considers all possible answers to each query, and sums their epistemic utility, weighted by their probability, to an expected information gain associated with each query (Meder et al., 2022).

In some instances, OED theories can capture human behaviour well, such as in logical hypothesis testing during a Wason card selection task (Oaksford & Chater, 1994) or eye-movements during search (Najemnik & Geisler, 2005). OED has been used to inform experiment designs (Rainforth et al., 2023; Valentin et al., 2024) or indeed intelligence gathering by government agencies (Timms et al., 2024). It is, however, of note that OED theories typically evaluate the optimality of a single step query, whereas ecological agents likely consider the cost and efficacy of a full sequence of queries (Meder et al., 2022). This highlights the cognitive cost and skill implicated in active learning under resource constraints. Active information sampling is a sequence of decisions between alternatives and thereby demands cognitive effort (Petitet et al., 2021). Ruggeri & Lombrozo (2015) proposed an *an ecological active learning framework* based on this presumption that no exploration strategy is cost effective and adaptive across all environments (which differ). The learner thus likely learns to choose wisely from a battery of search strategies. The study of actions as decision-making instances is fertile ground to study learning and cognitive development over time.

The study of active learning can also give insight into the mechanism of cognitive processes *in* time. Kirsh & Maglio (1994) introduced the concept of *epistemic actions*, which they define as actions that specifically facilitate or ease ongoing cognitive processes. This can include actions which achieve greater perceptual clarity or self-priming, or those that reduce computational load. By using the game of Tetris as an experimental environment (Allen et al., 2023), the authors report that participants took unprompted epistemic actions, which did not achieve game goals, but were tailored to support cognitive operations. This included rotating falling shapes, rather than mentally simulating the rotation to test a hypothesis (in this case whether or not a shape fits into a free space) or pushing shapes to a wall to estimate distances from the target location. The authors also stress the subjectivity of optimal action in Tetris: "the epistemic value of [an action], will depend crucially on the current state of the agent, as well as on how candidate placements are generated and tested" (p.543, Kirsh & Maglio (1994)). In other words, the optimal actions depend on what hypotheses come to mind.

We have highlighted papers from the active learning literature which show that humans can learn to take actions that are tailored to their cognitive processes. We will now add a review of a third and broader theoretical framework which has argued the case that thoughts depend on actions: this is the field of *embodied cognition*.

Embodied Cognition

The field of *embodied cognition* would posit that it is more reasonable to expect a tight coupling between motor learning and the learning of abstract relations, than to assume that they are separable. The proposal is that fundamentally, cognitive abilities develop through and serve to guide movement (Glenberg et al., 2013). Humans develop, learn and reason in a body. As already mentioned in the section on active learning, our movements (whether it be eye movements in infancy or the traversing of great distances to receive teachings from an admired scientist in adulthood) affect what we perceive and when we perceive it. The information we digest in the form of sensory experience shapes the representations, memories and models we build. Actions, in so far as they shape the experiences we make, thus directly shape the building blocks of our cognition. Theories of embodied cognition however take this hypothesis a step further (Butz & Kurz, 2017). They suggest that even the most abstract cognitive processes such as language and number comprehension, spatial awareness or emotions are built on and therefore tied to sensory-motor experience or memories of previous sensory-motor experiences.

Our earliest cognitive efforts are spent on learning to control and predict the body: walking, grasping, discriminating via taste, smell and eyesight. All cognitive development, even in the womb, coincides with motor behaviours (chapter 4, Butz & Kurz, 2017). This suggests that the first and most fundamental representations we acquire are purposed to control effectors. The hypothesis of embodiment is that cognition originates in, depends on and is inseparable from action and the range of actions *afforded* to the organism (Gibson, 1950). Some authors suggest that indeed the causal logic ought to be turned on its head: the role of cognition is foremostly to move the body (Glenberg et al., 2013; Olivers & Roelfsema, 2020). In that, embodiment appears to solve at least some aspects of the age-old symbol-grounding problem (Barsalou, 1999; Harnad, 1990), by conceiving of representations not as ready-made and in need of a relation to meaning and the physical world, but instead as derived and built from physical interactions between a body's range of motion and its environment (Glenberg, 1997).

This view has been fruitful to researchers in linguistics (Glenberg & Kaschak, 2002; Pelkey, 2023), philosophy (Varela et al., 1991) and artificial intelligence (Clay et al., 2021). It remains a useful framework to revisit as research in cognitive science and machine learning makes progress on a range of long-standing questions, both with and without bodies (Barrett & Stout, 2024). Proponents have been particularly interested in showing the effects of movement on language, in parts because language is a highly abstract manipulation of symbols, which does not in any obvious way relate to the body per se; and in parts because verbal communication is infused with reference to physical space and physical experiences, such as for instance the description of being 'trapped in marriage' (Glenberg, 1997) or that something is 'far-fetched' or a 'tight race'. An obvious point which supports a theory of embodiment in language is that linguistic symbols are transmitted through the body - whether it be a vocal tract or fingers. Fine-grained motor behaviour is paramount to communication, and some evolutionary theories suggest straightforwardly that all verbal communication evolved from gestures (chapter 13, Butz & Kurz, 2017). Secondly, there is evidence for anatomical linkages between regions of the brain that process movement and abstract linguistic concepts (Pulvermüller, 2005) and neuroimaging studies have reported that mental simulations of higher-level concepts elicit activity in the motor cortex. A meta-review by Watson et al. (2013) however disputes this. Modern techniques using distortions in virtual reality environments found significant correlations between motor planning and mental imagery capacity in individuals (Bennet & Reiner, 2022).

Distortions of cognition via bodily movements have indeed been found across many domains (Glenberg et al., 2013; Robinson & Thomas, 2021), including spatial perception (e.g. Proffitt, 2006), or numerical reasoning (Fu & He, 2025). Space becomes a scaffold for further reasoning tasks (e.g. Lerousseau & Summerfield, 2024). Our interest is substantially in memory, so we will consider results which pertain to that. Samuelson et al. (2011) showed that children's declarative word learning is supported by consistent spatial arrangements. Prima facie, such results are not obvious. As we have noted in the section on active learning, passivity spares the learner from costly decisions about how, when and where to take action, which is itself a computation that could distract and lead to reduced performance in the task at hand.

Glenberg (1997) relates the framework of embodiment to the purpose and function of memory, which he suggests is the learning of representations that are repeated and particular to a particular body and the particular affordances given to said body by those environments that are traversed frequently. Glenberg also discusses early results which showed that memory retrieval is facilitated when actions are spatially aligned with the remembered items and that memory encoding is often improved under enactment.

While a range of experimental results with the embodied cognition literature show that memory formation is distorted (both positively and negatively) by the body, evidence is often contradictory, and theories remain imprecise (Dijkstra & Post, 2015). Some of these results, while presented as evidence for embodied cognition, might also be explained by alternative hypotheses, such as that actions paired with memory items lead to a richer encoding due to the multi-sensory experience of an action or that congruent actions facilitate context reinstatement. For instance, Cook et al. (2008) found that gesturing supported children's retention of how to solve a math problem. However, this difference might derive from the comparison of a cognitive process against a cognitive process *and* a sensory-motor action, which means that the benefit from the latter could derive from an enhanced richness of the encoding, rather than from a presumed sensory-motor origin of understanding the concept of mathematical equivalence. Other studies however control for this by requiring an action in both conditions.

Lagacé & Guérard (2015) for instance asked participants to remember the order of manipulable items, which were either paired with a congruent or incongruent grasping action (power versus precision grip). The proportion of correctly recalled sequences was higher for items that were paired with a congruent action. In Downing-Doucet & Guérard (2014) participants committed a set of manipulable objects to short-term memory. Each object was paired with either distinct or similar motor actions. The authors found that distinct motor actions caused less interference and allowed participants to better remember distinct items. Guérard et al. (2015) extend this finding to show that motor suppression has a detrimental effect on recall, but exclusively for congruent items. The effect however only holds for a short list of items. Canits et al. (2018) also finds a congruency effect in reaction times, but not for long-term memory. In Van Dam et al. (2013), participants memorised nouns for objects that require a pressing (e.g. piano) or a twisting (e.g. pepper mill) motion or those that were not manipulatable. During a subsequent distractor task, participants then made a series of decisions with either a pressing or twisting motion, before they were given a delayed recognition task in which they discriminated previously seen nouns from new nouns. The authors found a significant effect of action congruency on recognition, meaning that nouns which described a twisting action were remembered better if the retention interval required a twisting action, than if it required a pressing action. However, Pecher, Wolters and Zeelenberg in Robinson & Thomas (2021), could not find a motor interference effect on the recall of manipulatable object (where the required finger movements were always incongruent with the objects of study), as would have been predicted by an embodied cognition framework.

The umbrella term of embodied cognition plausibly captures too many disparate phenomena (Hommel in Robinson & Thomas, 2021). Körner et al. (2015) categorises the reviewed effects into three suggested primary mechanisms: biasing, priming and sensory-motor simulation. While scholars generally no longer question *if* animal cognition arises from and is affected by bodily movement, there is a demand for specificity: *to what extent*, *how* and *when* is cognition influenced by action (Dijkstra & Post, 2015; Hommel, 2015; Ianì, 2021; Ostarek & Bottini, 2021)? One attempt at mechanism and formalisation is the theory of event coding (TEC), which proposes, in essence, that the same cognitive representation is used to perceive and trigger action execution (Hommel, 2015). In (Haazebroek et al., 2017) the instantiated connectionist model uses one and the same representation to perceive and act. Actions are understood as neural motor codes that are triggered and preceded by the prediction of a percept. The logic is that actions always lead to action-specific percepts and that percepts derive mostly from actions. TEC calls these shared representations ‘feature-codes’ which are both activated by sensory inputs and activate motor programmes. These feature codes bind multiple sensory inputs. TEC commits itself to a version of embodied cognition in which sensory-motor experience is essential for creating representations, but not for using them later (e.g. in simulating unseen feature combinations). In the model, percept-motor code associations are learned via trial and error (also referred to as motor babbling) as prescribed by an ideomotor theory (Sun et al., 2020). In this way, the model breaks with our perceived difference between action and perception (Barsalou et al., 2007; Prinz, 1990; Prinz, 1997).

Interim Summary

We have reviewed distinct memory types, learning from supervision and object-in-place learning. We considered three angles on how memory formation can be shaped through actions. This can lead to memory being strengthened or weakened, biased or distorted.

First, evidence for interactions between procedural and declarative memory systems show that learning how to do something is not independent of learning what is the case. Second, according to theories of active learning, active selections of (the order of) observations affect beliefs and memory. The idea of epistemic actions suggest that actions are sometimes tailored to a specific cognitive function, rather than a specific goal of how to change the environment. Third, the theory that cognition is fundamentally embodied, and representations arise from sensory-motor experience, would suggest that movements can shape memories.

Experiments in chapter II-IV will be situated in a supervised, object-in-place, associative learning task, in which we show that action selections can boost and diminish memory performance.

We will consider that distortions in memory, as well as an awareness of how such distortions come about, may bring an advantage to agents who learn and reason in constantly changing environments. To do this, agents need to be able to model their own cognition, at least to the extent that they know what actions support, rather than hinder, cognition. We have thus far looked at evidence showing that actions affect memory. We will now turn towards how memory affects actions.

Memory in Action

How do we learn to take epistemically advantageous actions? How is active learning learned? How do we select actions that help us make memories that last? To know what action we should take to support our own learning, we must memorise which of our actions led to improved learning outcomes in the past. We recall memories of actions to guide current and future actions. This requires that we learn about our own cognitive processes and their outcomes. The ability to apply our cognition to cognition itself, was aptly termed *metacognition*. The following section will introduce metacognition, and metamemory, and describe learning contexts without feedback, in which such insight into oneself may be of particular benefit. This will foreshadow the unsupervised task environment used in chapters V&VI and motivate our focus on behavioural metrics of metacognitive capacities in these chapters.

Metacognition in Action

In what order and for how long should the gardener study the names of seeds to do well in a botany class? Should she test herself with or without seed labels? How confident should she be that she is attending to the relevant features when classifying plants? The ability to learn about, monitor and construct models about our own cognitive processes was termed metacognition (Fleming et al., 2012).

The study of metacognition began by noting that children and monkeys learn how to learn (Flavell, 1979; Harlow, 1949). Early studies examined children's abilities to assess their readiness to recall learning materials and found that older students were better at knowing when their memory became reliable than younger students, and that such self-monitoring could generalise to new learning materials (A. L. Brown et al., 1978). In other words, children can learn to know if and when they know.

Metacognition over knowledge acquisition however is partial and faulty. One example of this partial access is for instance found in peoples' lack of understanding about the benefits of self-testing. The testing effect, also referred to as test-enhanced retrieval or retrieval practise, is well evidenced curriculum found to enhance memory. Recalling an item, rather than merely being re-exposed to an item, has repeatedly been shown to improve memory performance (Agarwal et al., 2021). In a review, Rivers (2021), describes that people tend not to be aware of this effect and tend not to use it to improve their memory performance. While people report the use of self-testing as a *monitoring* tool (to see how well they already know the material), testing is not preferred as a *memory* tool, over repeatedly studying the same material. Metacognition about the robust testing effect also stands as an example of the promise of metacognition: if leaners possessed self-knowledge over how testing boosts their memory, they would counterfactually be able to optimise their own learning.

The concept of metacognition has found application in the field of artificial intelligence (Botvinick et al., 2019) where (deep) meta-reinforcement learning algorithms (Hattori et al., 2023; Wang et al., 2017) were developed to train on instances of learning to optimise task performances. These algorithms implement learning on two different timescales, in which the slower learning process does not optimise learning to complete individual tasks, but instead learns about the similarities across tasks. A first-order system which completes individual tasks is adjusted by a second-order system (in Wang et al. (2017) this was a recurrent neural network) to enable one-shot learning on a future task.

This framing of hierarchical learning bears similarities to a model of metacognition in humans by Fleming & Daw (2017). In it, the authors distinguish between first and second-order cognitive processing. Metacognition is often measured as the degree to which a person’s self-judgement agrees with their true performance (while controlling for performance and biases) (Fleming & Lau, 2014). One approach to studying metacognition uses confidence ratings. Confidence ratings are uncertainty estimates that show a subject’s ability to judge the reliability of their own cognitive processes, such as e.g. those that underpin a binary decision or point estimate (Fleming, 2024). In Fleming & Daw (2017), the reliability of a first-order computation (such as a perception of an event that leads to an action) is evaluated by a second-order inference process that additionally takes the covariance between two variables into account: in this case the covariance between confidence and action. This means the model proposes that one’s own actions are an input to one’s own confidence judgements and that one’s actions are evaluated by accounting for the reliability of the actor (oneself) at the time at which the action was taken. Metacognition loops models over models to, for instance, account for the noise contained in an action selection process, to arrive at a measure of how reliable the output (action) of that process might be.

The importance of metacognition is highlighted by considering that many of the environments that humans navigate, do not contain labels of the correct answers and actions. The supervision signal is lacking. Scientific, philosophical or artistic progress for example is driven by discovering and approximating unlabelled truths. Recent theories of how humans accomplish this, suggest that the supervision signals are essentially self-generated and thereby reliant on memory. It may be of interest to consider whether improvements in metamemory and metacognition could be a factor that leads to improvements in rates of discoveries (Popp et al., 2024).

Metacognition in Unsupervised Learning

Learning progress without feedback is often referred to as *unsupervised learning*, which is a term from machine learning used to describe representation formation that occurs despite unlabelled data (Bengio et al., 2013; Hinton et al., 1995; Z. Lin et al., 2024). One recent formalisation of unsupervised learning in humans (Bröker et al., 2024) can illustrate the crucial role of memory in learning without feedback. We will introduce this formalisation starting with an example.

Consider a slightly amended scenario in which our apprentice gardener was given two sets of seeds, which she knows to be the species *Lathyrus odoratus* and *Canna indica*, but this time, she was not told where to plant them. Her task is to identify which seed should be planted into which of two flowerbeds that are already filled with a range of blooming species. She is quite sure that she does not know which order or clade each species belongs to, because she remembers being absent-minded in a few of her taxonomy classes. But she remembers that there are visual, structural features that are used to classify plants. She tries to recall them one by one while letting her eyes wonder back and forth from one patch of flowers to the other, when it suddenly dawns on her: all flowers in the left patch exhibit a bilateral symmetry in petal arrangements (zygomorphic), whereas species on the right appear to have an asymmetric morphology. She now knows what to do: plant the zygomorphic species *Lathyrus odoratus* in the left and *Canna indica* into the right flower patch. In knowing what she was likely to remember and in paying attention to the relevant, distinguishing feature, she relied on her metamemory to discover the right category boundary which informed her what actions to take.

Bröker et al. (2024) formalised a similar scenario. The authors propose that unsupervised learning depends on the degree to which the participant's representation of the category boundary is aligned with the true (or experimentally defined) category boundary. The unsupervised learning problem is that the learner needs to find the right representation (e.g. a category boundary) for the task. Learning without labels from the external world is here understood to in fact be a form of *self-supervised* learning. This theory is suggested as an explanation for when and why unsupervised learning improves performance - and conversely, why it sometimes does not (Bröker et al., 2022; McDonnell et al., 2012; Ramscar et al., 2013; Vandist et al., 2009). According to Bröker et al. (2024), people can gain in categorisation expertise without feedback, if they are already attending to the right features (or combination of features). The label assignment depends on already memorised representations that are more (self-supervised learning helps) or less (self-supervised learning does not help) aligned.

In the case of our apprentice gardener, the feature which she needs to attend to, to be able to assign the correct seeds to each patch, is symmetry: a spatial concept and percept which she refined through experience (Shao & Gentner, 2019) and which she memorised as being a particularly important classifier in botanical taxonomy. Part of the discovery of a solution to a problem might thus be the discovery (retrieval) of a memory. We recall Markant & Gureckis (2014) discussed in the section of active learning, in which self-directed sampling close to the perceived category boundary was beneficial to learning and thereby depended on having the right hypothesis in mind. Action sequences, optimal learning and representation are closely tied.

Bröker et al. (2022) reports experimental evidence for this theory of unsupervised learning. Categorisation performance improved under unsupervised exposure only if learners showed early signs of using the correct features as a category boundary (as indicated by correctly classifying stimuli which could only be categorised correctly if the relevant (but not the irrelevant) feature was used as the category flag).

The role of memory and right-representation is also often highlighted in the context of studies that are interested in discrete moments or problem solving or sudden discoveries. The conception of insight moments as the sudden connection between two separately memorised behaviours was one of the earliest hypotheses about how insights occur (see for instance Epstein et al. (1984) for an early study in pigeons). In the literature, sudden discoveries, are often characterised as discontinuous learning dynamics (Ash et al., 2012) in reasoning problem for which an average person's prior over the solution space is incorrect or inefficient (Maier, 1931). In a schematic three-step model of discovery, the reasoner's first attempt is followed by an impasse or by 'being stuck' with the failed solution (incorrect hypothesis or representation), before the problem representation is restructured (Chronicle et al., 2004). Descriptions of cognitive discovery tend to be vague (such as e.g. restructured representations, a shift in perspective or new connections between existing pieces of knowledge), but the idea that representational change is what underpins the difference between pre and post-discovery, resurfaces frequently (Kaplan & Simon, 1990; Kizilirmak et al., 2016; Öllinger et al., 2014; Weisberg, 2015). Despite early calls for programmatic explanations for problem solving via insight (Newell, 1972), the phenomenon is still elusive and demands for more quantifiable, high-resolution and multi-dimensional behavioural measurements of discovery persist (LeGris et al., 2024). We will answer this call in chapter V.

Unsupervised learning with the wrong hypothesis in mind might even reinforce incorrect priors. Bröker et al. (2024) suggest links to studies on biases and stereotypes (Cox et al., 2022), which show that predictions that are never corrected can strengthen even if they are incorrect (possibly via unsupervised Hebbian learning and self-reinforcement). Metacognition may include knowledge over one's biases and prejudices or how emotional states arise and affect one's judgements and decision-making. To know what cognitive factors led us to act in the way we did, may be instructive for making predictions about how we will likely act in the future if a similar situation arises. This self-knowledge in turn tells us how to bring about futures that serve the kind of person we are or serves the values we are committed to.

Our apprentice might know that her well-being always improves after a day in the botanical garden, so she is keen to pass her last botany exam and turn her back on classrooms. As she sits down to study, her sister calls with the intention to vent about a recent argument with their mother. The apprentice remembers that last time, such a call brought her to a rattled emotional state which made it difficult to memorise her flashcards. She asks her sister to call back later, thus acting on her *metamemory* (Dunlosky & Tauber, 2016) to accomplish both her goals of being a good sister and finishing her exams.

Metamemory is her awareness of factors that influence her ability to make lasting and reliable memories. Bein & Niv (2025) suggest that hierarchical reinforcement learning provides a theoretical framework for how compressed models of expected state transitions (such as for instance whether the apprentice will be able to study well after the call) might be formed in the medial prefrontal cortex. Metamemory includes knowing whether a memory is trustworthy or error prone. An accessible example of metamemory is the ‘feeling of knowing’ (Metcalfe & Shimamura, 1994) or ‘tip of the tongue’ phenomenon (Thomas et al., 2016), in which a reasoner is aware of the possession of a memory, that is however not accessible. The examples above show that metamemory (and metacognition more generally) is incomplete. Interior sensory evidence is noisy and incomplete and there may be much we do not know about ourselves, much that is not easily named or inaccessible, which might include the origins of our ideas or how we came to a believe a conclusion (Bear et al., 2020; Randall 2019; Nisbett & Wilson, 1977). Access to our cognitive processes, such as how we learn and memorise best, is partial and leads us to make choices that do not necessarily optimise learning (Bjork & Bjork, 2011). This is what led to the suggestion that metacognition, similar to (incomplete) sensory perception or learning about what others think, might be fundamentally inferential (Fleming, 2024). We model ourselves to help ourselves.

Reconstructive Memory

Our memories are not accurate reflections of what we experienced (Straube, 2012) and are subject to distortions and biases that mean that some experiences are etched into memory, while others are barely accessible (Schacter & Slotnick, 2004). We have reviewed evidence showing that movement can disrupt or boost memory. But many theories of memory go further and suggest that the nature of memory is fundamentally reconstructive, rather than an exact copy of past experiences stored (chapter 10 in Bartlett, 1932; Dudai & Carruthers, 2005). Recalling a memory, according to these theories, thereby queries a model that re-instates a previous experience from a compressed representation.

It serves us here to consider theories which understand memory as a mechanism for resource-efficient reuse of information. Resource-efficiency also demands the management of redundant information. This includes information that is contained across contexts and can therefore be re-used akin to a cached representations for future function executions (Dasgupta & Gershman, 2021). This framework suggests that memory formation too is a computation that is subject to optimisation. It may be more sensible to assume that the brain does *not* make and retrieve exact copies of experiences but instead *reconstructs* an approximation of past sensory inputs (or parts thereof). These reconstructions are then subject to a range of distortions, including false recognitions, inaccuracies and biases (Payne et al., 2009; Schacter & Slotnick, 2004; Sinclair et al., 2021). Metamemory assesses how memories are constructed, how they might be faulty, biased or reliable, is likely useful for optimising learning.

A recent paper by (Spens & Burgess, 2024) formalised this age-old idea of memory as being reconstructive by way of a generative neural network which compresses inputs and uses these learned latent variables to reconstruct (retrieve from memory) the probability distributions of observations. In their model, an autoassociative network allows for rapid encoding of sensory experiences and memories are consolidated by replaying these encodings to train a variational autoencoder, which then compresses and reconstructs observations until the reconstruction error (provided by the hippocampal autoassociative network) is sufficiently small. Interestingly, this memory consolidation model leads to predictable distortions in the representation of experience. Compression biases memory towards prototypicality, such that instances which are recalled become more similar than the original class of observations, a phenomenon that is also found in humans (Kerrén et al., 2024).

With this model, the authors not only explain how the brain might account for observational redundancy, but also why memory appears to be inextricably linked to the capacity to construct visions of the future. The relationship between a person's ability to imagine the past and the future has long been noted (Tulving, 1985b; Williams et al., 1996) but has become more integrated into theories of memory in recent years (Eichenbaum & Fortin, 2009; Hassabis et al., 2017; Schacter et al., 2012). Findings suggest that imagination (construction) and memory (reconstruction) indeed rely on the same neural substrates. Hassabis et al. (2007) and Addis et al. (2007) show that hippocampal damage leads to measurable impairments in imagination and according to Spanò et al. (2020), such bilateral lesions also reduce the frequency and richness of dreams. M. Levin (2024) suggests that it is precisely this modifiability of memory that allows for a cohesive sense of self. It is the *constructive* nature of both memory and our ability to simulate situations that have not yet occurred, which allow us to maintain cognitive stability in spite of a constantly changing environment and body. He expresses the proposal that re-invention (and to some extent confabulation) is necessary, by analogy to the auto-encoder model (also implemented by Spens & Burgess (2024)). The auto-encoder stores a compressed form of the original information, which then necessitates re-expansion and thereby re-interpretation. Under constant context shifts, the original information is thus adapted and modified to fit the needs of the organism. These frameworks and empirical results suggest that memory is ubiquitous (Eichenbaum & Cohen, 2001), and "a mode of knowledge […] concerned with the structuring and reconstitution of the past [which] cannot be dissociated from intelligence"

(Piaget et al. 1968, translated in Flavell, 1971), or in other words, a multi-facetted, foundational component of cognition.

In chapter V and VI we will empirically query peoples' models (which are partially memory-based) of their own cognition and behaviour. We showcase a new experimental paradigm which asks participants to re-instantiate their own learning dynamics in an unsupervised task with spatial structure. The experiment is tailored to elicit the generative model that participants have of themselves, by asking them to exercise the generative capacity of memory that we explained in this section. In our tasks, participants need to rely on their generative memory model of previously experienced instances of discovery, to guide their actions choices. This is memory in action.

Conclusion and Outlook

In this chapter, we reviewed a range of literatures that motivate and situate the empirical chapters of this thesis. We built a mechanistic understanding of how associative representations are formed via supervision signals and focussed particularly on declarative learning within a multiple-systems framework of human memory. We then considered how action sequences are learned via procedural learning, before reviewing studies that looked at interactions between declarative and procedural memory. This was presented as evidence that actions can affect representations, a view that we saw is shared by researchers in the field of active learning and embodied cognition. Active selections of observations and interventions in the environment can shape and construct representations and memory: actions can help us think. Metacognition involves learning what actions have helped or hindered learning in the past. The reconstructive model of memory suggests that memory is faulty and that memory underpins our ability to imagine how we will likely act in the future.

In chapter II, we study how actions affect memory by use of a paradigm through which we introduce what we termed the action alignment effect on memory. This is followed by two shorter chapters (III, IV) in which we add experimental controls and generalisations of this effect. Both chapters rely on the same experimental paradigm and methods which we introduce at length in chapter II.

In chapters V & VI we will show how reconstructive memory guides participants' actions and allows us to quantify metacognitive capacity (memory in action). Chapter V lays the methodological foundation for chapter VI in which we show an exploratory experimental pipeline that leans heavily on the task introduced in chapter V but extends it into the inquiry about metacognition.

All chapters require a form of spatial learning: Chapters II-IV study spatial learning in a supervised setting, whereas chapters V & VI study learning arising from an unsupervised discovery of a spatial pattern. All chapters are oriented towards how actions construct and are constructed from memory, how action selections aid memory and how learners might help their own learning by better understanding this interaction.

Chapter II - The Action Alignment Effect

Abstract

We present three experiments that capture the action alignment effect: a behavioural effect of action selection on memory strength. Participants learned an association between visual stimuli and locations on a Cartesian grid. They selected a single location in each training trial and received feedback at the correct location. Consequentially, the experimental manipulation constrained their active selection, thereby controlling whether their action would target the feedback location (aligned) or not (misaligned). The (mis)alignment between action and supervision signal during training significantly modulated memory performance in later recall trials.

Introduction

A logical line of argumentation posits that cognitive function can be moulded through actions, because brains are embedded in and develop in bodies. We reviewed three areas of research (see chapter I - action in memory) which study how and under what circumstances actions impact cognition. These studies are situated in a multiple-systems view of memory, where interactions between procedural and declarative memory are considered; in the active learning domain, in which observation selections affect the accessibility of inferences; and in the field of embodied cognition, where cognition is understood as arising from and ultimately purposed to guide actions. We also reviewed models that question the distinction between motor action and perception entirely.

We here report a series of experiments investigating whether the strength of associative memory can be affected by an unrelated action selection. We report results from an experimental paradigm that uses a classic object-in-place association task (Lhermitte & Signoret, 1972). Participants learned an association between a visual stimulus and one of 16 spatial locations arranged in a 4 x 4 grid, via a training signal. This training signal was fully informative, carried no rewards and marked the location which participants were later asked to recall.

Crucially, our task additionally required participants to take an action, but these actions were not pertinent to the goal of the task. Across two conditions, we solely varied whether these unrelated actions were *aligned* with the training signal (i.e. the action selects the location at which the feedback signal will be shown) or *misaligned* (i.e. the action selects another, incorrect location at the feedback signal will *not* be shown).

The persuasive line of argument with which we opened this chapter predicts that a difference in action will consequently yield a difference in memory performance. Our task was, however, designed such that there are also at least three plausible perspectives suggesting that we should expect no effect. First, consider that our task is principally a supervised learning task (LeCun et al., 2015), with a superfluous action sequence added on. The dominant field of machine learning, which supplies many fruitful models in the cognitive and neural sciences, and in which knowledge bases can and are routinely acquired in complete isolation from actions (or reinforcement learning), offers no presupposed theoretical reasons to expect our supervised learning task to be impacted by irrelevant actions. In fact, the field of artificial intelligence has made great progress on cognitive tasks (Summerfield, 2023) without providing embodiment to these cognitive systems. Since our feedback is fully informative, the supervision signal is in theory sufficient for an optimal solver to solve the task. Second, critics of the embodied cognition framework primarily think it possible that cognition, even if originating from sensory-motor experiences, can, once developed, operate fully or in parts, without any use of or interference by the body (Dijkstra & Post, 2015; Ianì, 2021; Robinson & Thomas, 2021). Researchers call for precision in identifying the exact conditions under which actions take effect on (which) cognitive processes but warn against a generalised expectation that all cognitive processes should be affected by action. Third, a historically well evidenced assumption of independence between declarative and procedural memory (as introduced in chapter I), also predicts that we will find no difference in how well stimulus-location associations are encoded. According to this classical model, actions and objects are learned separately. While we noted and summarised a recent turn towards studying some conditions under which there are interactions between the two memory types, this research direction is not about to overhaul decades worth of research showing independence between procedural and declarative memory, and far from suggesting that one should *expect* procedural memory to interfere with other types of learning. We have thus contrasted several, mostly separate fields of scholarship, which provide reasons to expect two opposing results in our study.

Our experimental design is reminiscent of studies in embodied cognition that manipulate whether an action is congruent or incongruent with the learning item. Glenberg & Kaschak (2002) for instance found sentence judgements to be impacted by whether the direction of the action taken (moving away or towards), was congruent with sentences describing movement. Van Dam et al. (2013) find that the presentation of an incongruent twisting action impairs memory for objects that require a pressing action and vice versa. Our design generalises the structure of such paradigms by pinpointing an effect to a precise and generalisable mechanism of (mis)-alignment between the selective action and the supervision signal. Furthermore, the field of embodied cognition has been critiqued for classifying too many disparate effects under one term (Hommel in Robinson & Thomas, 2021), some of which are indistinguishable from effects that stem from multi-sensory encoding or expectation violations, rather than the actions per se. We will return to this topic in the discussion to show that our study controls for and is not subject to these critiques.

Our design also bears similarities to recent publications by Schreiner et al. (2025) and Ren et al. (2025), both of which investigate how matched versus mismatched action-effect relations impact memory. Across two experiments, Schreiner et al. (2025), replicate Hon & Yeo (2021), by showing that recognition memory is enhanced for stimuli whose training presentation was preceded by a self-selected action that agreed with the perceptual event. Participants were presented with two locations (up and down) and asked to use their arrow keys to move a box on screen to their desired location. After the action was selected, the perceptual event either followed suit (the box moved to the selected location) or opposed the selected action (the box moved to the *other* location). A word stimulus then appeared in the location of the box. Half of the stimuli were thus encoded under a condition in which the perceptual event was congruent with a prior action and the other half was encoded under a condition in which the perceptual event was *incongruent* with the preceding action. The experimenters tested recognition memory for all words shown across both conditions. A similar structure of the experiment was repeated, using a different type of action: participants chose one of many colours, which would then either result in a visual item being shown in the selected colour (congruent) or a different colour (incongruent). The authors report a significant positive effect of action congruency on recognition memory across both experiments.

Similarly, Ren et al. (2025) present pilot data from a variation of this task, in which recognition memory for visual stimuli was tested in thirty participants. Here, participants were first presented with a stimulus, before they arbitrarily made an action selection to assign it to a location left or right of the presentation position. In congruent trials, the stimulus then indeed moved to that selected location, while in incongruent trials, the stimulus moved to the other location. The outcome was thus either aligned or misaligned with the participant's action. In test trials, participants later had to indicate whether the visual item they were shown had previously been presented in the training session *and* which location the presented stimulus had been assigned to. One hour later, the authors found a reaction time (RT) effect of faster responses for items learned in congruent trials during recognition tests and a positive effect of congruency on accuracy of spatial memory. Memory was altered by actions that aligned with perceptual events.

Our studies extend this line of inquiry and make several useful contributions. First, the existing literature is inconsistent and offers questions about the conditions under which effects such as the ones reported above occur and we show a new experimental design which reliably elicits the effects. Tsuji & Imaizumi (2022) for instance failed to replicate Hon & Yeo (2021) who showed that self-selected congruent action increases recognition memory. Our results and discussion section will contribute to elucidating this inconsistency and the authors interpretation of the results. Second, we study the effect under a condition in which there is no trial-wise switch between congruence and incongruence as for instance in Schreiner et al. (2025) or Ren et al. (2025) and in the context of immediate (rather than long-term) memory recall (rather than recognition) tests. Third, we show that the effect likely does not depend on self-selected actions, as previous interpretations have suggested. Fourth, we relate the effect to the supervision signal, thereby offering a generalisation of the result. Firth, we study the effect of incongruency on an action sequence, rather than a single action selection. Sixth, we isolate the effect of action congruency from the predictability of the perceptual event. Finally, we experimentally contrast the effect with other established learning curricula in the literature.

In Chapter II, we specifically test if action selection at training (with fully informative feedback) affects memory recall during immediate tests trials in an object-in-place task. If stimulus-location associations are learned in complete isolation from actions, we would expect no effect. If, however, action sequences can be shown to modulate memory, despite fully informative feedback, it suggests that human memory of relations is shaped by what action choices were executed during the encoding phase. Based on literatures that have showed interactions between action and cognition, we hypothesised that training with action misalignment (an action selects a location that is *not* the location where feedback will be shown) would lead to worse memory performance during test trials, than prior training under action alignment (an action selects a location that *is* the location where feedback will be shown). We will also test the hypothesis that the type of action taken at training is not the modulator of the effect and that the effect does not derive from learning the correct action sequence which then transfers to test trials. We will further test the hypothesis that a misaligned action interferes with memory formation, even in participants who make use of the correcting information contained in the feedback signal. We will then test the hypothesis that participants, when given the option, show a preference to align their action with the location where feedback will be shown.

We will now present the first three empirical studies on the action alignment effect. We then deepen those investigations in chapter III and extend and generalise the effect in chapter IV. All three chapters are concerned with the action alignment effect and build on each other. Our aim is to provide empirical evidence for how actions can be used to improve learning and study curricula. We situate our results amidst a multitude of psychological traditions, in the hope to be informative to researchers interested in, for instance, interactions between procedural and declarative memory, active learning, or embodied cognition. But we also stress the relevance of our results to a wide range of educational settings and a long line of psychological research assessing the role of supervision signals in learning (p.50, Druckman et al., 1994). We offer a curriculum structure that uses the alignment between actions and supervision signals to significantly improve memory for abstract relations.

Methods

Participants

Participants were recruited from the online platform Prolific (Prolific, 2024). All participants for all our studies on the action alignment effect first completed a two-back attention check experiment which lasted 5 min. In this experiment, participants were shown a sequence of large, capitalised letters on screen (one letter per trial) and were asked if the current letter matched the letter shown two trials earlier. They responded by clicking a button with the word *yes* or *no* written on it. After their response, the next trial began.

We permitted all participants to complete this attention check with the only selection criterion being English as a first and fluent language, an approval rate of above 90% and at minimum five previous submissions on Prolific. Participants who answered correctly in 75% of all trials of our attention test were invited to partake in the study.

All experiments were conducted under ethics clearance through the University of Oxford Central University Research Ethics Committee (code R50750/RE009) and all participants gave informed consent before entering any of our studies.

Participants were paid at an hourly rate of £9, plus an advertised reward of up to £4 depending on their test performance. The attention check was completed in a median time of 5.5 minutes. Each condition (with three train-test cycles, see section Design) was completed in a median time of 25 minutes. Participants were informed about the length of each experiment prior to participation, and they were able to take self-paced breaks in-between conditions (max 5 minutes).

Exclusions

Exclusions were based on self-reported use of memory tools. Participants were asked not to use their phone, pen or paper or any other tools to complete the task before they began. After the game, they were again asked if they had used any of these tools (see section Procedure) and told that their answer would not affect their bonus payments. Participants who gave confirmatory answers were excluded before any data analysis.

We additionally excluded participants who showed signs of inattentive behaviour (too slow responses or inattentive, repeated responses), by excluding participants whose average reaction times across all test trials lay 2 standard deviations above the population mean. We also computed the maximum number of times each participant repeatedly clicked on the same location during a test trial sequence and excluded participants whose highest number of repeats was 2 standard deviations above the population mean.

The same exclusion criteria were applied in all studies in chapters II-IV. This led to some sample size variability due to the variability in how many participants met our exclusion criterion in each study. We aimed to collect at least 30 participants for each condition and scaled or repeated data collection in repeated studies or more complex designs (such as Exp.2.b, c, d). For this chapter, we present Exp.1.a (n=50), Exp.1.b (n=106) and Exp.1.c (n=56). For Exp.1.c, we pre-registered (Cremer & Summerfield, 2024) our sample size based on a power analysis and pilot results.

Stimuli

We constructed four categories of semantically distinct, visual stimuli to present on screen. Figure 3 A shows a full set of 16 stimuli (category plants) that we used. In addition to a set of 16 plants and a set of animals, we also constructed a set of images of 16 household items and a set of 16 food items. Each stimulus was visually unique and coloured. All images were manually selected from Hebart et al. (2019).

Procedure

Participants were first given detailed instructions about the task and shown explanatory images for each part of the task. They were told their task was to learn the association between a unique visual stimulus (e.g. an image of a tree, see Figure 1 Top Row A) and a location on 4x4 grid of squares. Each location was associated with one unique visual stimulus. On each trial, participants would be shown one stimulus on the top of the screen and the same response grid of 16 squares (Figure 1). On each trial, they were asked to take an action to select one of the 16 squares. Trials were divided into training and test trials. Participants were instructed that they could study the associations during training trials and would be able to earn monetary bonus reward, only during test trials.

Training Trials

A training trial proceeded as follows. Initially, participants were shown a unique visual stimulus (e.g. a plant) and a response space of 16 squares (Top Row A, Figure 1). A pair of locations (a, b), diagonal and equidistant from one another, appeared highlighted in grey. One of these grey squares was framed in yellow (location (a) in Top Row A, Figure 1). Participants were instructed (both before the experiment and during the trial via a message on screen) that one of the grey squares was the correct location. They were never explicitly told which one of the grey squares (a or b) was the correct one.

One other square was coloured blue (Figure 1 Column A or B). Participants had been instructed that the blue square represented ‘their’ current location. They could move the blue square, within the response space, by use of all four arrow keys on the keyboard. Pressing the arrow key UP would cause the blue square to move 1 row higher, arrow key LEFT would move it one column to the left, etc (Top Row A, Figure 1 shows the direction options as arrows, not visible for participants and only added here for illustrative purposes). All remaining squares appeared white.

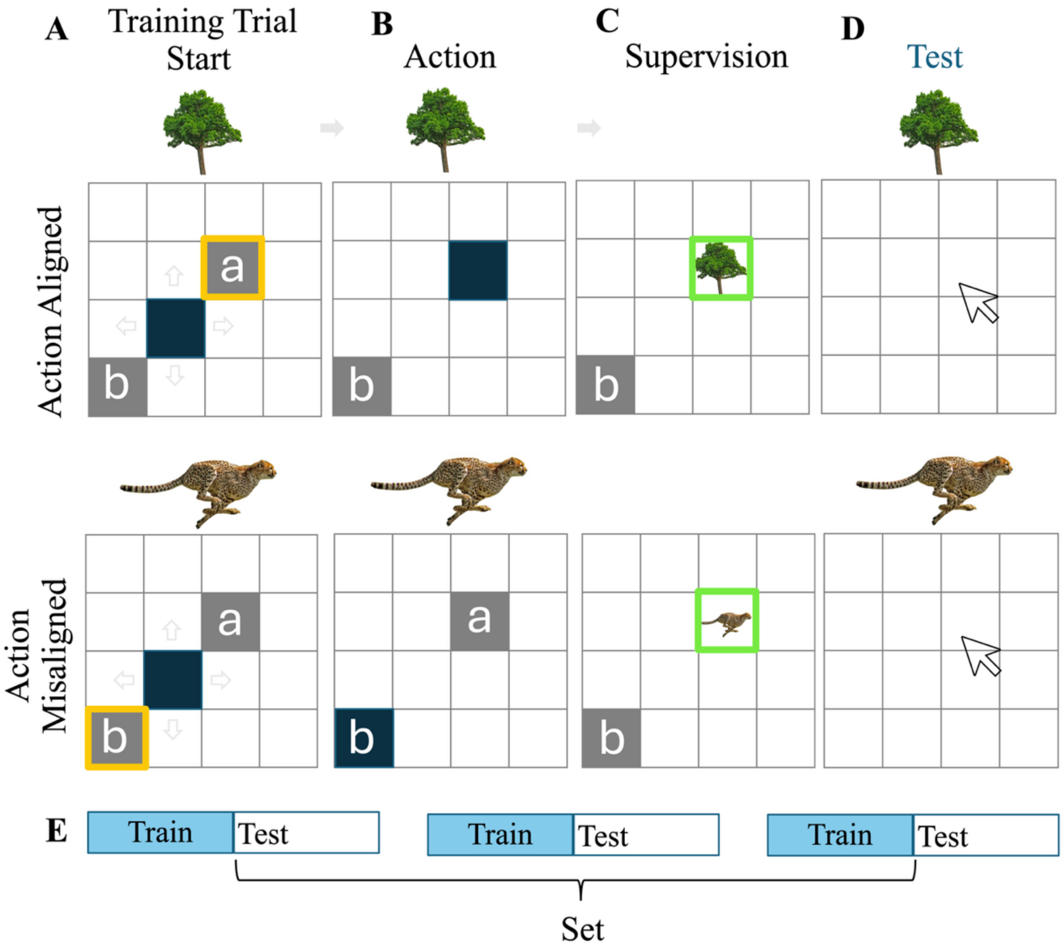

Figure 1: **Experimental Design: The Action Alignment Task** Participants learn a deterministic mapping between unique visual stimuli and unique x, y coordinates. **Top Row:** Condition Action Aligned **Bottom Row:** Condition Action Misaligned **Top A:** Start of a training trial. A visual stimulus (tree) appears above a 4x4 response grid. Two squares are highlighted in grey. One of the grey squares is framed in yellow. A blue square, which signifies the participant's location, appears randomly in one of the remaining white locations. An instruction message (not shown here) tells the participant to navigate the blue square to the target square (here **a**) with the yellow frame. **Top B**: The action sequence is completed once the participant has navigated the blue square to the target location **Top C**: The supervision signal shows which location was correct: the stimulus moves into the location it is associated with. Here, the correct location is **a**, so the action is aligned with

the supervision signal. **Top D**: A test trial in which 1 of 16 pre-trained stimuli is presented at random and participants select what they recall being the correct location via a mouse click. **Bottom A**: In the action misaligned condition, a training trial is the same as in Top A except that the target location is now **b**. Additionally, all stimuli derive from a different semantic category. **Bottom B**: Participants move the avatar to the target location **b**. **Bottom C**: The supervision signal shows the correct location, which is **a**. The action sequence in Bottom C is therefore misaligned with supervision. **Bottom D**: Test trials are the same as described in Top D. **E**: Experimental Structure. Each condition is associated with a new *set* of semantically distinct images. Each *set* is learned via 3 alternating cycles of training and test. Each block of training/test trials shows each stimulus once.

During training trials, participants were asked to navigate their avatar (the blue square) to the square framed in yellow (square *a* in Top A and square *b* Bottom A in Figure 1) by use of their arrow keys. In a series of key press actions, participants navigated the blue square into the one grey square that was highlighted with a yellow frame (Top Row B, Figure 1). They were given no time constraint during the navigation phase. The starting position of the avatar was randomised on every trial. Participants were able to navigate to any location, but the trial would not proceed, and no feedback would be shown unless they landed in the target location.

Once the blue square had been positioned within the yellow frame (Figure 1 B), the action sequence was complete. Participants were then given feedback and shown which of the two grey squares the stimulus was associated with (Figure 1 C). This feedback/supervision signal was that the stimulus moved (900 milliseconds after the blue square reached the target) into the location it was associated with (square *a* in Top and Bottom Figure 1) and by framing the correct square in green (Column C in Figure 1). Feedback remained on screen for 1.3

seconds before a new trial began and the other grey square remained on screen (see Appendix A for described variations in how we implemented feedback, through which we checked the robustness of our results to minor user-experience design feature variations). In a new trial, a new image was shown on top of the response space.

In a training trial, participants experienced one of two conditions (Top Row vs Bottom Row in Figure 1). The conditions differed by whether the square which participants were asked to navigate to (grey + yellow frame) was the correct square (square a in Figure 1) or an incorrect square (square b Figure 1). In an *action aligned* training trial (Top Row Figure 1), the square with the yellow frame (target) was always the correct square (*a*) and the other highlighted square (*b*) always a distractor (non-target). In an *action misaligned* training trial (Bottom Row Figure 1), the square with the yellow frame (target) was always the incorrect, distractor square (*b*) and the correction location (*a*) could not be acted upon (non-target).

We explicitly told participants that they would sometimes navigate to an incorrect square, but that this was irrelevant and they only had to remember the correct association. The instructions read: "The CORRECT square will always be one of the squares highlighted in GREY. Careful: sometimes the yellow frame tells you to go to the incorrect square. The task is to remember the correct square, not the incorrect square. After you navigate to the yellow square, the image will appear in its CORRECT square with a GREEN frame. That is the square you should try to remember." They were reminded of their task during the trial via message on screen. In training trials, a message appeared above the response space: "Training Trial! Use ARROW Keys to move YOUR BLUE square to the YELLOW framed square. The image will appear in the square it belongs to. Remember that this image belongs to this square.".

Participants were exposed to each training regime in blocks of 3 rounds of 16 trials each, which we here call a *set* (Figure 1 E): each condition was associated with a new set of

16 stimuli that were visually and semantically distinct. Figure 1 Top Row shows an example of learning one stimulus from the set of plant images (Figure 3 A) and Figure 1 Bottom Row shows an example of memorising a stimulus from the set of animal images. Participants first completed all rounds of training on one stimulus set before being exposed to another set and condition. In between each block of 16 training trials, participants completed 16 test trials (Figure 1 E).

Test Trials

We tested participants on how well they encoded the association in the same manner across both conditions (Figure 1 Column D). Immediately after each training block of 16 trials (in which each stimulus from the set, e.g. each plant, was studied once), participants were tested on the stimulus set that they just studied. In a test block of 16 trials, the same set of stimuli (i.e. plants) were each presented in a random order.

Participants received an indication of the switch to a test trial by being shown a message above the response space: “Test Trial! Which square does this image belong to? Click on the square which you think this image belongs to!”

At test, all squares remained white, and participants were asked to select a square via a mouse click (Figure 1 Column D). Upon a click, no other click was permitted, the text above the screen switched to ‘Response Recorded’ and the square that had been clicked turned blue. The next trial began after 1.3 seconds *without any feedback* being shown. All squares turned white and a new stimulus from the set appeared above the response space.

After one block of 16 training trials followed by 16 test trials on the same set of images, participants were given a self-paced break, in which they were shown how many bonus points (out of 80 points) they had earned during the last test round. They could proceed to the next train-test cycle by pressing the space bar.

After the full experiment was completed, participants answered two questions:

1. Did you like or dislike it when you had to navigate to the incorrect square (when the yellow frame was around the incorrect square)? Why?
2. Did you use pen/paper or any memory tools to do this task? (please answer honestly - your payment will NOT be affected!)

Task Design

We tested the effect of action alignment using a within-subject design. Participants experienced three train-test cycles in each condition. The order of conditions was randomised across participants. Each condition trained participants on a new set of images. For instance, a participant may have learned the association between plants (Figure 3 A) and grid locations by training trials that always required them to navigate to the correct location (Figure 1 Top Row). This was then followed by a condition in which they learned the association between images of *animals* and grid locations, where training always required the participant to navigate to the incorrect of the two grey locations. Conditions (and their associated stimuli sets) were never interleaved. Stimuli sets were assigned to conditions at random and the assignment between stimulus and location was randomised also.

We constructed two conditions with two different training curricula to manipulate action alignment. In our *action aligned* curriculum, participants exclusively take actions that are congruent with the supervision signal: they select the correct location on every training trial, and their action therefore targets the location in which feedback will be shown. On the contrary, our *action misaligned* curriculum restricts participants' actions to always select an incorrect location that will be misaligned with feedback on every training trial.

This design thus allowed us to test memory in after training blocks in which participants either always took the aligned action or always took the misaligned action before they were tested. The conditions were indistinguishable except for whether the action taken at training was aligned or misaligned with the feedback signal. Our supervision signal was the same and valanced neutral to positive (green frame) across conditions. We showed no ‘punishment’ signal after participants navigated to an incorrect target location in the misaligned condition. Our design ensures that a significant difference in memory retrieval at test can be attributed to action alignment during the preceding training block.

Importantly, the correct location was always (in both conditions) predictable from the start of a training trial. Once participants understood that they were in a training block which always required them to navigate to the correct location (action aligned) they could infer which of the two grey squares was correct, from the moment the trial started: it would be the grey square framed in yellow (square *a*, Figure 1 Top). If instead they were in a training context which consistently asked them to navigate to the incorrect location (action misaligned), they could infer that the correct location was the (only) other highlighted square (square *b*, Figure 1 Bottom), even before feedback was shown. The feedback signal (Figure 1 C) was the same in both conditions. The information provided and the valence of the feedback was therefore matched across conditions. Note that participants were instructed at the start that even if they were asked to navigate to an incorrect square, they would only ever be tested and monetarily rewarded for remembering the *correct* square (at test) but never be required to remember the location they navigated to.

We will report three datasets. In Exp.1.a (n=50) we compared the above conditions within participants. In Exp.1.b (n=106) we repeated this study with one additional important feature in the design: in the second and third block of each condition we introduced feedback prediction trials in which we asked participants to predict in which one of the two highlighted squares the feedback signal would appear. Exp.1.c (n=56) was a pre-registered repetition of Exp.1.b.

Feedback Prediction Trials

In both conditions, participants had access to the same information. In the action aligned condition, the correct location was always the single grey square with the yellow frame, while in the action misaligned condition it was always the single grey square without the yellow frame. Our additional feedback prediction trials thus tests whether participants understood this. In six random trials, participants were shown two buttons, right after they took an action and before feedback appeared on screen. A text above the stimulus (which remained on screen) asked whether the feedback will next appear in the grey square that they *did* (action aligned condition) or *did not* (action misaligned condition) just navigate to. They indicated their prediction about the feedback (i.e. the correct location) by clicking one of the buttons. Within each condition, the correct answer was the same across all feedback prediction trials shown in that condition. It was always either the grey square that they had navigated to (action aligned condition) or the one they had not just navigated to (action misaligned condition). These predictions allowed us to test whether participants could distinguish the actions they took from predicting the correct location. Note that this is not the same as our test trials because we do not ask them to recall which one of 16 locations is the correct location. Instead, we ask them which of 2 greyed, highlighted locations, is the correct location.

Analyses

Software

Data analyses were conducted using *Python* 3.10.13 and the following packages: *NumPy* 1.26.4, *SciPy* 1.13.1, *Pandas* 2.2.2, *Matplotlib* 3.9.2, *Seaborn* 0.13.2 and *Statsmodels* 0.14.2. We also used R version 4.4.2, and the R package *lme4*, version 1.1.35.5. Experiments were hand-coded in Javascript. This chapter was written using Quarto (Allaire et al., 2025).

Statistical Tests

We used generalised mixed effects logistic regression models (Bates et al., 2015; Harrison et al., 2018; Matuschek et al., 2017; Wilson & Collins, 2019; Yu et al., 2022) to assess statistical differences across binomial observations (i) where e.g. 1=correct square selected and 0=incorrect square selected. Generalised linear mixed effects models with maximum likelihood estimation were used to fit continuous measures (such as reaction times or detours). We conducted these tests in R (R Core Team, 2024), using the package *lme4* and functions *glmer* and *lmer* respectively. Where appropriate we confirmed that our main findings were consistent with a standard t-test.

Barr et al. (2013) recommend that a model with the maximal random effects structure (that is justified by the experimental design) should be assessed to reduce the chance of Type I errors. We thus started by modelling expected variation with random effects considering both intercepts and slopes and the nested structure of our experiment. We selected features of our experiment which could reasonably be expected to lead to variation: individual participants, block number and stimulus set exposure. We then structured the maximal model according to the nested experiment design. Since each participant experienced multiple instances of curriculum, block and stimulus set, our maximal model included a random intercept for each participant, a random intercept for each block, nested within participant, a random slope for each curriculum and for stimulus set per participant and a fixed effect for curriculum. This maximal model was compared against iteratively simpler versions of the model using a likelihood ratio test and the data from Exp. 1.a. At each random effects parameter reduction step, we tested if each factor made a significant difference (p-values were obtained by assuming the test statistic follows a chi-squared distribution) in model fit and only excluded the factor to avoid overparameterisation (Bates et al. 2018) and convergence issues if the difference between models was not significant. We found no significant difference in model fit when we (separately) excluded the slope parameters for curriculum and image set per participant (e.g. the comparison between the maximal model versus the maximal model without the image set slope yielded $\chi^2(3) = 0.33$, $p=.95$), but we found a highly significant difference ($\chi^2(1) = 609.33$, $p < .001$) between a model which only included a random intercept per participant (j) and a model which combined a random intercept per participant with a random intercept for each block, nested within participant (jk). The latter was thus our final model.

$$\text{logit}\left(\Pr(Y_{ijk} = 1)\right) = \beta_0 + \beta_1 x_{ijk} + b_{0j} + b_{0jk}$$

Significance levels indicated on all figures show the linear mixed effects model results. We maintained the same model structure to ensure comparability between our experiment versions. For the assessment of performance differences in feedback prediction trials, we used repeated-measures t-tests. Significance indicators on figures showing feedback prediction trials thus indicate those respective test results.

Measurement of Visits

We illustrate the measure of visits in Figure 2 A. We define a visit as when the participants' navigation path traversed the non-target grey square during a training trial. In Figure 2 A, the shortest navigation path from the (blue) starting location to the (yellow) target location is marked in yellow. The participant's navigation path is illustrated in blue.

For every training trial we recorded each location that the blue square was navigation through. If one of the locations in the navigation path is equal to the location of the non-target highlighted location on training trial t, this trial will be labelled as a trial in which a visit occurred (1). If the navigation path did not include the non-target highlighted location, trial t will be logged as containing no visit (0). These Bernoulli trials will be used to fit a mixed effects logistic regression to each training trial in which visits are predicted by the fixed effect of condition (see the section on mixed effects logistic regression for details). Some trials may include the other highlighted location if it lies in the most efficient path. Most navigations to the other highlighted location j will however be a detour. Any detour away from the most efficient path can also affect the time taken to complete the trial.

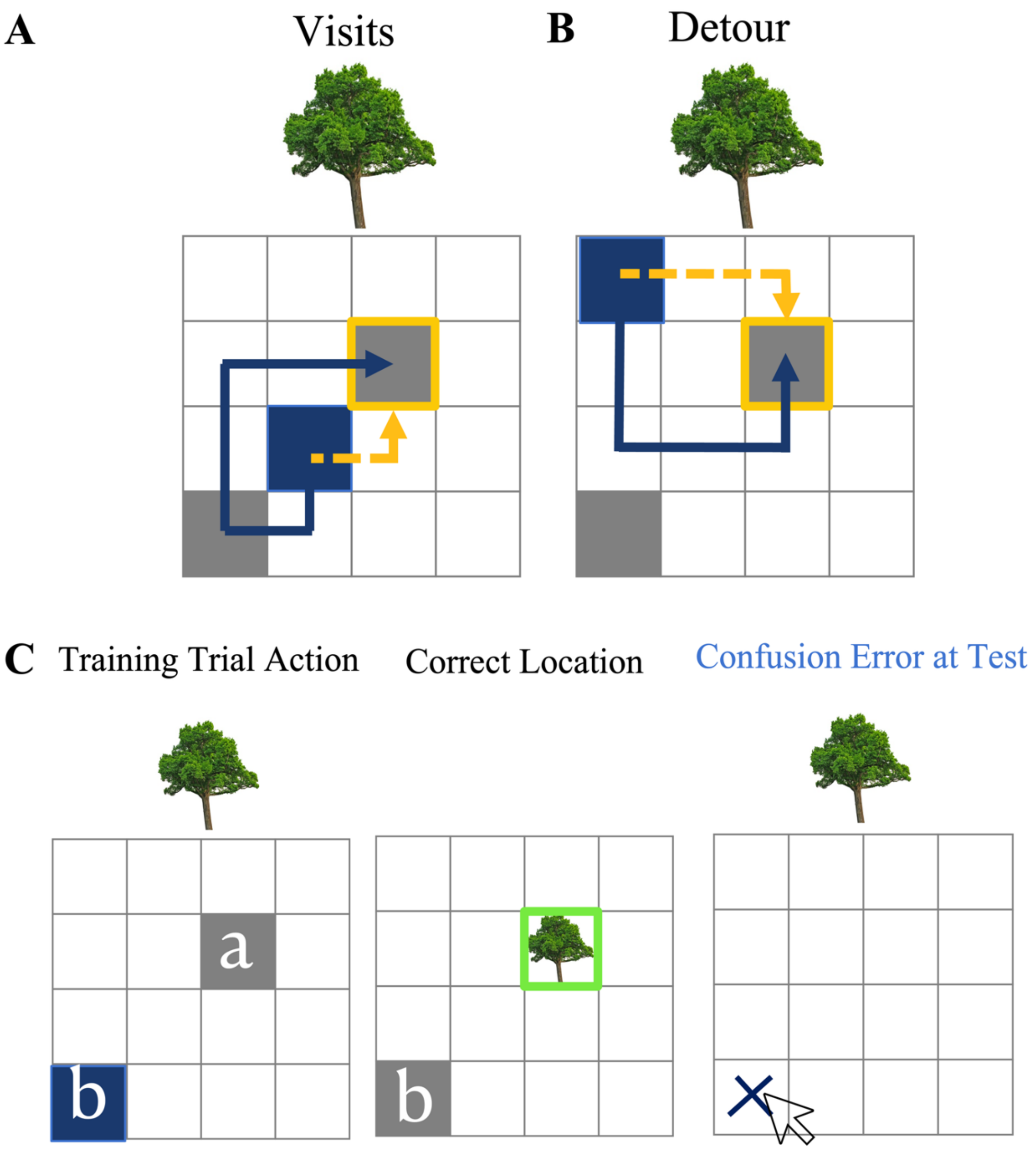

Figure 2: **Illustration of Analyses Chapter II-IV A**: Illustration of a training trial visit: the blue navigation path from the blue square to the yellow target square traverses the other, non-target grey square. The shortest possible distance to target is illustrated via a dashed yellow line. **B**: Illustration of a training trial detour. The shortest possible distance

to travel is indicated by the dashed yellow line. If the true distance travelled (indicated by the blue line) from starting point (location of blue square) to target is longer than the shortest possible path, we count it as a detour defined as the difference between the two path lengths. **C**: An illustration of confusion errors. Left panel shows action taken during training trial: the blue square was navigated into the location b. The middle panel shows that the correct location was location a, therefore the condition shown is the action misaligned condition. The right panel shows the choice made by a participant during a test trial: the cross shows where they clicked to indicate which location they believe to be associated with the test stimulus (tree). This location is the same as the one they navigated to during training. They therefore erroneously confused the correct location a (non-target at training) with the action location b (target at training).

Measurement of Detour Behaviour

We illustrate the concept of a detours in Figure 2 B. Here, the shortest navigation path from the (blue) starting location to the (yellow) target location is marked in yellow. A detour occurs when the participant's navigation path (illustrated in blue) is not the same as the shortest possible path.

We conducted detour analyses which quantified whether the misaligned condition reduced the efficiency of navigation during training. We compared participants' full navigation path lengths (measured in Manhattan distance) to the shortest possible paths between starting and target locations.

The Manhattan distance is defined as follows:

$$md(p, q) = |x_p - x_q| + |y_p - y_q|$$

In our analysis we compute the difference between the total absolute distance travelled minus the absolute shortest distance required to move the blue square (from its random starting point in one of the white squares) into the target location. Each navigation step, no matter the direction, counts as a Manhattan distance of 1 to the total count.

$$D = \mathrm{md}_{\mathrm{travelled}} - \mathrm{md}_{\mathrm{shortest}}$$

A larger detour value (D) indicates less efficient navigation, while a minimal value of 0 shows that the participant took the shortest possible route.

Detouring was predicted by the condition and the block b = {1,2,3} within a condition, which expressed how much prior training the participant had at trial t.

Measurement of Confusion Errors

We were particularly interested in a specific type of error at test. Figure 2 C illustrates what we call confusion errors. In the action misaligned condition, these are errors during a test trial that confuse the correct location with the location that was acted on. The left panel of Figure 2 C shows the action taken during a training trial: the participant moved the blue avatar to the target location *b*. The middle panel shows that the correct location was however the non-target location a. The right panel shows a test trial in which the participant needs to give their best guess of the correct location in association with the same stimulus (tree). If their choice is the incorrect location (error), which they navigated to during a previous action misaligned training trial (left panel in Figure 2 C), then we count this as a confusion error. What we refer to as confusion errors here is often referred to as swap errors in the visual attention literature (Alleman et al., 2024), and while these types of errors are related we chose to label them as confusion errors because the term seemed less committed to an association of repeated errors (for which we do not test), because we do not know how the erroneous choice relates to the encoding of the location that would have been the correct alternative and because swap errors are often studied in a context of two visual stimuli options, while our experiment technically presents more than two locations on screen.

Confusion errors are identifiable in both conditions: they are all errors at test during which the location chosen is the same location as the location that was highlighted during the prior training trials, but which was *not* the correct location. We will compare (across conditions) the number of test trials in which the error committed was a confusion error, as a proportion of all errors across all test trials.

Results

We present sample populations who partook in the experiment detailed above and present the data from each experiment in a separate figure: Exp.1.a (n=50) is shown in Figure 3, Exp.1.b (n=106) is shown in Figure 4 and Exp.1.c (n=56) is shown in Figure 5.

Accuracy

Participants learned the association between visual stimuli and grid locations. Consider panel B and C in Figure 3, which show learning curves of all participants in Exp.1.a (n=50): panel B plots test block against the average accuracy of each participant recalling the correct association after being trained in the action aligned (plotted in light blue) condition. Each level on the x-axis shows the increasing proportion of test trials (out of 16 test trials in the block) during which the participant selected the correct location associated with the stimulus shown during the trial. Average accuracy across the sample (black line) increases with each test block. Some participants reach optimal recall on later blocks and can recall the correct location for each of the 16 images. This is also the case after the same participants learned a new set of images by training in the action misaligned condition: Figure 3 C shows test accuracy increases over consecutive blocks. We plot the same learning curve for the other two experiments in Figure 4 (C, D) which plots data from Exp.1.b (n=106) and in Figure 5 (C, D) which shows data from Exp.1.c (n=56). All experiments show that participants can learn the association in both of our conditions. We thus turn to ask whether they do so better in one condition compared to the other.

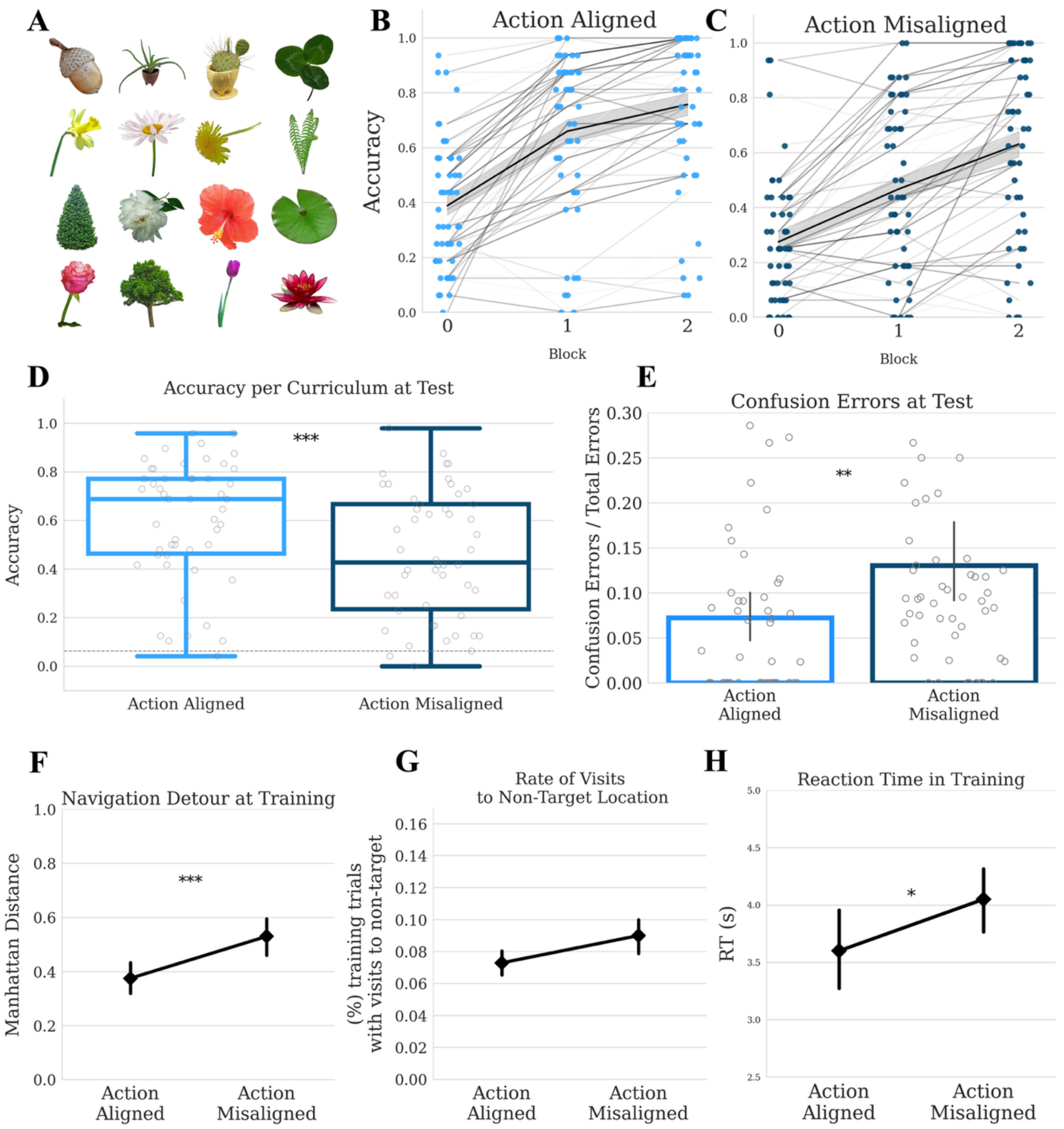

Figure 3: **Results Exp.1.a (n=50) A**: Example of one set of visually distinct stimuli within one semantic category (plants). **B**: Learning curve within the action aligned curriculum across three test blocks. Datapoints indicate each participants' average accuracy (number of correct trials across 16 test trials). Dark line shows the mean accuracy across

the population. Grey area shows the +/- standard error of the mean. **C**: Learning curve within the action misaligned curriculum. **D**: Boxplot of accuracy across all test trials per curriculum (within-participant). Each datapoint represents the average accuracy across all test trials of an individual participant: one in action aligned and one in action misaligned **E**: Ratio of confusion errors at test per curriculum. Probability of erroneously selecting the other square that was highlighted in a training trial during a test trial. **F**: Average length of navigation detour across training trials per curriculum. **G**: Proportion of training trials during which the navigation path traversed the other highlighted square, out of all training trials, per curriculum. **H**: Average time and standard error of the mean (in seconds) taken to complete a training trial, shown for each curriculum.

In Figure 3 D, we show all test trials from Exp.1.a collapsed across blocks after training via action alignment (light blue) versus action misalignment (dark blue). In Exp.1.a, participants reached an overall higher accuracy in memory recall trials that were preceded by training via action alignment. We quantify this using a mixed effects logistic regression model in which we modelled participants and blocks (nested within participant) as random effects (see Methods section for the details of the implementation of mixed effects models used in this chapter). We predict test accuracy (with accuracy coded as correct or incorrect), based on action (mis)alignment at training. In Figure 3 D we indicated statistical results in the figure using stars. The regression model found a significant effect of action alignment on the probability of a correct response: $b = 0.90$, $SE = 0.18$, $p < .001$, 95% CI [0.55, 1.26]. In Exp.1.a, action alignment during training improved test performance relative to misalignment.

We show the data from Exp.1.b and c in the same format in Figure 4 A and Figure 5 A respectively. Visually, the same effect is apparent. In Exp.1.b the same statistical analysis finds main effect of action alignment (b = 0.45, SE = 0.10, z = 4.50, p < .001, 95% CI [0.25, 0.64]) on overall test performance and the same holds true for Exp.1.c (b = 0.56, SE = 0.14, p < .001, 95% CI [0.28, 0.84]). In three experiments, we find that action alignment at training benefits memory recall at test (Figure 3 B, D).

In Exp.1.b and c we used feedback prediction trials to test if the effect still holds when we subselect for participants who could accurately predict where the feedback was going to be shown. We plot the per person average accuracy across all feedback prediction trials per condition in Figure 4 E. A repeated-measures t-test showed no differences (t(105) = -0.55, p = .59) in predicting feedback locations in Exp.1.b. Participants in the *action misaligned* condition learned that the correct location *b* was fully predictable from the moment the trial started. They learned this just as well as in the action aligned condition. This result shows that participants were not just confused about the supervision signal in our task and learned to infer which of the two highlighted locations in a training trial was correct, irrespective of the action they had to take. Feedback prediction trials show that participants knew where feedback was going to be shown, before it was shown.

In Figure 5 E we show the test accuracy of feedback prediction trials from Exp.1.c. In this experiment the action misaligned condition did lead to lower accuracy in predicting the location in which feedback was going to be shown (t(55) = -4.08, p < .001), suggesting that those participants may have been less sure about the correction location when they had to navigate to an incorrect square, even though the information about the correct location contained in the trial was the same in both conditions. We wanted to exclude the possibility that the action alignment effect stems from a detrimental effect on memory that derives from a confusion about which location ought to be remembered. Rerunning the above analysis of test accuracy on all participants from Exp.1.c whose accuracy across feedback predictions was above 90% (n=31) in *both* conditions (top performers in Figure 5 E), showed that the effect was robust and does not stem from people being unsure about the identity of the correct location: action misalignment in Exp.1.c (Figure 5) incurred a significant cost even if we run the analysis considering only (n=31) participants who had knowledge of the correct location (b = -0.50, SE = 0.18, p = 0.006, 95% CI [-0.86, -0.14]).

Errors

When actions selected the correct location, the locations were better remembered. We considered whether, if an action selection leads to an improvement in encoding the selected location, this increase in encoding might also be measurable in the memory for a distractor location after an action selected the distractor location. In action misaligned trials, participants learned the association between stimulus *s* and location *a*, but always took action which selected the distractor *b*. We turn our focus to test trials that followed the misaligned training blocks to measure if participants were more likely to have encoded an association between *s* and the (distractor) location that they acted on (*b*) than between *s* and its correct location *a*. If action selections benefit encoding, we hypothesised that actions will also lead to stronger encodings of incorrect locations, thereby skewing the associative representation between image and location towards the distractor location in the misaligned condition.

Participants who align their representation with the (incorrect) action taken, are expected to commit *confusion* errors during test trials: the location associated with stimulus *s* will mistakenly remembered to be *b*. Confusions between *a* and *b* can also occur after action aligned training, in which participants navigated to the correct location *a*, but saw a visually salient distractor *b* highlighted in grey. We measured the fraction of errors in each condition, in which participants chose distractor *b* rather than correct response *a* upon presentation of *s*.

We show the proportion of all confusion errors out of all erroneous test trials in Exp.1a in Figure 3 E. This type of confusion error was significantly more frequent after action misaligned training in comparison to action aligned training (Exp.1a: b = 0.64, SE = 0.23, z =

2.77, p = 0.006, 95% CI [0.19, 1.09]). We again applied a mixed effects logistic regression model (see Methods) to quantify this effect and mark the significance level in the figure.

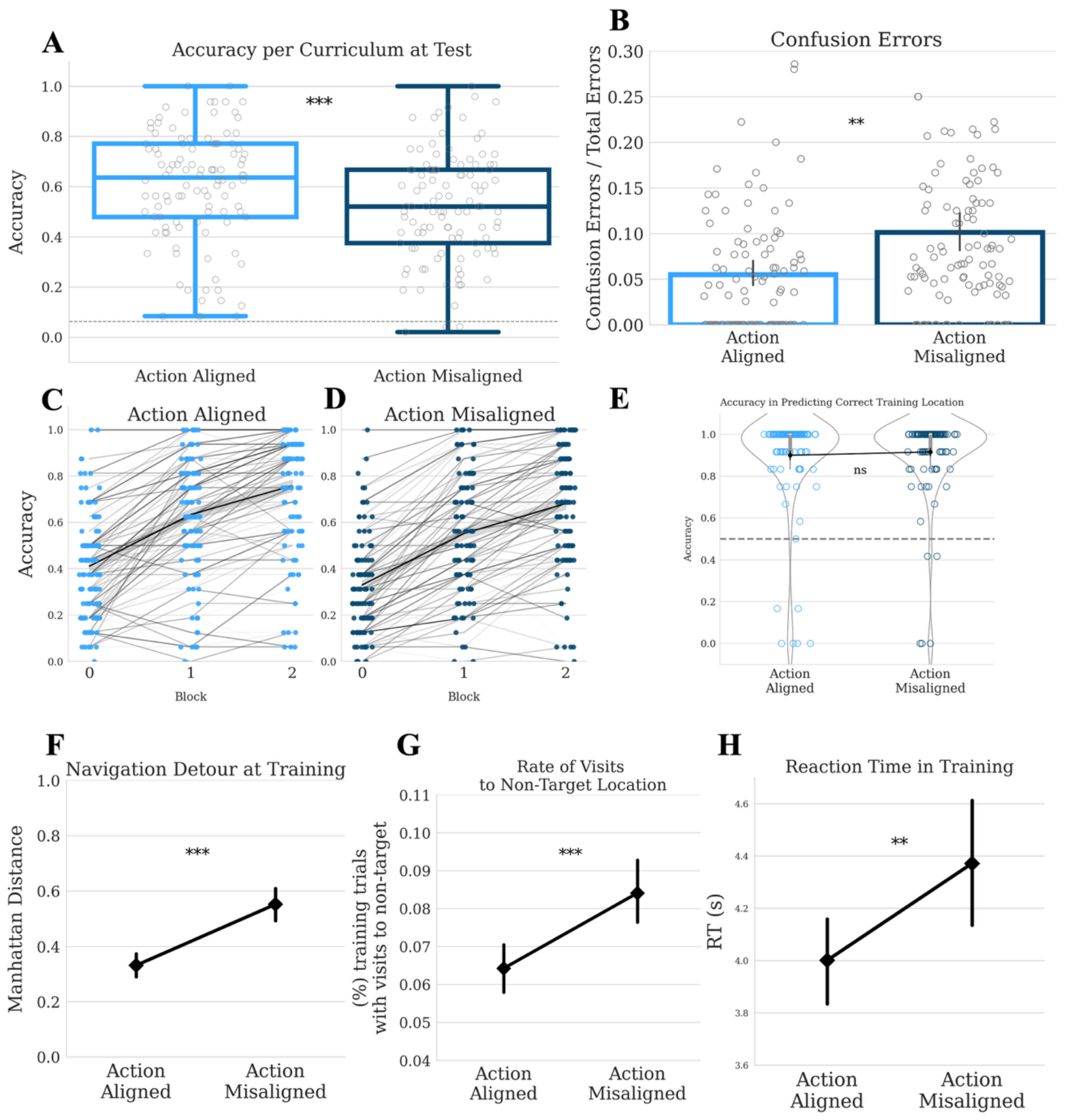

Figure 4: **Results Exp.1.b (n=106) A**: Accuracy across all test trials **B**: Rate of confusion errors **C**: Learning curve within the action aligned condition across three test blocks. **D**: Learning curve within the action misaligned condition. **E**: Accuracy across six (2AFC) prediction error trials in each condition. Pointplot in black shows that the average accuracy is no different across the curricula. Dashed line shows chance performance at

50%. **F**: Population mean of the average unnecessary Manhattan distance travelled during a training trial. **G**: Population mean of the probability of navigating through the other highlighted square during training blocks. **H**: Average time (in seconds) taken to complete a training trial, shown for each curriculum.

The same test showed the same effect in **Exp.1b** (b = 0.35, SE = 0.13, z = 2.72, p = 0.006, 95% CI [0.10, 0.60]) and **Exp.1c** (b = 0.53, SE = 0.24, z = 2.21, p = 0.027, 95% CI [0.06, 0.99]), which we plot in Figure 4 B and Figure 5 B respectively.

This was also the case in Exp.1c after we selected participants who accurately predicted the correct locations during training (**Exp.1c** n31: b = 0.69, SE = 0.29, z = 2.40, p = 0.017, 95% CI [0.13, 1.26]). Confusion errors were already more frequent in the misaligned condition after a single training block in **Exp.1a** (b = 0.70, SE = 0.32, z = 2.20, p = 0.027, 95% CI [0.08, 1.32]) and **Exp.1b** (b = 0.62, SE = 0.19, z = 3.30, p < .001, 95% CI [0.25, 0.99]) in Exp.1c n31: (b = 1.05, SE = 0.42, z = 2.49, p = 0.013, 95% CI [0.22, 1.88]) and Exp.1c n56: (b = 0.88, SE = 0.37, z = 2.35, p = 0.019, 95% CI [0.15, 1.61]).

Navigation

We next report the effect of action misalignment on navigation behaviour. During training trials, participants were able to move the avatar (blue square in Figure 1) into any of the 16 locations and were given no time constraints. This gave participants the opportunity to align their actions with the correct location (i.e. to navigate towards or to the correct location), before they had to navigate to an incorrect location to proceed in the game. They were however not incentivised to do so. We labelled training trials in which participants navigated to the (non-target) second highlighted square (square *b* in Figure 1) as *visits* (see Methods).

We also measured cognitive conflict between encoding location *a* but navigating to *b* via the degree to which efficiency of navigation was compromised. For every training trial, we computed the difference between each participant's travelled path length (measured in Manhattan distance) and the shortest possible path (see Methods). The shortest possible path was the Manhattan distance between start (randomised) and target locations. We called this measure *detouring*. We predicted more frequent *visits* and *detouring* in the misaligned condition.

Navigation paths indeed passed through the non-target location more often while participants underwent training in the action misaligned condition. The requirement to navigate to the incorrect location *b*, made participants navigate through the correct location *a* before navigating to *b*, in more trials than under action aligned training. According to a mixed effects logistic regression, this trend was significant in **Exp.1.c** (Figure 5 G) in both the full sample (Exp.1.c n56: b = 0.38, SE = 0.12, z = 3.31, p < .001, 95% CI [0.16, 0.61]) and the subselection according to feedback prediction trials (Exp.1.c n31: b = 0.60, SE = 0.16, z = 3.63, p < .001, 95% CI [0.27, 0.92]). This same trend was also significant in **Exp.1.b** (Figure 4 G; Exp.1.b: b = 0.30, SE = 0.09, z = 3.56, p < .001, 95% CI [0.14, 0.47]) and was marginally significant in **Exp.1.a** (Figure 3 G; b = 0.24, SE = 0.12, z = 1.93, p = 0.053, 95% CI [-0.00, 0.47]).

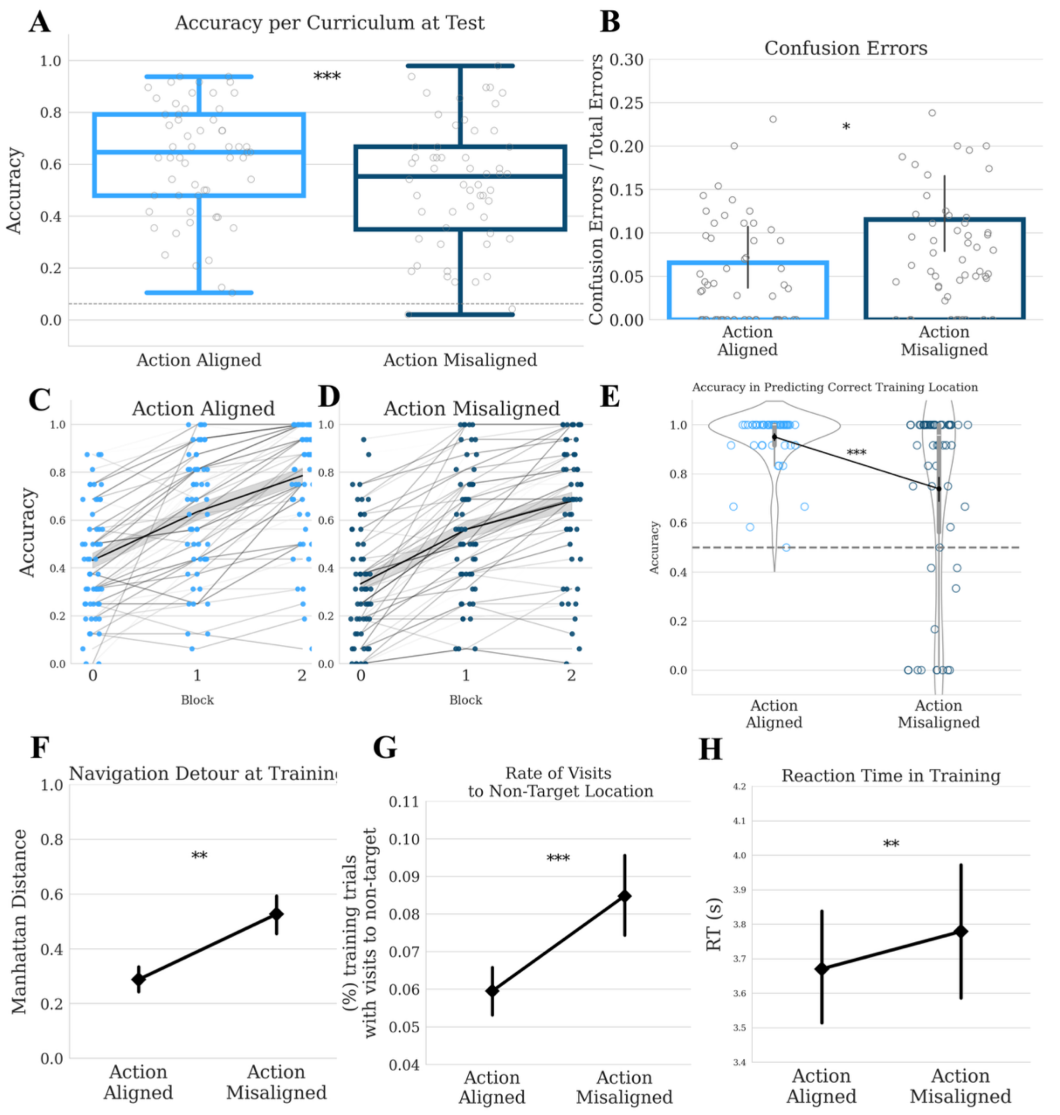

Figure 5: **Results Exp.1.c (n=56)** **A**: Average accuracy across all test trials for each curriculum **B**: Rate of confusion errors **C**: Learning curve within the action aligned condition across three test blocks. **D**: Same learning curve within the action misaligned condition. **E**: Accuracy across six (2AFC) prediction error trials in each condition. Pointplot in black shows that the average accuracy differs. Dashed line shows chance performance at 50%. **F**: Population mean of the average unnecessary Manhattan distance travelled during a training trial. **G**: Population mean of the probability of navigating

through the other highlighted square during training blocks. **H**: Average time (in seconds) taken to complete a training trial.

We posited that cognitive conflict would be reflected in more inefficient (unnecessarily long) navigation paths and that inefficiency would correlate with increasing certainty over the identity of the correct location. This led us to expect an increase in detouring for later blocks, since later blocks are preceded by a high number (e.g. 32) of exposures to training trials in which the correct location is consistently the non-target option.

We computed the unnecessary Manhattan distance travelled in each training trial. The average shortest path did not differ between conditions, but according to a linear mixed effects model with an interaction term between time passed (block) and the efficiency of the path taken (see Methods), detours significantly (shown in panel F for each experiment) and positively interacted with block in the action misaligned condition across all experiments (**Exp.1.a:** b = 0.12, SE = 0.04, t(4750) = 3.46, p < .001, 95% CI [0.05, 0.20]; **Exp.1.b:** b = 0.10, SE = 0.02, t(10070) = 4.29, p < .001, 95% CI [0.05, 0.14]; **Exp.1.c** (n=56): b = 0.10, SE = 0.03, t(5320) = 3.10, p = 0.002, 95% CI [0.04, 0.16]; Exp.1.c (n=31): b = 0.11, SE = 0.04, t(2945) = 2.82, p = 0.005, 95% CI [0.04, 0.19]). Across all our experiments, action misalignment at training lead participants to make more detours the more training trials they were exposed to.

Reaction Times

Less efficient behaviour and uncertainty can increase the time taken to complete a trial. If detouring inflates trial length, we expect the time between trial start and feedback (RT) to increase in misaligned training trials. We found that this measure agrees with our finding that participants navigated less efficiently during misaligned training. Reaction times are plotted in panel H in each figure.

According to linear mixed effects model which predicted trial length (seconds) based on condition, participants were significantly slower to complete a trial in the action misaligned condition in **Exp.1.a** (b = 0.45, SE = 0.15, t(250.00) = 3.07, p = 0.002, 95% CI [0.16, 0.74]) and **Exp.1.b** (b = 0.37, SE = 0.15, t(530.00) = 2.50, p = 0.013, 95% CI [0.08, 0.66]). In **Exp2.c** this difference was not significant (n31: b = 0.09, SE = 0.17, t(155.00) = 0.56, p = 0.580, 95% CI [-0.23, 0.42]; n56: b = 0.11, SE = 0.13, t(280.00) = 0.84, p = 0.402, 95% CI [-0.14, 0.36]).

Discussion

Summary

We have presented a supervised learning task with superfluous actions, in which participants learn to memorise associations between visual input and spatial locations. When training required participants' actions to select the response location in which the fully informative supervision signal would appear, participants had better test performance than if the action selected an incorrect location. We observed this effect in all test blocks and across all test trials. We term this finding the *action alignment* effect.

This result is not due to participants being confused about the location of the correct location: additional test trials confirmed that this effect was present even when participants recognised that the information available to them was the same across conditions. Participants were able to learn the associations in both conditions, however, participants in Exp.1.c (but not Exp.1.b) did find it harder to predict the correct location during action misaligned training. This too may be a reflection of the action alignment effect, but the effect also remains when we exclude participants who did not reach optimal performance in feedback prediction trials.

We also found that participants were tempted to align their actions with what they knew to be the correct location, as evidenced by unnecessary navigation detours during the action misaligned curriculum. Interestingly, detouring towards the correct location increased with increasing blocks, showing that more training, i.e. greater certainty over the correct locations, tempted participants more to align their actions with the correct location in the misaligned condition.[1]

In action aligned training, the rate of false recall of the distractor was significantly lower than in the action misaligned condition. In the misaligned condition, the incorrect location was encoded because it was acted upon.

The action alignment effect is not due to a correct motor sequence being learned in the action aligned curriculum that subsequently provides an advantage during test trials. We controlled for this possibility by changing the type of action taken between training and test, such that no procedural knowledge of the training sequence could carry over to test. Participants navigated during training and clicked during test.

---

[1] Li et al. (2023) also find this self-directed alignment between action and an incorrectly encoded distractor location. In their task, participants are presented with $n = \{1,2,3\}$ coloured shapes at once and after this presentation are tasked to recreate one of these shapes. To do this, they moved a mouse cursor across a circle in which the distance from the center determines the colour and the angle determines the shape. Li et al. (2023) found that participants moved the mouse trajectory towards a distractor color-shape combination, akin to our measure of ‘visits’ to the non-target location.

Notably, we found the effect in a task in which participants have complete access to which location is correct from the start of each trial across both conditions. Participants were instructed multiple times (see methods section) that they would *only* be rewarded for their ability to recall the correct associations between image and location, *not* their navigation paths. Their navigation paths and trial lengths were significantly longer than necessary in action misaligned curricula, even though they were only able to earn monetary reward in test trials and were thereby incentivised to complete training trials swiftly.

In chapter III and IV, we will add controls and generalisations of this effect. We will turn discuss a subset of the plausible interpretations and explanations of the action alignment effect.

Interpretation

We suggest that our effect derives from a (mis)-alignment of the option which the participant actively selected, versus the option that was indicated as correct by the supervision signal. When actions are aligned with feedback, (declarative and procedural) memory processes interact to support learning, while when actions are misaligned with feedback, they produce interference. This is our proposed interpretation. But there are a few alternative hypotheses we should briefly consider. We begin to do this here and continue this exploration in chapter III and IV.

If we consider the action in the aligned condition to be a ‘success’ and the action in the misaligned condition to be a ‘failure’, similarities between our studies and the affective literature become apparent. Could it be that participants learn more in the action aligned condition, because they interpret their action as successful as opposed to a failure in the action misaligned condition?

Humans have a positivity bias. For instance, in learning about partially observable variables by trial and error (such as payout probabilities in a bandit task), updates from rewards tend to be greater than those from punishments (Cazé & van der Meer, 2013; Gershman, 2015b; Katahira, 2018). Positive outcomes or descriptions are attended to and memorised and used by preference, a phenomenon sometimes called the Pollyanna Principle (Lewicka et al., 1992). Telling people about successes, rather than failures, has been shown to promote learning, even during binary choices where these feedback signals offer identical information (Baumeister et al., 2001; Eskreis-Winkler & Fishbach, 2022). One reason as to why people sometimes learn more from successes, is that the information contained in successes can be higher than in failures (Eskreis-Winkler & Fishbach, 2022), but this explanation does not apply here. Our participants had complete information about the correct location in both conditions. The asymmetry is also thought to (at least partially) derive from a discouragement of goal-pursuit. This too is difficult to accept as an explanation for our effect, because participants' explicit 'goal' is to navigate to an (at times incorrect) target. By virtue of how we set up how they can precede to the next trial, they thus always accomplish their immediate goal. In another view, participants are not given a choice about goals in our task and thus have no reason to see their navigation path as a reflection of their ability to achieve their own goal. We also 'warn' participants that they will sometimes need to navigate to an incorrect location without any effect on their reward at test. Participants additionally know that training trial performance does not count toward their bonus in our task. Finally, we measure their true ability to predict the correct location via feedback prediction trials (thereby giving participants an opportunity to show their true ability). Our results hold even for participants who are fully competent at learning which of two squares is correct. The effect does not derive from participants being discouraged or less able or interested in learning the correct location during training trials.

Guitart-Masip et al. (2014) discuss diverse evidence for an interaction between action and valence: withholding a response to gain reward and taking an action to avoid a punishment are known to be more difficult than taking an action to gain reward and withholding an action to avoid punishment. However again, in our task participants take the same action in both conditions. This explanation too does not seem to be a satisfying account.

The above variants of affective hypotheses for what might explain the action alignment effect, presume that our feedback signal is experienced as more rewarding in the action aligned condition, even though we ensured that the feedback signal is the same. We crucially do not implement any punishment or negative visual feedback when participants navigate to the incorrect, target location. These hypotheses thus depend on the assumption of a covertly generative reward by action alignment. It is equally plausible to suggest that participants experienced an internally generated reward signal for hitting their target square (including the incorrect square), regardless of the condition, since this was the task they were assigned. The proposal that the effect is due to a covertly generated reward signal, might face the retort that it is neither behaviourally verifiable, nor the simplest explanation.

Alternatively, might our effect originate in participants' expectation about action outcomes? In directional agreement with the embodied cognition literature, researchers find that action affects cognition via expectation. Humans spent much of their early development learning to predict the outcomes of their bodily actions. These expectations accumulate over a lifetime and thus shape our expected percepts (Press et al., 2023b), which can result in either an attenuation (Lubinus et al., 2022) or sharpening of sensory input (Press et al., 2023a; Yon et al., 2021). It is possible, that a participant who knows *a* to be the correct location, experiences a violation of an expected co-occurrence when they find themselves forced to consistently choose *b*. It is plausible that such a violation accounts for our effect. Maybe 'expectation alignment' leads to better encoding and 'expectation violation' leads to worse encoding in our task, but further research would need to investigate this hypothesis. It is of note, that this account presumes that humans' prior expectation might be that their actions tend to be successful, i.e. that they probabilistically receive positive feedback for the action selections they tend to make. This is a tenuous assumption, since in most contexts, there are more ways to fail than to succeed, and feedback is not always available.

In the introduction we summarised work by Schreiner et al. (2025), who found action congruency (or alignment) yielded better recognition memory for items (and Ren et al. (2025) found better spatial memory) than action incongruency. These effects bear similarities to our own but were interpreted by the authors as related to and originating from a sense of agency (SoA). Agency (or the perception thereof) is thought to arise when an agent's predictions about the outcomes of goal-pursuit or actions, matches the observed outcomes (Haggard, 2017). Schreiner et al. (2025) indeed found that a self-reported SoA was higher in their congruent, compared to their incongruent condition, leading to the suggestion that agency supports memory formation (see also Hon & Yeo (2024), who also reported supportive results). Our work also helps to elucidate some of the inconsistencies in the literature.

Unlike Hon and Yeo (2024), Tsuji & Imaizumi (2022) had found no effect of congruency in (know and remember) recognition tests for word stimuli. They did however find their reported measurement of participants' SoA to increase for congruent trials. Tsuji & Imaizumi explanation for the difference between their own results and that of Hon and Yeo (2024) focused on the difference in learning strategy for the two studies: Hon and Yeo (2024) did not tell participants that they had to remember the stimuli, while Tsuji & Imaizumi (2022) did. Tsuji & Imaizumi (2022) suggest that the intentionality in learning increased the encoding of stimuli that were seen during incongruent trials, such that the difference between the two trial types was reduced. It is possible that this is the case for their experimental setup and the number (63) of stimuli (two-character Japanese kanji compounds) they used to test their participants (n=31). Our studies were however able to confirm that even when participants are told that they need to memorise the stimuli associations, the congruency/incongruency effect on memory (recall) is present. While a direct comparison between our studies and Tsuji & Imaizumi (2022) is not possible to due to variation in too many design features such as the stimuli, length of the study, testing strategy, user experience and latencies, our results suggest that learning intention might not be the cause of the null result. The authors' alternative explanation was that the SoA was not sufficiently strong, but our study finds a congruency/alignment effect without a variation or implication of SoA, suggesting that the effect can be present without SoA being implicated.

A major difference between these studies and ours is that our participants lack agency during training. While Schreiner et al. (2025) and Ren et al. (2025) allow participants to determine which action they want to take, we consistently enforce an action (or at least its endpoint). Misalignment in the aforementioned studies is achieved by an experimental manipulation of the perceptual event. This retains the participant's agency. We, however, achieve misalignment by requiring an incongruent action at the expense of agency. It is thus unclear, whether the concept of agency is relevant to explain our results. The general suggestion, that agentically experienced behaviours are better remembered, seems sensible: agents need to have precise representations about how that which is under their control, e.g. how their choices or motor actions yield changes in the environment. As we discussed in the introductory chapter, this ability is paramount to the ability to identify causal relationships or exclude hypotheses, which in turn is useful for staying alive. It is not impossible that our effect might stem from what might be called *counterfactual agency*: participants in the misaligned condition predict the correct location and are prevented from benefiting from the memory boost of self-selecting the correct action. While this an interesting proposition, it is not a straight-forward argument to make and fails to explain why or how agentic selection would yield improvements in object-in-place learning above and beyond the theoretically sufficient feedback signal. We will return to this suggested interpretation again after we report results from chapter IV, where we report some evidence contrary to this hypothesis. In chapter III and chapter IV, we consider and exclude further explanations for the action alignment effect. The following two chapters will stress-test and expand what we have presented thus far.

Limitations

Our behavioural investigations do not permit us to say much about neural mechanism. It is one of the reasons why we must keep in mind the varieties of scholarship in psychology which might relate to or explain the result.

Some questions however can be addressed behaviourally. These questions and limitations motivate the empirical work in the following two chapters. We have shown a benefit of action alignment over misalignment, but we have not shown if alignment confers a benefit over a baseline study condition or whether misalignment is detrimental relative to baseline. What is the baseline in our task? What is the role of confusion errors and the role of attention (as opposed to action) in our task and what is their relationship? We will investigate these and further questions in chapter III to further pinpoint the action alignment effect.

Chapter III Controls for the Action Alignment Effect

Abstract

In this chapter we report from studies in which we experimentally controlled for alternative explanations and sources of the action alignment effect on memory. Two experiments, in which we introduce two baseline conditions where participants do not take actions inside the response space, indicate that the action alignment is beneficial relative to merely observing the study material. Observing a distractor square, relative to acting upon it, reduces the effect. These experiments also provide evidence that confusion errors do not account for the detriment of action misalignment. Two further experiments, in which we then modulated attention in a comparison of action alignment and misalignment showed that attention modulation alone does not modulate the action alignment effect. We suggest that selective action, not merely selective attention, plays a role in generating the effect.

Introduction

Chapter II gave rise to a set of follow-up questions: Does action alignment confer a benefit relative to a baseline supervision task? Can reduced accuracy in the action misaligned condition be fully attributed to confusions errors, in which the target location, rather than the correct location is recalled? Is it warranted to label this finding as the action alignment effect, or is attention all you need? We report four experiments which address these questions. We will then expand on the discussion on the action alignment effect from chapter II, by expanding our considerations to the informative results from chapter III.

In Exp.2.a and Exp.2.b we generate two baseline conditions in which we compare participants' performance after training with action (mis)-alignment to performance after they simply observe the supervision task and take an irrelevant action. This provides insight into the benefit of an aligned action relative to pure study and allows us to assess the importance of (mis-)aligned *actions*, as opposed to feature-based *attention*/distraction. We are also able to discern under what conditions confusion error arise and if they fully explain the performance difference between alignment and misalignment. Here we specifically test the hypothesis that the action, rather than observation alone, yields the effect we found in chapter II. We expect the difference between alignment and misalignment on memory recall to be greater in conditions in which participants take an action in the response space, compared to the difference between alignment and misalignment when participants do not take an action in the response space.

In Exp.2.c and d we turn our attention to how much attention alone might explain our results. We cannot fully disconnect spatial movement from the attention that is required to complete the motion. Attention and action are inextricably linked (Olivers & Roelfsema, 2020) and attention deepens the encoding of a spatial location (Muzzio et al., 2009). Can this explain our effect? While we cannot dissociate attention and action in our task completely, we can ask participants to complete the same action under different attentional constraints and see if this modulates our result. We modulate attention via the size of the distance between the two highlighted squares, drawing from a spotlight theory of attention (Lindsay, 2020; Posner, 1980), which assumes that attention to one area will improve processing of that area relative to others. If attention can modulate our result, we expect a greater effect of misalignment when the target square is further away from the correct square, and/or a greater rate of confusion errors between the two highlighted squares when they are spatially proximate. If the effect is resilient against attentional manipulations, it is more likely that the benefit of selecting the correct location is not merely derived from attention being drawn to it.

Methods

If not described otherwise, the task, analyses and methods should be assumed to be as described in the methods section of chapter II. Studies in chapters II-IV are intentionally constructed to present variations of the same experimental theme on action alignment. We spare the reader a repetition of design features that are kept constant and refer the reader to the methods section of chapter II for all details. Unless otherwise specified, participant selection, exclusion criteria and software packages are the same as in chapter II. The instructions which participants received before each experiment, were slightly adapted to each experiment. Below we report all elements of the procedures and designs that are new or that differ from the procedure and design in chapter II.

Procedure Exp.2.a

We added a third condition to our within-subject design of chapter II (Figure 6 A and B). Here we describe the procedure of this additional condition. In this third, baseline condition, participants took an action *outside* of the response space (Figure 6 C). Specifically, on every training trial, a yellow, squared frame appeared in one of 4 locations left or right of the response space. Which of the 4 locations the frame appeared in was randomised across trials (top left, bottom left, top right, bottom right on the screen). Participants are asked to use their mouse to click on this yellow square next to the response space (Figure 6 C). After the click, feedback is shown just like in chapter II. During this condition, all squares in the response space remain white at all times. No squares within the response space are highlighted (neither grey nor blue). The task is otherwise the same as before (see Figure 1).

We make two minor changes to the action aligned and action misaligned condition, which are otherwise also exactly the same as in chapter II. First, participants now use their mouse to select the target location at training and test. Second, we removed all feedback prediction tests across all conditions (to make the action aligned and misaligned conditions similar to the baseline condition).

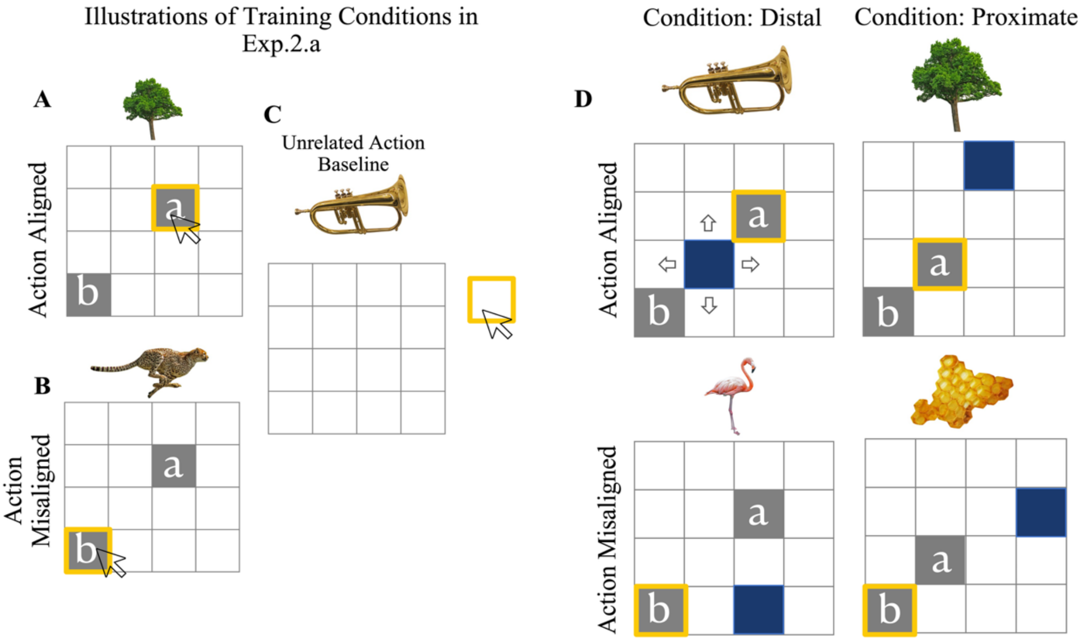

Figure 6: **Illustration for Experimental Designs Chapter III A Exp.2.a** Condition Action Aligned in Exp.2.a: illustration of training trial in which participant clicks on a target location (labelled *a* here, letters not visible for participants) **B:** Condition Action Misaligned in Exp.2.a **C:** Condition with Unrelated Action taken outside of the response space in Exp.2.a. In all training trials, participants click on a yellow frame that appears next to (left or right) the response space at the start of the trial. **D**: Illustration of 2x2 design used in Exp.2.c and d.

Design Exp.2.a

Participants in Exp.2.a (n=54) undergo three conditions: the (new) baseline condition, the action aligned condition and the action misaligned condition. As before, each condition is associated with a new set of images (stimuli as described in methods section chapter II) and the order of conditions was counterbalanced across participants.

In each condition, participants completed three cycles of training and test blocks of 16 trials each. Test trials are the same across all conditions and are equivalent to test trials in chapter II.

Procedure Exp.2.b

We add a different baseline condition, in this case an observation condition, which is experienced as paired with action alignment and misalignment. In this condition, participants experienced the same task as in chapter II (action aligned versus misaligned training blocks followed by tests), except that they never took any actions (they neither clicked or navigated) within the response space.

During the observation condition, on each training trial, they were shown two highlighted squares as before. One of them was framed in yellow (the correct square in the action aligned condition and the incorrect square in the action misaligned condition), just as before. To retain alertness, they were instructed to press the space bar (an irrelevant action) to see the supervision signal (after 900ms as in the other conditions). Just as before, the stimulus moved into its correct location. Instruction messages above the response space alerted the participants to the action they must take (press space bar).

The trial message was adjusted to reflect the given task. E.g. in observation condition of Exp.2.b, it read: “Once you press the SPACE BAR, the image will appear in the CORRECT square with a GREEN frame. That is the square you should try to remember.”

Design Exp.2.b

In **Exp.2.b** (n=86) we used a between-subjects design, in which we compare a group of participants who learn our task by taking a relevant action against a group (n=36) which learned by observation (paired with an irrelevant action). Each participant thus experienced training via two conditions: action alignment and misalignment paired with a relevant action in the response space (a mouse click); or alignment and misalignment paired with an irrelevant (space bar press), which were our pure study conditions. Within each cohort, conditions order was randomised across participants.

Design Exp.2.c & d

In **Exp.2.c** (n=39) and **Exp.2.d** (n=48), we tested if the action alignment effect was modulated by attention. The procedure of Exp.2.c & d is as described in chapter II, so we introduce the only feature which differs below:

We constructed a 2x2 within-subject design, in which participants completed the same task as in chapter II, except that we now vary the distance between the two highlighted, grey squares (Figure 6 D). In two of the conditions, which we term *distal*, participants undergo the conditions action aligned vs misaligned, while each highlighted pair is a Manhattan distance of *4 apart*. These are the two conditions introduced in chapter II (Figure 1).

In the other two conditions (again action aligned versus misaligned) each highlighted pair is a Manhattan distance of *2 apart*, which we term the *proximal* condition.

Each condition is again associated with a new set of semantically distinct stimuli, the assignment of which is randomly counterbalanced across participants and conditions. The order of conditions was randomised across participants, bar the constraint that proximal and distal would alternate.

We repeated this experiment (Exp.2.d) to ensure that the effect was consistent regardless of what motor responses (key navigation or mouse click) are required at training and test. In Exp.2.c participants always made selection using mouse clicks at training and at test, whereas in Exp.2.d we required participants to click at training but use their arrow keys to navigate an avatar square (as described in chapter II) at test.

Analyses

In Exp.2.b, c, d introduce the use of a mixed effects logistic regression model with the same structure as in chapter II, but here we add an additional interaction term:

$$\operatorname{logit}\left(\Pr(Y_{ij} = 1)\right) = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \beta_3 (X_{1ij} \cdot X_{2ij}) + b_{0j} + b_{0jk}$$

The predictor X_{1ij} is the curriculum (e.g. action aligned vs misaligned) and X_{2ij} is the predictor *action vs observation* (in Exp.2.b). In Exp.2.c and d, X_{2ij} is the predictor *proximal or distal*.

For Exp.2.b, we additionally assessed the sensitivity of our model fit to our unequal sample size differences, by use of a resampling procedure. Specifically, we computed the average estimated effect size across 100 separate model fits, each based on a balanced subsample: all 36 participants from the observation condition, and a random sample of 36 participants drawn (with *set.seed(123)* for reproducibility) from the 50 participants in the action condition who passed exclusion criteria. We report the results from both a single model fit to the full dataset (n = 86) and the average effect size obtained from the 100 resampled, size-matched datasets (each n = 72). Contrasts are computed using the package *emmeans* (Lenth, 2024).

Results

In Exp.2.a, test accuracy (shown in Figure 7 A) in the action misaligned condition was at level with the baseline condition (in which participants took an unrelated action outside of the response space) and both the action misaligned and baseline condition were trumped by the action aligned condition (b = 0.55, SE = 0.16, p < .001, 95% CI [0.23, 0.86]). Action alignment at training improved memory performance relative to a study condition in which participants did not take an action inside the response space.

Test confusion errors (Figure 7 B), in which participants selected the distractor location, were higher in the action misaligned condition (b = 1.19, SE = 0.26, z = 4.64, p < .001, 95% CI [0.68, 1.69]) relative to the baseline condition (irrelevant action) and the action aligned condition, but in the action aligned condition these errors were not higher (b= -0.07, SE = 0.29, z = -0.24, p = 0.812, 95% CI [-0.63, 0.49]) than they would be if no grey squares were shown (as was the case in the baseline condition).

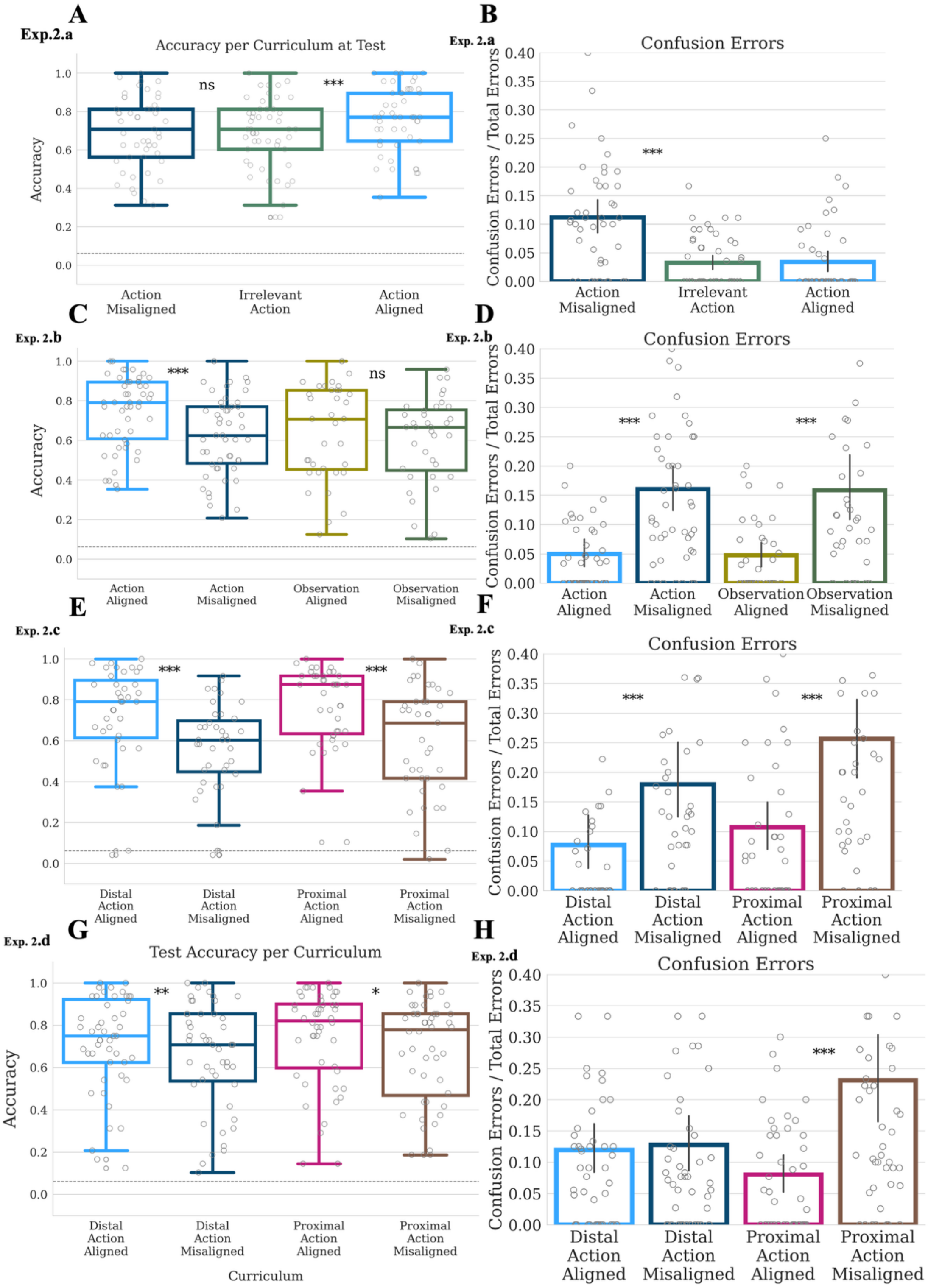

Figure 7: **Results Chapter III A: Exp 2.a** Test accuracy per condition. Each datapoint represents the average accuracy across all test trials per person (resulting in 3

datapoints shown per person in this panel). Coloured boxplots show population summary statistic. New, added baseline condition plotted in green. Significance indicators in plot reflects results from glmms pairwise comparisons. **B: Exp 2.a** Ratio of confusion errors at test per condition. Probability of erroneously selecting the other square that was highlighted in a training trial during a test trial trying to recall that same stimulus. **C: Exp 2.b** Coloured boxplots of accuracy across all test trials per curriculum. Blues represent the active conditions. Each datapoint represents the average accuracy across all test trials of an individual participant: one in an action aligned condition and one in an action misaligned condition. **D: Exp 2.b** Ratio of confusion errors at test per curriculum. **E: Exp 2.c** Coloured boxplots of accuracy across all test trials per curriculum. Blues again represent the active conditions. **F: Exp 2.c** Ratio of confusion errors at test per curriculum.

In the following experiments we investigated the effect of attention modulation in the action alignment effect (e.g. Figure 7 **C**). In Exp.2.b, we used a mixed effects logistic regression model (again adding participant ID and block nested within each participant as random effects. See methods section in chapter II) on the full sample (n=86), with an interaction term between alignment (aligned vs misaligned) and relevant action (action vs observation).

We once again found a significant main effect of misalignment (b = -0.76, SE = 0.16, z = -4.78, p < .001, 95% CI [-1.07, -0.45]). We then looked at this effect within the observation condition. A pairwise comparison between the estimated marginal means of alignment versus misalignment was as expected, reduced (yet to marginal significance: b = 0.34, SE = 0.18, z = 1.86, p = .062, 95% CI [−0.02, 0.70]). A comparison between alignment in the action condition and alignment in the observation condition, shows that *action* alignment produced superior performance compared to *observation* alignment: b = 0.60, SE = 0.29, z = 2.06, p = .039, 95% CI [0.03, 1.17], see (Figure 7 C blue boxplot versus yellow boxplot).

We next tested for an interaction effect to test if the effect of alignment versus misalignment is significantly reduced if we condition on observation. The effect of misalignment was directionally reduced when we tested for an interaction with the observation condition, but was marginally significant (b = 0.41, SE = 0.24, z = 1.71, p = 0.088, 95% CI [-0.06, 0.89]). We added an additional analysis (see methods) to test the sensitivity of this result against the inequality of our samples sizes (observation condition: n=36; action condition: n=50). We randomly sampled 36 participants from the pool of 50 participants in the action condition and ran the same mixed effects logistic regression analysis on a matched sample size (n=72). We repeated this process 100 times and here report the summary statistics of 100 logistic regressions: the main effect of misalignment was statistically significant (<0.05) in 100% of all regressions (average of estimate: b = -0.90, SE = 0.18, z = -4.93, p < .001, 95% CI [-1.25, -0.54]), the main, negative effect of studying by observation was significant in 87% of all regressions (b = -0.71, SE = 0.32, z = -2.24, p = 0.029, 95% CI [-1.33, -0.09]) and the positive interaction effect was significant in 83% of all regressions, with an average effect size estimate of b = 0.56, SE = 0.26, z = 2.18, p = 0.034, 95% CI [0.06, 1.06]. Attention alone so far does not appear to explain the action alignment effect. In Exp.2.b we found that a condition in which attention, but not action, is drawn to a distractor square, produces a reduced effect (Figure 7 C) than if participants take an action to select the square.

We observed the same main effect of misalignment on confusion errors, where misalignment leads to more confusion errors (b = 1.61, SE = 0.25, z = 6.42, p < .001, 95% CI [1.12, 2.10]), see (Figure 7 D). The observation misaligned condition led to as many confusion errors as the action misaligned condition (b = -0.33, SE = 0.37, z = -0.88, p = 0.379, 95% CI [-1.06, 0.40]).

In Exp.2.c (Figure 7 F), we also found no evidence in error rates (however some evidence in error types), supporting the hypothesis that attention plays a role in the alignment effect. Here we tested a manipulation of distance between the true square and distractor square paired with the aligned and misaligned condition at training, on test accuracy. We use a mixed effects logistic regression model with an interaction term between alignment and distance and find the same, main negative effect of misalignment (Exp.2.c: b = -1.13, SE = 0.19, z = -5.90, p < .001, 95% CI [-1.50, -0.75]), but no modulation of this alignment effect via the proximal or distal manipulation (interaction term: b = -0.05, SE = 0.27, z = -0.19, p = 0.850, 95% CI [-0.58, 0.48]).

Against the attention modulation hypothesis, which would have predicted proximal conditions to lead to a higher error rate, we instead find a positive trend (approaching significance) of proximity on accuracy: b = 0.35, SE = 0.20, z = 1.80, p = 0.073, 95% CI [-0.03, 0.74]. We found no main effect of proximity on confusion errors (b = 0.10, SE = 0.33, z = 0.30, p = 0.763, 95% CI [-0.54, 0.74]) and no interaction between proximity and misalignment (b = 0.35, SE = 0.41, z = 0.85, p = 0.393, 95% CI [-0.45, 1.14]), though a marginally significant decrease in confusion errors in the distal, relative to the proximate misaligned condition once we considered paired comparisons (b = -0.4465, SE = 0.242, z = -1.846, p = 0.0649, CI [−0.92,0.03]). As expected, we found a main effect of misalignment on confusion errors (b = 1.66, SE = 0.30, z = 5.59, p < .001, 95% CI [1.08, 2.24]).

In Exp.2.d we repeated this experiment (Appendix A Figure 24 shows accuracy and error rate data from Exp.2.d in the same format as seen in Figure 7 E and F) and the ran the same analysis and confirm the result that attention alone does not modulate the main action alignment effect on accuracy. We again found no interaction effect (b = -0.14, SE = 0.25, z =

-0.56, p = 0.577, 95% CI [-0.63, 0.35]), no effect of distance (b = 0.23, SE = 0.18, z = 1.29, p = 0.195, 95% CI [-0.12, 0.59]), but the same main effect of misalignment (b = -0.35, SE = 0.18, z = -1.95, p = 0.051, 95% CI [-0.70, 0.00]). Confusion errors[2] in Exp.2.d show a significant, positive interaction between proximity and misalignment (b = 0.93, SE = 0.30, z = 3.10, p = 0.002, 95% CI [0.34, 1.52]), no main effect of proximity (b = -0.35, SE = 0.23, z = -1.55, p = 0.122, 95% CI [-0.79, 0.09]) and in this case a non-significant effect of misalignment (b = 0.05, SE = 0.21, z = 0.26, p = 0.798, 95% CI [-0.36, 0.47]). According to a pairwise comparisons analysis, there was no significant difference between alignment and misalignment on confusion errors within the distal condition (b = -0.05, SE = 0.21, z = -0.26, p = 0.798, 95% CI [-0.47,0.36]), but a significant difference within the proximal condition (b = -0.98, SE = 0.21, z = -4.62, p = <.0001, 95% CI [-1.39,-0.56]).

Discussion

We established two baseline conditions and investigated the contribution of confusion errors and attention to our findings from chapter II. Our results indicate that action alignment is beneficial relative to two baseline study conditions: in Exp.2.a an active selection of the correct location led to better performance than a selection of an irrelevant location and in Exp.2.b, active selection of the correct location led to better performance than attending to but not selecting the correct location.

---

[2] Note that confusion error rate in Exp.2.c and d is generally slightly higher across all conditions. This is owed to a manipulation of temporal predictability that we added and which is in line with our expectations. In Exp.2.c and d, we added a temporal correlation between the highlight squares (see Appendix A for details), but since this did not affect the main results or comparison of interest, we do not comment on it further here.

The performance detriment under action misalignment is unlikely to be caused by interference alone. The effect was reduced in the observation condition of Exp.2.b, even though confusion errors were the same in both misaligned *observation vs action* and in aligned *observation vs action*. A baseline condition (Exp2.a), in which interference was not facilitated (actions were outside the response space and no location was highlighted), showed the same error rate as an action misaligned condition, even though confusion errors are significantly lower than when participants selected a distractor. Additionally, in one of our experiments (Exp.2.d, Appendix A Figure 24) we found a drop to baseline in confusion errors within the distal, misaligned condition, but accuracy was still significantly lower than the distal aligned condition, suggesting that confusions errors are unlikely to make up for the bulk of the effect.

The presence of a distractor location does not lead to interference (Exp.2.a) when the distractor is not acted on or visually salient (Exp.2.b). Participants are able to ignore a salient distractor location if they learn in an action aligned curriculum or if the correct location is visually salient in an observation condition. The action of selecting a correct location *a* (either via visual attention or action) potentially protects against distracting features that might have interrupted the encoding of *a*.

Attention alone does not appear to account for the alignment effect, but attention does play a role. In Exp.2.b participants attended to but did not actively select the correct location and we saw that the difference between alignment and misalignment was smaller than if a relevant action was taken. In Exp.2.c and Exp.2.d we also found that attention manipulations do not modulate the action alignment effect in terms of accuracy. This result held up when we ensured that action patterns (clicks) learned during training could not transfer to test action sequences (navigation). If attention was the main reason for the alignment effect, we would have expected the effect to vary with our attention manipulation and/or lead to increased confusions when highlighted squares were proximal to each other. While we found no effect on accuracy, we did observe an increase in confusion errors under the proximate misaligned conditions. Considering that Exp.2.b showed the effect on accuracy is reduced, but not nullified in the observation condition, and that we found no modulation of the accuracy difference, but a modulation of confusions, via attention modulation in Exp.2.c and d, we believe this series of experiments shows that attention is involved in, but not exclusively responsible for the action alignment effect. Our attention controls are not bullet-proof, in parts because we cannot separate action from attention completely and a replication with a greater distance between the squares in the distal condition might find an attentional effect. We discuss the interplay between attention and action in respect to our results below.

Our study bears some resemblance to research on the Attentional Boost Effect (ABE), which is classically found in the recall of items, for which, during the encoding phase, participants are additionally tasked to give or withhold a response when a target or distractor (respectively) is presented concurrently with the item (Swallow et al., 2022). Items paired with the target, i.e. items that required a response, are remembered better than items which were not responded to during study. This has been linked to the benefits of active learning in the classroom (Prince, 2004) and relates to our results, in so far as in our task an action modulates event encoding. The ABE is prima facie a surprising effect, because it shows that participants do better while completing two tasks rather than one (Prull et al., 2023). Under an assumption of cognitive resource limitation, such an effect demands explanation.

Pupil diameter and fMRI measurements suggest that the origin of the effect stem from a noradrenergic release by the locus coeruleus during goal-directed action, thereby upregulating attention and alertness (Sajedin et al.,2019). Our study again differs significantly from these experiments in that it does not measure the difference between withholding and taking an action in memory tasks, but instead *always* requires the same action (same except for its alignment with the supervision signal). ABE experiments, while similarly interested in how actions can boost memory, are likely detecting an effect of modulated alertness. Since we require an action in both conditions, there is little reason to expect this to explain our results, as this would require an assumption of higher alertness when action is aligned with feedback.

The ongoing search for an explanation of the ABE however highlights an interesting distinction that is relevant to us. Do the reported memory boosts require an active response? Or does attention suffice? It appears not. Some researchers stress that the ABE is only found when participants make an active response (Toh & Lee, 2022), going so far as to describe the effect as *action-induced memory enhancement* (Yebra et al., 2019), while other scholars still find the ABE in experiments that merely require covert target detection among distractors (J. Y. Lin et al., 2010; Swallow & Jiang, 2013) or even when the response is an action-cancellation upon target detection (Makovski et al., 2013). It has been suggested that action vs target-detection induced memory enhancements might be two separate phenomena (Yebra et al., 2019), or indeed that a more general mechanism is at play. Swallow et al. (2022), try to unify these desperate results under an account of attentional boost as event boundary detection, as increased alertness during a shift from one cognitive state to another, or more generally described as “the transient mobilization of effort to compensate for differences between internal states and the external world,” (p.15 Swallow et al., 2022).

Note that this interpretation too does not immediately lend itself as an explanation of our findings, since it actually predicts boosted memory following our action misaligned curriculum. Their theory does however deal with the question of whether attention or action is the culprit of a memory boost, in parts by making way with a neat distinction between attention and action (p.6, Swallow et al., 2022). The contrast is instead understood as a continuum in which attention is considered a kind of covert, internal movement, stemming from the same circuits that evolved for overt effector activation.

They are not alone in relying on this line of argument. In an article on visual working memory (vWM), Olivers & Roelfsema (2020) consider actions to be planned program executions of a goal-directed motion by an effector, which, crucially, do not necessarily have to be overtly executed. In line with the embodied cognition tradition, the authors consider vWM to fundamentally be in service of actions and their outcomes, whereby 'successful action must reinforce the sensory and memory representations that led to them' (p.181, Olivers & Roelfsema, 2020). They thus propose a bi-directional relationship where task-relevant actions or action plans modulate what is perceived and memorised and where percepts that previously led to successful action outcomes are retrieved and consolidated more readily. Attention is modelled as an action-evaluation mechanism, which gates cognitive resource allocation (Lindsay, 2020) towards those percepts that are likely to be useful for actions. The authors present this mechanism as a relevance-rating: since perceptual acuity and attention consume limited cognitive resources, action-relevant items are prioritised. They suggest this might apply both for actions that are being taken at that time, and importantly, for actions that are likely to be taken in the future.

While this proposal was put forward to explain findings in vWM, the listed evidence showing how actions affect memory encoding, bears similarities to our study. This evidence ties memory encoding, attentional and perceptual resources to items that are action-relevant, thereby offering corroborating evidence and an explanation for why our enforced action regime might modulate memory. Deubel & Schneider (1996) showed that an intended, planned eye-movement to a spatial location increased discrimination accuracy for visual stimuli which were presented in that location *before* the eye-movement was made. Visual discrimination was enhanced by a planned action. Heuer et al. (2017) constructed an vWM experiment in which, in between a memorisation and a recall phase, participants pointed to a cued target location. Even though the memory items were no longer shown when the movement was made, memory for the items which had been in the location that was pointed to, was better than if the test item was in a location that was not pointed to. The effect was stronger when more items needed to be remembered, suggesting a selective, prioritising effect of action when the option space is large.

Considering these results and theoretical proposals, we suggest it may be possible that in our action misaligned condition, we observe upregulated encoding of location *b* in memory (Figure 1), because it was made action relevant. We find evidence for this coupling theory in our confusion error rate. Participants do not only show a reduced effectiveness in encoding the correct (non-target) location *a*, by diverting action/attention away from the correct location. They also show an increased recall of the incorrect (target) location *b* at test, which tells us that the action selecting location *b* leads to upregulated consolidation of *b*. Future studies might test controls in which the location *b* is made task relevant but not action relevant.

In this framing, participants in our action aligned condition reap the benefits of action-relevant perceptual acuity and/or increased consolidation strength for the correct location. In our action misaligned condition however, participants must additionally learn a gating operation in which they use expected future reward to gradually allocate more attention to the correct location, which they know will be *action-relevant in the future*. In Olivers & Roelfsema (2020), attention is proposed as a credit-assignment mechanism, mapping a simulated, future action (in our case the selection of the correct location at test) to the reward and the stimulus. Attention is known to stabilise encoding (Muzzio et al., 2009), suggesting that a version of our experiment with eye-tracking would help disentangle the effect of stable attention (aligned condition) from less stable attention (misaligned condition). While an overt action will inevitably drag some attention to the incorrect location, humans are also known to be able to dissociate overt action (eye-movement) and attention (covert action) to some extent

(Lindsay, 2020; Morishige et al., 2021). Gating is required (to ignore the distractor square), but much easier to accomplish in the action aligned case, in which we observe that the second highlighted square (a salient feature that could draw bottom-up attention) is not recalled more often than in a baseline condition (Exp.2.a) in which no highlighted squares are shown. In debrief responses across our experiments, participants routinely reported that the action misaligned condition felt more challenging to them, which is consistent with the idea that they are learning a more effortful gating operation. The fact that participants can learn this gating operation, in which they inhibit or override the strong encoding of the incorrect, target location by recalling the correct, non-target location at test, is evidenced by the fact that participants in our action misaligned condition still learn the correct associations over time. Their memory is not fully (but partially) at the mercy of their actions.

Chapter IV Generalisations of the Action Alignment Effect

Abstract

We present five empirical studies which generalise the action alignment effect and place it in context of curriculum design and teaching. First, we report a study that allowed us to compare the action alignment effect against the classical effect of the retrieval practise test and find, surprisingly, that action alignment achieves superior memory performance compared to a condition in which participants can test themselves. Second, we use action alignment to design a curriculum through which we are able to prioritise the encoding of a subset of learning materials. Lastly, we investigate if the origin of an action (origin self versus origin other), modulates the action alignment effect and find that students might even be able to benefit from action alignment, when skilled teachers select the actions.

Introduction

This chapter presents empirical work which extends our knowledge about the action alignment effect by generalising it to adaptation of the task and by comparing it with other classical effects in the memory literature. Our interest lies with the applicability and generalisability of the effect. In what ways can action alignment assist learners and how does the effect fare in comparison with other curriculum designs?

One of the most well-evidenced behavioural effects on memory strength is the so-called testing-effect, already introduced in chapter I. Test-enhanced learning (RoedigerIII & Karpicke, 2006), has repeatedly been found to boost long-term memory, and this effect stems from active retrieval of the learning materials (Adesope et al., 2017). Retrieval is more beneficial even if compared against the same number of exposures to the study material (RoedigerIII & Karpicke, 2006). For instance, a repeated attempt to solve the same multiple-choice test, positively affects memory in comparison to being shown the correct answers the exact same number of times. The testing effect has stood the test of time, translates to classroom conditions and has been found to support learning in a variety of task and feedback settings and retention delays (Agarwal et al., 2021).

The action alignment effect appears to be procedurally at odds with the ability to benefit from retrieval practice. In our task (across both conditions), the information about the correct location is available as soon as the (re-)training trial begins, thereby providing no chance for participants to test themselves via retrieval. Yet, we have shown that participants benefit from action aligned training. How does the action alignment effect fare in comparison to a condition in which retrieval practise is possible? We will attempt a comparison by use of an experimental design which, in essence, presents an abstract version of a multiple-choice test (often used in classrooms). We will vary the number of choices from multiple to few, to only a single choice, thereby obtaining training trials that are akin to our action alignment condition, in which participants can only select the correct choice. In previous chapters, the probability of action alignment at training was deterministic: an action was aligned or not. The testing effect however depends on a scenario in which the probability of a correct choice is less than certain. We will introduce uncertainty into our design, to give way to the testing effect. Participants will have an opportunity to test themselves before making a choice. We will specifically test if a combination of the action alignment effect *and* the testing effect (k=2), yields better memory performance than the action alignment effect by itself (k=1).

Curricula in schools introduce teaching materials in a sequential order. Sometimes, learning a concept is easier when pre-cursors of the concept have already been learned (Bengio et al., 2019). Teachers implement curricula that teach students one thing before another and do so because pieces of knowledge depend on other knowledge or representations. We here ask whether the action alignment curriculum could assist in such efforts. Many tasks and generalisation conditions require humans to chunk and group items or decompose stimuli, action plans or complex tasks, into subcomponents. Some tasks cannot be accomplished without doing so. For instance, research has shown that a decomposition of stimulus dimensions (such as colour and shape) tends to occur in a sequential order, such that one dimension is encoded before another (Dekker et al., 2022). In this chapter, we report a variation of our previous design, through which we can show that action alignment can be used to prioritise learning of an arbitrary subset of stimuli. We specifically test if action alignment could be used to support curricula for sequential knowledge acquisition. To this end, we construct a curriculum in which a subset of items is studied under action alignment (k=1), whereas another subset is studied without any helpful constraints (k=16) which would align action and feedback. We will test the hypothesis that items which were studied under action alignment are learned by preference.

In a last experimental design, we then turn to question whether the origin of the action determines the action alignment effect. Another classic effect on memory encoding is the so-called self-reference effect (Symons & College, 1997), which shows that learning materials associated with one's representation of self are remembered better than those associated with others (Ganellen & Carver, 1985). If self-reference plays a role in the action alignment effect, its application is limited to all learning contexts in which the learner can take the actions themselves. In our designs so far, the actions taken are simple, but outside the lab, this is not always the case. Is the action alignment effect flexible enough to transfer to scenarios in which skilled teachers need to take the actions?

We will test if the action alignment effect differs depending on if the action originates from oneself versus another participant. We do this for two reasons. This study allows us to further pinpoint the origin of the action alignment effect. Does it elicit the self-reference effect at all? How relevant is the execution of the motor action itself? The actions taken in an online experiment are small motor actions, eye-movements and head movement. Is it the motor movement that modulates the effect on memory or is it merely the representation of an action being taken? It allows us to test a generalisation of the effect. Do students always need to take the correct action themselves to benefit from action alignment?

We will use a within-subject design that combines the design from chapter II, with a condition in which participants observe reconstructed actions (both congruent and incongruent) of another person, instead of performing the actions themselves. We will test the action alignment effect once more, conditioning on the origin of the action.

Methods

All unspecified details in the following sections are as described in the methods section in chapter II. This includes participant selection, exclusion, basic structure of the procedure and design on action alignment, statistical analyses and software packages used. Below we introduce all new variations and amendments which were not specified in chapter II.

Procedure & Design Exp.3.a-c

On each training trial, participants see a subset of $k = \{1, 2, 4, 8, 16\}$ squares highlighted in grey (Figure 8 C), one of which is the correct location. In previous chapters, $k = 2$ and participants were only able to select only *one* of the highlighted squares. Here we now vary the number of highlighted options and leave the choice of the target up to participants. Participants click on one of the grey locations to make their choice. No grey square is framed in yellow. Upon a click during training, feedback was now visualised as a green square in the correct location, while the stimulus remained located on top of the response space. Test trials were as before.

We manipulate the probability of an aligned action during training. By varying k, we effectively vary the number of possible actions and thereby control the probability that the action will align with feedback. When $k = 1$, the probability of an aligned action is 1. When k

= 2, the chance of a correct choice at first try is 0.5. When k = 16, it is as if the option set was not reduced and the probability of an aligned action is 1/16. This is another baseline condition, using self-selected responses.

In **Exp.3.a** (n=35), participants experienced four conditions (in random order), each of which is distinguished by the number (k = {1, 2, 4, 8}) of grey squares visible on each training trial (k stays constant within each condition) and a new set of stimuli. **Exp.3.b** (n=29) repeats Exp.3.a and in **Exp.3.c** (n=44), participants experienced three conditions k = {1, 4, 16}. The order of subsets presented within a condition was randomised across blocks and members of each subset were kept constant throughout a condition. In all experiments, each condition is still experienced as three cycles of training and test, where test trials are the same across all experiments.

A variation in visual feedback allowed us to test that the action alignment effect is not dependent on the specific way in which we had presented feedback previously. We here also put the testing effect to the test: if the testing effect is more beneficial for encoding a spatial declarative memory, we should see improved performance for conditions in which the probability of the outcome of a choice is not guaranteed to be correct, i.e. the probability of action alignment is less than 1. If however, k=1 leads to strictly better performance than any of the other conditions (including k = 2), the benefits of the testing effect do not outweigh the benefits of action alignment in our task.

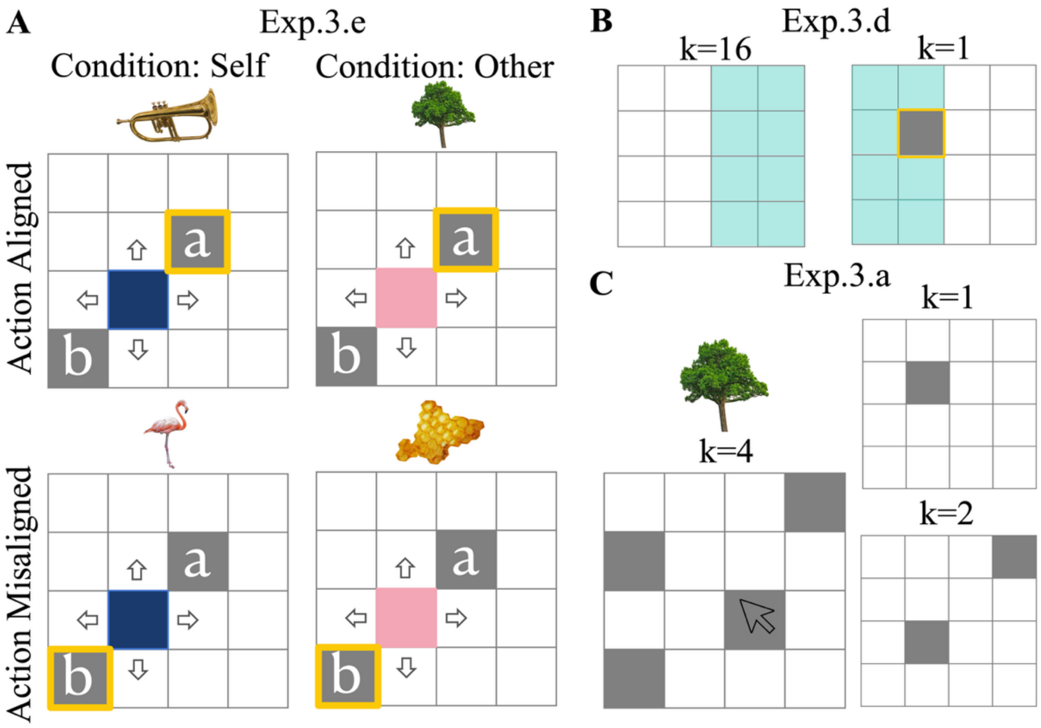

Figure 8: **Illustration of Experimental Designs Chapter IV A: Exp.3.e** 2x2 design crossing action aligned vs misaligned (rows) against *self* and *other* (columns). In condition *self*, the participant navigates the blue avatar. In condition *other*, the participant observes the reply/reconstruction of the navigation behaviour by a previous participant. **B: Exp.3.d** Half of all stimuli are trained using the action aligned condition with 1 square highlighted (k=1). The other half are studied without any highlighted square. **C: Exp.3.a** shows three conditions, varied by k = the number of highlighted squares on a training trial. E.g. on a training trial for the tree stimulus, when k=4, four squares will be highlighted. The correct square is always one of the highlighted squares. Participants select their choice using a mouse click. When k=1, participants have certainty over the correct location and thereby no ability to test themselves. When k>1, they have varying degrees of uncertainty.

Procedure & Design Exp.3.d

In the first block of each set, participants were shown an all-white response space without any highlighted squares and could select any of 16 squares using their mouse (k=16) (Figure 8 B). Upon a mouse click, the square that was clicked would turn blue for 100ms to indicate that a response was recorded, followed by the usual feedback signal (stimulus moves into its correct location and is framed in green after 1.3 seconds). The next trial began as usual after 1.3 seconds.

In the second block, unbeknownst to participants, stimuli-location pairs were then assigned at random to two groups of size 8 (S_i, S_j). In the second block, half of all stimuli (according to group selection S_i) were presented with the correct square framed in yellow (Figure 8 B, k=1). This square was now the only response option (k=1), essentially forcing the action to be aligned with the ensuing supervision signal. If participants clicked anywhere else on the screen, nothing would happen. All remaining stimuli (S_j) were shown without any restriction of responses, just like in the first block. The order of stimuli presented at training was randomised, which meant that participants could not predict if the next trial would present a stimulus with one (k=1) or all (k=16) response options. In the third block, the same half of the stimuli (S_i) were again shown with it's correct square as the only option, while the remaining half (S_j) was shown again without any restriction (k=16).

In **Exp.3.d** (n=39), participants played two sets of three cycles of alternating training and test blocks, in which they made a mouse click to select a square. This split of stimuli into two groups (S_i, S_j) was randomised across participants. We set out to measure the interaction between block (={1,2,3}) and action alignment (k={1,16}) on test performance.

In Exp.3.d we test if the action alignment effect can be used to make selective, preferential and significant progress in learning a chosen subset of stimuli. We ask if an arbitrary selection of stimuli will be encoded faster than the remainder of stimuli, if only the subset is trained via action alignment.

Procedure & Design Exp.3.e

At the start, participants were instructed that they (represented as a blue avatar) would play the game together with a partner (represented as a pink avatar) (Figure 8 A). In every training trial, they were told, the avatar had to be moved into the target square, but in some training trials it was the participant's task to do this (blue avatar shown, condition=self), whereas in others it was the partner's task to navigate (pink avatar shown, condition=other). No matter who the action originated from, participants were told they had to learn the association as best as they could and that they would afterwards be tested (and earn bonus points) separately from their partner.

As in previous designs, participants were instructed that sometimes they or their partner would be required to navigate to the square with a yellow frame even though this "is not the correct square for that image", but that even so, their task was still "to remember the correct square, not the square in yellow".

At the start of a trial, a trial message and the colour of the avatar would tell participants whether they or their partner had to navigate to the target square (a, framed in yellow in Figure 8 A). Conditions in which participants navigated themselves, proceeded as in chapter II (e.g. Exp.1.a). Conditions in which participants did not need to navigate, still asked them to press the return key to start the training trial. They were then asked to watch the navigation sequence of a partner. A message read: "Training Trial! Press ENTER to let your partner start the trial." And after they had pressed the key: "Watch while your partner moves THEIR PINK square to the square framed in YELLOW! The image will then move into the square it belongs to. Remember this square. You will be tested later!" Test trials preceded as usual:

participants click on any of the white squares, the square turned blue to capture the response and no feedback was shown.

Exp.3.e (n=128) was a pre-registered (Cremer & Summerfield, 2024) experiment, designed to test if the action alignment effect is also present when participants watch a game partner take the action, rather than take the action themselves.

They never actually played with a partner. We used the recorded navigation data (paths and reaction times) by 25 participants from Exp.1.a and reconstructed their behaviour in the game environment so that participants in Exp.3.e could observe an aligned or misaligned action being taken. A participant's experience from Exp.1.a (the order of locations queried, the time taken for each trial, the order of condition aligned versus misaligned), was replayed to a participant from Exp.3.e in their experience of condition 'other'. We did this by using the data recorded from training trials of participants in Exp.1.a (the exact locations they had moved their avatars through and their reaction times) to determine where and when the pink avatar (which indicated the partner) would move during each training trial. Each participant from Exp.3.e thus watched a faithful reconstruction of real human behaviour, rather than a simulation of behaviour. Each participant from Exp.3.e was assigned one of 25 participants reconstructions from Exp.1.a. They watched the same participant's reconstruction in both the action aligned and misaligned condition.

Exp.3.e included four conditions (Figure 8 A) resulting from a 2×2 factorial design crossing action alignment (aligned vs. misaligned) with action origin (self versus other). Conditions were randomised but alternated between *self* and *other*. Each condition again included three cycles of training and test.

Note this is different from other baseline conditions (Exp.2.a and Exp.2.b) in which participants observed the supervision signal without any relevant action being taken or being presented on screen. In Exp.3.e participants observed relevant actions, but they did not take them themselves.

Each condition also included feedback prediction trials as in Exp.1.a. To assess performance differences in feedback prediction trials, we used repeated-measures t-tests (and conservatively made no correction for multiple comparisons). Since each participant was assigned to one of 25 unique reconstructions of navigation patterns by a participant in training trials of Exp.1.a, we captured individual variation between those ‘partner players’ by adding the unique participant IDs from Exp.1.a as an additional random effect in our statistical analyses.

Appendix A details minor variations in the designs which did not affect the results, but which can be consulted to see how the effects generalise across differences in the presentation and appearance of the task.

Results

Exp.3.a-c

Exp.3.a (Figure 9 B) and Exp.3.b (Appendix Figure 24) both showed that k = 1 resulted in superior memory at test compared to $k = 2$ (**Exp.3.a**: b = -1.07, SE = 0.18, z = -5.92, p < .001, 95% CI [-1.42, -0.71]; **Exp.3.b**: b = -0.63, SE = 0.19, z = -3.27, p = 0.001, 95% CI [-1.01, -0.25]), or $k = 4$ (**Exp.3.a**: b = -1.20, SE = 0.18, z = -6.65, p < .001, 95% CI [-1.55, -0.85]; **Exp.3.b**: b = -0.88, SE = 0.19, z = -4.57, p < .001, 95% CI [-1.25, -0.50]) or $k = 8$ (**Exp.3.a**: b = -1.71, SE = 0.18, z = -9.39, p < .001, 95% CI [-2.07, -1.35]; **Exp.3.b**: b = -1.42, SE = 0.19, z = -7.39, p < .001, 95% CI [-1.79, -1.04]). The action alignment effect supersedes a condition of uncertainty in our spatial declarative memory task.

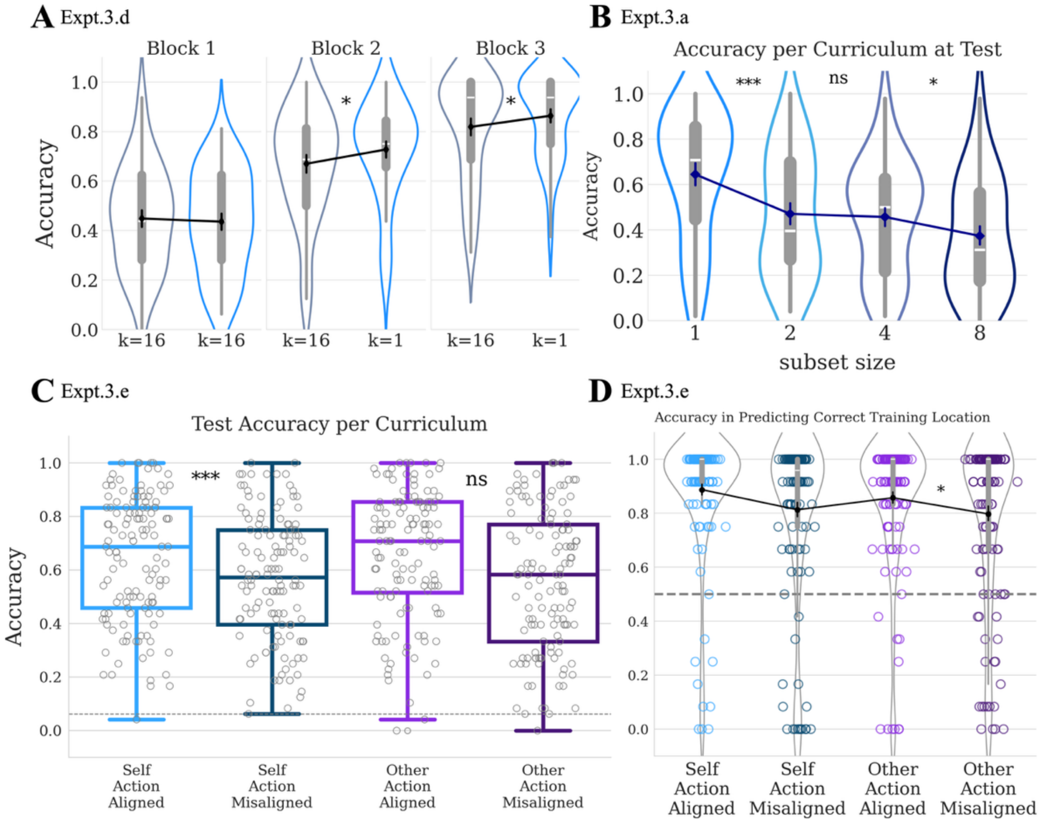

Figure 9: **A:** Test Accuracy in Exp.3.d per block for all stimuli presented under condition k=1 or k=16. **B:** Test accuracy Exp.3.a across all test trials, after training under no uncertainty (k=1) versus increasing uncertainty (k>2). **C:** Test accuracy Exp.3.e for self (blues) vs other (purples). Each shown per action aligned and misaligned condition. **D** Feedback Prediction Test accuracy Exp.3.e.

In **Exp.3.a**, *k = 2* did not lead to significantly better test performance than *k = 4* (b = 0.13, SE = 0.18, z = 0.76, p = 0.450, 95% CI [-0.21, 0.48]), but we found a significant difference in *k = 4* > *k = 8* (b = -0.51, SE = 0.18, z = -2.87, p = 0.004, 95% CI [-0.86, -0.16]). The same pattern holds in **Exp.3.b** where *k = 2* was not superior to *k = 4* (b = -0.24, SE = 0.19, z = -1.30, p = 0.195, 95% CI [-0.61, 0.13]), but *k = 4* was significantly better than *k = 8* (b = -0.54, SE = 0.19, z = -2.91, p = 0.004, 95% CI [-0.91, -0.18]). In **Exp.3.c** (Appendix Figure 24), we find again that a higher frequency of action alignment benefits memory encoding, where *k = 1* > *k = 4* (b = -1.31, SE = 0.16, z = -8.33, p < .001, 95% CI [-1.62, -1.00]) > *k = 16* (b = -1.53, SE = 0.16, z = -9.74, p < .001, 95% CI [-1.84, -1.22]), showing that action alignment achieves superior memory performance over pure study (*k = 16*) and that it trumps conditions in which both action alignment and the testing effect could potentially benefit learning (*k > 1*).

Exp.3.d&e

In Exp.3.d we found a significant and positive interaction effect between block 2 and action alignment (b = 2.19, SE = 0.18, z = 12.17, p < .001, 95% CI [1.83, 2.54]), as well as between block 3 and action alignment (b = 0.48, SE = 0.21, z = 2.25, p = 0.024, 95% CI [0.06, 0.89]) showing that the subset of stimuli learned via an action aligned curriculum (which we introduced in blocks 2 and 3) were encoded faster than stimuli that were learned via trial and error (Figure 9 A). Across two sets, we expectedly found no difference in performance for stimuli from S_i vs S_j in block 1 (b = -0.06, SE = 0.12, z = -0.50, p = 0.616, 95% CI [-0.31, 0.18]), showing that mere exposure did not advantage or preferentially encode a subset of stimuli. Irrespective of the stimuli or locations in subset S_i or S_j, action alignment led to faster encoding of stimulus-location pairs.

The action alignment effect is present even when participants do not take the action themselves but instead watch a game partner make the selection (Figure 9 C). A linear mixed effects model applied to test accuracy in Exp.3.e, shows the same main effect of action misalignment (b = -0.45, SE = 0.10, z = -4.33, p < .001, 95% CI [-0.65, -0.25]), but no main effect of *other* (b = 0.07, SE = 0.10, z = 0.67, p = 0.505, 95% CI [-0.13, 0.27]) and no interaction between misalignment and *other* (b = -0.17, SE = 0.15, z = -1.17, p = 0.241, 95% CI [-0.46, 0.12]). Feedback prediction trials (Figure 9 D) showed that participants were more unsure about the correct location during training in the action misaligned condition in both *self* (t(127) = -1.80, p = .074) and *other* (t(127) = -2.19, p = .031). We found the same pattern when selecting for participants who were above chance in feedback prediction trials: the interaction (b = -0.17, SE = 0.15, z = -1.17, p = 0.241, 95% CI [-0.46, 0.12]) and the effect of *other* was insignificant (b = 0.07, SE = 0.10, z = 0.67, p = 0.505, 95% CI [-0.13, 0.27]), but the main action alignment was effect was still present (b = -0.45, SE = 0.10, z = -4.33, p < .001, 95% CI [-0.65, -0.25]).

Nor were confusion errors affected by the origin of the action. When we applied a linear mixed effects logistic regression to the number of congruent errors out of all errors, we observed the same positive main effect of action misalignment (b = 0.70, SE = 0.13, z = 5.19, p < .001, 95% CI [0.43, 0.96]), but no effect of *other* (b = 0.15, SE = 0.15, z = 1.01, p = 0.315, 95% CI [-0.14, 0.44]) and no interaction effect between action origin and misalignment (b = -0.08, SE = 0.19, z = -0.45, p = 0.654, 95% CI [-0.45, 0.28]). We conclude that action alignment is an effective learning curriculum even when the learner does not perform the action themselves.

Discussion

We found a condition in which learning via action alignment trumps a condition in which participants were able to test themselves. We found that we can induce preferential encoding of an arbitrary subset of stimuli using the action alignment effect, and that the effect is present even if participants do not take the action themselves.

Interestingly, conditions which in theory leave room for both the testing and the action alignment effect, such as k=2, also led to worse performance than k=1. When two options are presented, participants have the option of testing themselves. If they make the correct choice, they can benefit from an alignment between the action and the feedback signal. We might therefore expect a double boost of the memory encoded in trials in which participants made the correct prediction in k=2. This could have a larger effect on accuracy than k=1 where self-testing is not an option. However, this is not what we observed, suggesting that action misalignment in k=2 is more detrimental than the benefit from self-testing and action alignment in our task.

Note that we implemented no test to verify that participants did in fact test themselves. It is possible that they behaved randomly during training, but our previous chapters on action misalignment suggest that they are unlikely to click on a location which they do not believe to be correct. One other explanation for why we do not observe a testing effect might be that the testing effect tends to apply in delayed memory tests and has sometimes been found to be stronger after delayed rather than immediate recall tasks (Ferreira & Wimber, 2023; Rivers, 2021). Future studies might examine temporal dependencies of action alignment and test-enhanced retrieval. We might for instance test the hypothesis that our design yields an advantage for k=2 over k=1 if tests are administered after longer delay periods (e.g. 1 week), after which the testing effect might override the action alignment effect. Our results obviously do not suggest that self-testing should not be used but instead suggests that future research might identify how learning environments can make optimal use of both the testing effect and the action alignment effect and identify how they might interact or how they should be sequenced (Rawson & Dunlosky, 2012).

In Exp.3.d, we found that we can use the action alignment effect to predictably encode a particular and arbitrary subset of stimuli before another set of stimuli. As a proof of concept, we suggest that action alignment could be used to shape learning and representation. Dekker et al. (2022) found that generalisation through visual stimulus decomposition is better learned through a curriculum that temporally correlates training trials along one stimulus dimension (e.g. shape) before training on variations of another stimulus dimension (e.g. colour). We here present evidence that action alignment could also be used as a curriculum in a similar fashion. Future research might for instance investigate whether a temporally *uncorrelated* curriculum, which however uses action alignment in combination with trials that only vary a single dimension, can achieve similar generalisation performance as a correlated curriculum.

In the discussion section of chapter II, we considered the suggestion that the action alignment effect might be related to agency. Schreiner et al. (2025) and Ren et al. (2025) reported results in a similar paradigm, in which participants were allowed to select their action freely. For trials in which the self-selected action was aligned with feedback, memory was improved, which led the authors to suggest a memory effect derived from a sense of agency. In this present chapter we reported results that seems to suggest this explanation does not apply, at least in our task. We found that a condition in which participants had no action choice or agency over the navigation path (because a simulation of navigation was observed rather than executed), still yielded an action alignment effect.

In our self versus other comparison, we told participants that the other navigator was a game partner. Clarkson et al. (2022) found the self-reference effect could be reduced by a salient association with another participant. Moreover, the action sequences in the reconstruction (other) looked like the action sequences of the participant (bar the colour of the avatar). This was a deliberate choice which allowed us to the pinpoint the effect to whether the action was taken or not. We cannot exclude the possibility that our framing led participants to associate themselves with the game partner's navigation path to the extent that a self-reference effect contributed to memory encoding. Future studies might investigate if the effect holds of game play *opponents*, if the effect remains if participants are told that they are watching a simulation rather than a person or if the effect holds when an action *sequence* is omitted. We have not shown that actions do not matter, but we have shown that at least in the task presented here, the action does not need to be taken by the learner themselves. This makes the action alignment effect potentially applicable in a wider variety of teaching settings. Future research might discern what actions, what types of feedback, what testing delays and what action-origins are able to modulate the action alignment effect. This finding raises questions about the precise mechanisms within procedural memory that may be at play. While the effect is framed as an *action* alignment effect here, it is worth asking whether and to what extent motor action are implicated. The finding that observing the action of a game partner is sufficient to induce the effect, suggests that a motor action *representation* may be sufficient. Daprati et al. (2019) indeed found increased recall of enacted actions in a patient who lacks proprioceptive feedback, suggesting that the related enactment effect can occur without proprioceptive feedback. The authors offer the suggestion that the neural representation of intending to move is the neural mechanism that underpins the memory effect. In support of this suggestion, we find the effect under minimal motor movements (such as mouse clicks), which lack distinction between trials and which we keep constant across all trials. Future research could pinpoint the effect to what neural mechanism in motor learning and execution is responsible for the majority of the behavioural effect we observe.

We presented an investigation into the proposition that the representation of self makes a difference to the action alignment effect. We will now flip the question and consider the role of memory in the representation of self. We suggest that the representation of self is, like all memory, a compressed model which can be used to reconstruct one’s own action sequences.

Chapter V

Abstract

In this chapter we lay the groundwork to study participants' metacognitive access to their own learning, by introducing behavioural markers of learning in a spatial rule discovery task. We introduce a task in which we are able to capture behavioural reflections of a cognitive transition from ignorance to knowledge about the spatial rule in behaviours. Behavioural markers reflect participants' epistemic state across accuracy, reaction times and mouse cursor movement. We discuss the method of *dynamic reconstruction* of participants game play, which we used to generate and refine hypotheses and behavioural measures.

Introduction

The promise of metacognition is that learners know themselves well enough to select actions that aid their efforts. As we will see, metacognition is usually measured with respect to specific percepts and via confidence measures, but recently, scholars have begun to study people's models of their own learning at a higher level of abstraction and have begun to search for task environments, tools and measures that allow for the study of metacognition applied to more multi-dimensional cognitive phenomena. We will briefly discuss how people's models of themselves relate to memory and how we can study these models by eliciting them. In this chapter (V), we present an experiment and new behavioural measures of learning via *discovery* and discuss the tool of *dynamic reconstruction*. This will provide the basis of chapter VI in which we extend this task to study peoples' models of their own discovery process, by eliciting their self-model to help them guide their action choices.

Metacognition

The discipline of psychology exists because humans have incomplete introspective access. If we could know the inner workings of our minds by introspection alone, we would not need researchers to look into it. Why do we find some research results in psychology surprising, but others less so (Mazor et al., 2023)? It is because they appear 'counter-intuitive', in seeming contradiction with what we think we know about our minds from experience (Ward & Scholl, 2015). They are surprising, because they violate our model of ourselves. Thus, self-knowledge exists, albeit in incomplete form and as systematically false.

The field of metacognition - the study of how the mind applies itself to itself - is ever in search of the boundaries of what we do and do not know about ourselves. Fleming et al. (2012) distinguish between the metacognitive access to cognitive content and process. The distinction might be analogous to perceptual or representational inputs and the procedure or algorithm that operates over those inputs. In the following two chapters, we are interested in peoples' cognition about the latter.

Metacognition is studied with respect to specific cognitive domains, such as memory, theory of mind or perception. The very assessment of whether one's memory is reliable is a metacognitive process, indeed one of the earliest types of metacognition to be scientifically studied (A. L. Brown, 1978; J. Brown et al., 1977; Flavell, 1971). Researchers found that metamemory is possible, but faulty and develops over time. Metacognitive research is ever since on the search of those boundaries of what parts of ourselves are and are not subject to metacognitive access and improvement (Flavell, 1979; Morris, 2025; Rivers, 2021; Rouy et al., 2022).

The main measurements for metacognition, are indicators of confidence (the degree to which the participants believe their cognitive process or content to be trustworthy).

Confidence is a second-order judgement that reflects the reasoner's evaluation of their own cognitive system. Measurements often control for performance and response-biases.

In Mazor, Moran, et al. (2024) for instance, confidence ratings were used to study peoples' inferences over their own perceptual system. Participants were found to treat their perceptual evidence differently, depending on whether they judged a stimulus to be present or absent. Other studies have used participants' predictions about their own cognitive abilities. In Mazor et al. (2023) for instance, participants made predictions about their visual search process (such as the effects of distractors on search times). Participants' awareness of their own visual system again showed metacognitive accuracy, as well as blind spots. The authors found for instance that participants' predictions about search times were indeed correlated with their own true search times more than with another participant's search times, reflecting the possibility that cognition about oneself does indeed use predictive information that only the participant in question is privy to at the time. Participants however failed to account for difference in search times induced by different types of distractors.

These are a few of many examples of research papers that find (imperfect) metacognitive accuracy across different domains. Katyal & Fleming (2024) review recent progress in metacognitive neuroscience but warn that domain-specific measurements of metacognition could narrow the rich and encompassing range of experiences that are associated with metacognition. This, they suggest, includes broader representations and beliefs about oneself, such as one's confidence to be able to gain new skills. These global self-performance estimators (Rouault et al., 2019), are suggested to hold influence over what actions we decide to take, such as embarking on learning a new skill or not. Researchers have stressed the importance of measuring and amending such self-beliefs in the context of mental health (Seow et al., 2021).

Models of Self

Mazor (2025) suggests we not only study self-beliefs, but self-models too. Self-models are introduced as structured representations that compress experiences of one's own cognition and actions (such as attentional biases, preferences or sensory perception) into a generative model through which the agent can make predictions about their own cognition and actions. The suggestion is that our biases, preferences and needs are noisy and require the agent to pose hypotheses and models about their inner workings. It is thus a model akin to other models that agents build of dynamic, structured environments to predict perceptual events, simulate counterfactual scenarios, infer latent causal structure and generalise to variations of experience. The difference is that in this case, the environment is our own body and mind, both of which are stochastic processes with causal and mechanistic constraints that permit predictably.

Self-models are partly based on memories of one's skills, emotions or preferences, (consider 'I always preferred to spend a day in the garden than to be sat inside'). They are refined by experience (consider the differences between how I thought I would react and how I did react). We remind the reader of the resource-efficient and compressed (Dasgupta & Gershman, 2021) reconstructive theory of memory introduced in chapter I, where Spens & Burgess (2024) propose that new episodic memories are formed when the prediction error between the generative model prediction and a new observation is sufficiently large (surprising). We reviewed literature which suggests that compressed models of experience are used to imagine and plan for the future (Schacter et al., 2012). Mazor (2025) suggests that similarly, model(s) of oneself can be used to predict what choices and feelings one is likely to have in future contexts that share similarities with previous contexts, can be used to infer what one is likely to have done in the past, should a memory be unreliable. Can the apprentice gardener know if she is likely to have been the one who left the garden unlocked?

Is she the kind of person who would be absent-minded and forget to do it or the anxious type who never would? In chapter I we also reviewed evidence for the crucial role of memory in imagining future episodes (Hassabis et al., 2007). McCoy et al. (2019) posit that self-models can also be updated by learning from self-simulations across scenarios that never happened or never will happen.

In the following two chapters, we will attempt to study one domain of self-models: peoples' model of their own learning process. The goal is to move beyond traditional measures (such as confidence) of local metacognition and instead require participants to roll out a full self-simulation of their actions (Mazor, Firestone, et al., 2024) during learning. Their self-model is built on memory and the actions they emulate lie in the past. We thus examine memory in action.

Unsupervised Rule Discovery

In chapter V, we introduce the setting in which we will probe participants' model of their own learning dynamics. We selected a domain in which we believe metacognition might be particularly helpful to learners (Rivers, 2021): an unsupervised learning context in which the learner is left to their own devices and needs to adjust their actions to support their own learning process. In chapter VI, we then present the same task to participants twice (with temporal distance). Crucially, in the second presentation, we ask participants to repeat their own past behaviour exactly. We thereby require them to behave as they would if they would learn about the task for the first time. In this way, we attempt to elicit their own representations of how they generate actions to learn. This also provides us a different measure of metacognitive accuracy because we can compare their re-instantiated behaviour to their true behaviour (Mazor, Firestone, et al., 2024).

More specifically, we study what people know about their own process of cognitive discovery. We previously introduced the idea of *unsupervised* learning, when feedback is absent. Humans are known to be able to learn through unsupervised (Bouchacourt et al., 2020) and semi-supervised (Bröker et al., 2024) learning and can do so both gradually and continuously (Eichenbaum, 2011b) as well as suddenly and discontinuously (G. Jones, 2003). Discontinuous learning dynamics are commonly referred to as 'insights' (Epstein et al., 1984), or moments in which the right representation for the task at hand (such as a matchstick arrangement) suddenly comes to mind and allows the reasoner to solve a problem which they were making little progress on prior to the representational shift (Chronicle et al., 2004; G. Jones, 2003; Maier, 1931). It is of practical interest to know what actions and policies might make it more likely that cognition yields discovery (Popp et al., 2024). Discoveries occur at the forefront of collective knowledge, where correct answers are unlabelled and curricula cannot be prepared. Metacognition might help learners know which of their actions helped them make a discovery. They might then be able to repeat similar actions in similar learning tasks.

In this chapter, we introduce a task environment, in which participants can take a series of actions to learn about a spatially determined rule, the discovery of which allows them to adjust their action sequence and gain more points. Participants are presented with a board of yellow cards, which they need to turn by mouse clicks, with the goal to uncover all black cards whilst revealing as few white cards as possible. Participants who, by themselves, discover that black cards are always arranged symmetrically left and right of the vertical midline, are able to complete the game with fewer white cards revealed.

LeGris et al. (2024) consider that one of the reasons why cognitive discoveries still lack theoretical and computational grounding might be that studies have traditionally been low in resolution. Experiments tend not to "generate rich real-time behavioural data that could be used to help pinpoint the origin of insights" (p.1, LeGris et al., 2024). We join researchers such as LeGris et al. (2024) and Berckley & Hattie (2023) in presenting an attempt at a high-resolution, continuous capture of cognitive discovery. This present empirical chapter should also be read as a methodological chapter, in which we explore ways in which cognitive scientists can make use of the benefits afforded to them by now common-place online data acquisition (Allen et al., 2023), while regaining and retaining some of the rich observations that were available in the days of in-person testing.

Continuous Measurements and Dynamic Reconstruction

Conducting psychological experiments online has many advantages (Allen et al., 2023), but with their adoption, scientists have lost what knowledge comes with being in the same room with their participants. It is common-place to test new game-like experimental designs on willing colleagues, family members or friends, who are asked to play the game while the experimenter is watching. This allows the experimenter to learn what part of the game is understood easily or what needs more explanation, in a process of inferring on ongoing cognitive process in another mind from mere observation. They might note their friend slowing down, unnecessarily and repeatedly hitting the enter key or frown or make spontaneous vocal expression ('Aha!'), all of which might reflect cognitive motion.

While experimenters devise a new online task, with the intention to test a specific hypothesis, they often apply their intuitive theory of mind. Will this task be interesting enough to keep participants' attention? Will they understand what they have to do? Will they think they have to do this fast, rather than with care? Theory of mind is used to design the task in a way that shapes what cognitive processes and behaviours it elicits. Tasks incentivise some behaviours and disincentive others. Sometimes, finding the right incentives and task constraints is accomplished by the experimenters asking themselves: what would I do and think when confronted with this task?

Of course, this multi-dimensional reflection of cognition is not accessible as soon as data collection is mediated via an online platform. Yet, as suggested above, there may be cognitive processes of interest that are not predictable via low-dimensional or low-resolution data, such as recorded binary decision at a low-resolution timescale (to give an extreme example to illustrate the point: asking participants at the end of a task whether or not they made a discovery, cannot provide insight into how they made the discovery.) Stephen & Dixon (2009) understand moments of insight as discontinuous phase transitions that occur as an agent follows a non-linear gradient in a high-dimensional continuous cognitive phase space. Others reiterate that cognition is always continuous (McClelland, 1979) and should be studied that way: "As any scientist (or vinyl audiophile) will tell you, if you want to understand what is going on in a continuous process, the highest- fidelity signal will come from a continuous measure of that process" (Spivey, 2023). One question is thus how researchers can retain as much dimensionality in our measurements and analyses that are commonly considered, while also reaping the benefit of expanding participation via online platforms. One easily accessible answer is the inclusion of mouse cursor movement into the standard analysis repertoire. Some researchers of course already do this. Zgonnikov et al. (2017) summarised some of analytic and visualisation options and discuss mouse movements as a 2D representation of traversing a high-dimensional decision landscape. Li et al. (2023) use an analysis of a mouse cursor trajectory to show that motions are biased towards distractor items (and more so if there are more distractor items). Wilcox et al. (2024) analyse mouse trajectories while people read. Similarly, LeGris et al. (2024) implement the mouse cursor as a spotlight illumination during a dark game, thereby forcing the mouse cursor to be the search tool that directly correlates with what participants can perceive. They then analyse regions visited, time spent or distances travelled.

A second concern is that hypotheses, including those by experimenters, are not plugged from thin air, but from the observations we make (Bear et al., 2020). What hypotheses have cognition researchers missed due to the lack of high-dimensional observations of the behaviours that we study? Hypotheses predetermine what measurements we take and analyse or what computational models we fit. While this precision is obviously desirable, it can also result in a kind of tunnel-vision that precludes alternative explanation or measurements. The hypotheses that come to mind are rarely exhaustive or imaginative enough to do justice to reality.

Hypotheses usefully and necessarily constrain the experimenter's search and measurement space, especially when the phenomena are multi-dimensional and complex. The art of (collectively) identifying the main predictors of a behaviour or the first principal component of a cognitive phenomenon, lies in doing the dimensionality reduction correctly. Experimenters need to make a choice about which dimensions of behaviour (eye-movement? decisions? verbal report? pupil dilation?) should be considered and measured in the first place. This step in the scientific process is refined by collective knowledge, but also the theory of mind of researchers, their hunches and intuitions. This applies particularly to cognitive experiences that have remained somewhat evasive of quantitative measurement, such as for instance cognitive discovery or models of the self.

So how can we, as experimenters, observe more, so that more comes to mind? We propose the method of *reconstruction* (Figure 10), by which we use the data generated by an online participant to replay the game play to the experimenter afterwards. We achieve this by displaying the fixed game elements, overlaid with the dynamically moving elements, that are guided by the exact recorded choices that the participant made (see methods section). Figure 10 A shows a static figure that captures the dynamic reconstruction visualisation. We show one round in the game in which a participant's cursor movement is traced across time (blue = earlier; red = later). Their click choices are visualised as white and black cards, that were revealed by a click at that location. We encourage the reader to watch one dynamic reconstruction by following this link: https://carlacremer.github.io/single_reconstruct_V/.

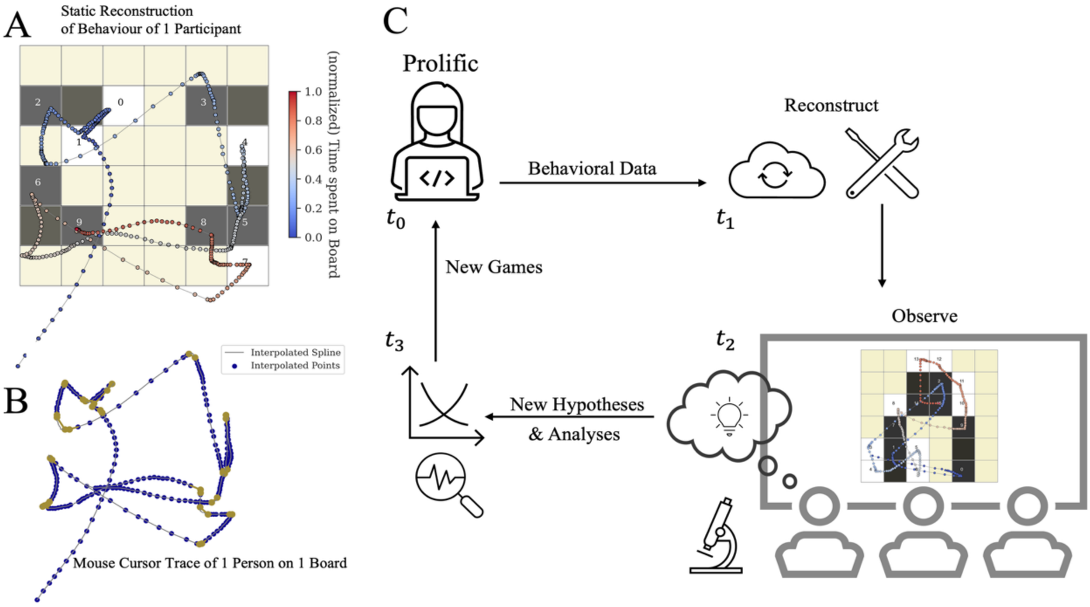

Figure 10: **Illustration of the Use of Behavioural Reconstruction** t_0: Participants partake in the online experiment outside the lab, providing behavioural markers (such as mouse clicks, reaction times and mouse movements) of their cognitive processes. t_1: Behaviour is captured using pre-specified measurements (such as cursor locations, speed, decision moments etc.). Data is saved and used to reconstruct the full behaviour of each individual participant. t_1: After data collection and reconstruction, researchers can sample individual participants and observe a reconstruction of how they played the game start to finish. Crucially, behaviour that does not pertain to the goals of the game (such as decisions or errors) is also included in the reconstruction games (e.g. mouse movement). Reconstructions allow researchers to apply their theory of mind in an attempt to infer the cognitive processes behind the behavioural patterns. This qualitative estimate becomes an input into the process of identifying quantitative analyses. Observing high-dimensional data *in action* allows researchers to use their inference abilities, to generate (new or better) hypotheses about what was going on in the minds of their participants. This allows

researchers to adjust measurements; it can inspire new analyses and adjustments to the game environment to better measure hypothesis-relevant behaviours (t_3).

To our knowledge, reconstruction is rarely deployed or reported as a method and is not, as of yet, part of an accepted and recommended repertoire of methodological steps in online studies (Gagné & Franzen, 2023). We want to suggest it should be. We have used this tool for the past few years in the lab and is has proven useful in several ways which we will list. We can also showcase the benefits here on the only other independent project (LeGris et al., 2024) that we know of that appears to have used reconstruction explicitly in a study of cognitive discovery. LeGris et al. (2024) constructed a task in which participants use 2D tools from a repertoire to direct shooting balls into target areas on the screen. The solutions which participants needed to discover lay in what physical constraints (implemented through the tools), achieved the highest number of balls falling into target zones. Solutions were designed to have sub-optimal, easier instantiations as well as more complex, optimal instantiations, thereby mimicking the classical impasse-insight scenarios, in which an elegant solution cannot be reached with most peoples' first attempt (Maier, 1931) with which they run the risk of being stuck in a local minimum. The authors set out to identify what behavioural markers preceded an insight moment. Amongst their range of quantitative approaches, the authors also report the use of "qualitative inspection of participant replay data" which "revealed a wide range of creative solutions for each problem". The authors observed participants being stuck or failing to notice a solution in spite of exploration around the right areas. The qualitative observation that participants' solutions were either static or dynamic, then informed quantitative measures of exactly those categories of the solutions space.

This example highlights how dynamic reconstructions can be useful in identifying behaviours of interest or categorising and labelling behaviour at the right level of granularity (discoveries about discovery so to speak). Note that this is particularly the case when researchers venture into unchartered territory, where it is not yet clear what behavioural marker they should be on the look-out for. Most importantly, in our experience dynamic reconstructions allow researchers to generate new hypotheses that were not yet part of their initial hypotheses set. In Figure 10 we suggest that some unexpected behaviour (*b*) that was observed in a sample of participants, can inspire new measures for said behaviour. In our experience, dynamic reconstructions can be helpful in verifying that a new quantitative measure (of e.g. a particular mouse movement pattern) is well tracked by the analysis or if the measure of choice does not fully capture the behaviour (b), by either including similar but distinct behaviours or failing to label behaviour *b* in other participants, thereby needing readjustment to achieve generalisation. Once the measure is sensitive to satisfaction, it can be scaled and statistically validated on the full sample. Note that this means that some biases that the researchers introduce into the analysis by watching and *overinterpreting* a dynamic reconstruction, will be corrected. Experimenters will for instance be prone to see signal in noise and might interpret meaning into mouse movements which in fact stemmed from a participant knocking over their coffee cup. This is not a problem, as such biases are weeded out, at least by the time that they try to re-identify the behaviour in the next pre-registered sample. A measure of a mouse movement associated with a falling coffee cup is unlikely to translate into a statistically significant association with cognitive discovery. While observations of participants might introduce more false hypotheses that can be excluded, this appears to be the lesser evil if we balance it against the inefficiency of a research community that is stuck on a narrow range of hypotheses or measures (Bröker et al., 2024). In addition, it was our experience that dynamic reconstruction also allowed us to notice non-compliant behaviours, data contamination (such as in our case tablet usage) and verifications of quantitative exclusion measures. The use cases will however differ from study to study.

Dynamic reconstruction brings the experimenter some way back into the room. While the dimensionality reduction of complex behaviour is of course still drastic (no eye tracking, limb movements or vocal expressions), dynamic reconstruction has three advantages over observing participants live: reconstructions are privacy preserving, can be made accessible to anyone who might care to watch, can be observed repeatedly and in groups, thereby allowing for collective inference.

This chapter will present a task in which participants have the opportunity to make a discovery (we will however make no claim about whether this discovery was discontinuous). We introduce this task in chapter V to discuss methods of dealing with the complexity of capturing discovery behaviour; to show quantitative measurements of behavioural markers of discovery and to provide the basis of understanding our study in chapter VI, in which we use the same task to study metacognitive access to cognitive discovery.

In the following sections we will describe the task and procedure, as well as the method of reconstruction which helped to inform a variety of quantitative measures of discovery, which make use of reaction times, number of clicks, position of clicks, error patterns and dynamic cursor movements. We introduce a sample of these measures in chapter V and extend the list in chapter VI. Our goal in chapter V is to familiarise the reader with the task we use and show that we can quantitatively measure behavioural markers of a representational shift across knowledge states in participants who discovered the rule.

Methods

Participants

The research complied with all relevant ethical regulations and was approved by the University of Oxford Central University Research Ethics Committee (code R90986/RE002). Participants were recruited via Prolific and gave informed consent prior to their participation.

To be eligible to take part in this study, participants' Prolific approval rate had to be 90% or higher and their reported first language had to be English. Participants were instructed that they could only enter the game if they were not using a tablet/phone (this ensured we could record mouse movement behaviour from all participants).

The median time to complete the experiment was 9 minutes. All participants were compensated for their time at an hourly rate of £10 and were incentivised via a bonus payment that dependent on their performance in the game (the top 30% of participants were given an additional £0.8).

Exclusions

We collected 103 participants and excluded 4 participants who used a tablet (as per our pre-registered exclusion plans), resulting in a sample of 99 participants. Participants were excluded based on self-reported tablet usage as well as an additional measurement that used the amount of distance travelled on screen. This quantitative alert of tablet use was then verified by watching a reconstruction of one round by those participants.

Task Design

Our experiment was implemented via a new online game environment (Figure 16), which participants (n=99) accessed via Prolific. In the game, they were presented with a board of 36 cards. Yellow cards hid black (hits) and white cards (misses). Participants' goal was to 'turn' yellow cards via a mouse click to uncover all black cards, whilst revealing as few white cards as possible. Participants were told that 10 black cards were arranged in groups of size 2, 3, or 4, that group corners can touch and that a single black card per group would be revealed at the start of each round (Figure 16 Top Left). Every white card revealed would deduct 1 point from their initial score.

There was one additional rule which participants were *not* told about: black cards were always arranged symmetrically left and right of the vertical midline (Figure 11 Bottom Middle).

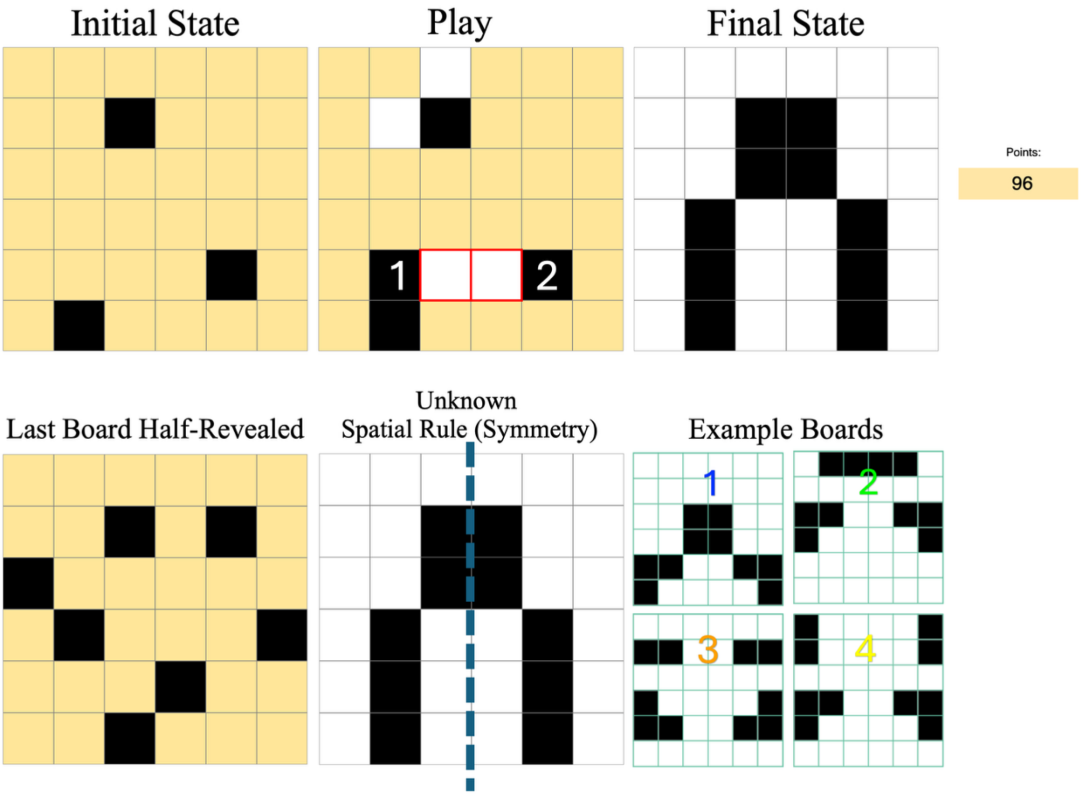

Figure 11: **Illustration Design of Discovery Task Chapter V**: **Top Left**: A board of 36 cards as seen by a participant at the start of a new round. Yellow cards indicate an unrevealed state. One from each group of black cards is already revealed. **Top Middle**: The same board during play. Participants click on yellow cards to find hidden black cards, while revealing as few white cards as possible. 4 white cards were revealed, each of which leads to a subtraction of 1 point from the initial (100) points shown on the right. The unknown spatial rule of the game was that all black cards have a mirror symmetric twin. We illustrate this here, with the number 1 and 2 placed on members of one symmetric pair (not shown to participants). A symmetric twin of a white card can never be a black card. A participant who has discovered this symmetry rule can avoid revealing one of the two white cards

highlighted in red here. **Top Right**: The board in its final state, showing the true arrangement of black cards. There are 10 black cards on every board. Upon revealing the last black card, all remaining yellow cards will momentarily fade, to highlight the symmetric arrangement to the viewer. **Bottom Left**: Half-revealed board, played last. This board contains 14 black cards, 7 of which are already revealed. Unbeknownst to the participant, one of each mirror-symmetric twin is already revealed, which means that a player who knows the symmetry rule does not need to incur the cost of touching any white cards. **Bottom Middle**: Illustration of the solution to the game. The black cards are deterministically arranged to be mirror-symmetric across the vertical midline. **Bottom Right**: 4 example boards sampled from the 6 unique boards played by each participant. Each board is associated with a coloured number which indicates the game round. The assignment of game round and board is randomised across participants.

Participants first received instructions, answered comprehension tests and had one practise round[3]. They then begun the game by playing 6 unique boards (see Figure 11 Bottom Right for examples of board arrangements). Each participant played each board once. The order in which they were played was randomised across participants. In between boards played, an image of a coloured number was shown to mark the identity of the upcoming round (Figure 16 Bottom Right). After the first six rounds, participants were then exposed to a final board, which differed from the others in that half of the (14) black cards had already been revealed (Figure 16 Bottom Left). This was known to participants. It was however not

---

[3] The practise board was just to get participants used to the user interface and they were not arranged like the real board to avoid participants learning about the board arrangement beforehand. In the practise board they revealed a two-rowed horizontal bar of black cards (which were thus symmetric along the x and y axis).

revealed to them that one of each mirror-symmetric twin had been revealed. We constructed this board to distinguish participants who had discovered the symmetric rule arrangement, since in this final round, a person who made choices according to the rule would not need to incur any errors. The locations of the missing black cards were fully predictable from the initial arrangement of the board (Figure 16 Bottom Left). Participants who had maximally a single error (only revealed 1 white card) on this last board, were designated as participants who had discovered the rule (Cremer & Mazor, 2025a).

After the game, participants were asked if and when they noticed a pattern of how the black cards were arranged.

In our game, participants could commit three types of errors: 1) errors which were illegal according to the pre-specified rules of the game (e.g. uncovering a white card adjacent to an already revealed island of size 4); 2) errors which were avoidable (Figure 16 Top Middle in Red) if they knew the symmetry rule (i.e. clicking on a covered card that is the symmetric mirror of an already revealed white card) and 3) unavoidable errors when no half of a twin remains covered and it is uncertain where the next first revealed of a new twin is located. These errors do not exist on that last boards in which half of each twin is already revealed. Participants were never told about the existence of a secret rule and were never given feedback (other than the subtraction of points upon a white card being revealed) about how their performance could be improved. They were however encouraged to improve in between rounds: their total games score and the error count of the last round was shown and a message read “Can you do even better?”, before the next round commended.

Dynamic Reconstructions

We briefly explain the reconstruction method we used in step t_1 shown in Figure 10. We recorded the position of participants’ mouse cursor with a high frequency (ca. every 16ms). This data was combined with records of the board played, click locations and reaction times, to reinstate the behaviour of participants for each board played. This was achieved by programming a replica of the game in JavaScript. The recorded board and point score determine the fixed, contextual elements of the game, whereas click locations, click times and mouse cursor coordinates dictate dynamic elements that represent the participant's choices.

One such reconstruction (of a pilot participant playing the last board) can be viewed under this link: https://carlacremer.github.io/single_reconstruct_V/. We will provide more reconstructions from another sample of participants in the following chapter (VI).

In our reconstructions, a red dot dynamically moves across the board, tracing the trajectory of the participant's recorded mouse cursor, matched to the time point at which each coordinate was passed during the game. A timer shows the milliseconds passed since the time point at which data recording commenced. At the time point at which the participant made a click, the card that was revealed by the participant at the time, is revealed in the reconstructed game and the score board is updated.

In this way, we were able to observe the game play of all participants across all rounds they played. We used repeated and random observations of individuals' (reconstructed) behaviour to generate hypotheses about what behaviour might be correlated with rule discovery.

We noticed that participants make a lot of mouse movements that are not obviously related to the achieving the goal of the task. One perspective is that these are random and irrelevant or unrelated behaviours. Another perspective would be to consider these actions as reflections of ongoing cognitive processes and therefore of interest. Some of the behaviours we observed included left-right movements with the mouse cursor in between clicks, akin to a compare and contrast behaviour, potentially to ease the difficulty of error-prone identification of the symmetric twin within the same row. This behaviour was reminiscent of what Kirsh & Maglio (1994) called an *epistemic action*, which is an action with the purpose to ease a cognitive effort (as opposed to affect the environment). Other behaviours included differences in reaction times that were associated with uncertainty differences defined by the symmetry rule, differences in hesitation before committing to a click or, when uncertainty was high or repeatedly moving to the same location, but not committing to revealing it.

We discussed a range of tentative hypotheses, prioritised between them and then generated quantitative measures using data from specific individuals, in which we had observed the behaviour and on which we could verify that our behaviour was indeed captured by our quantitative measure. We then scaled the same measure to all participants.

In this chapter we will showcase two measures of discovery as reflected in performance. This includes a straightforward measure of the number of errors per round and a non-parametric assessment of decisions by two models, one of which assumes a symmetric arrangement of the board and one which does not. We describe the details of the model evaluations in the methods sections below. This chapter then proceeds to demonstrate two additional behavioural markers of discovery which are also explained in detail in the methods sections: the difference in reaction time for revealing the first versus the second member of a symmetric pair and the tendency to move the mouse cursor left-right versus up and down.

Pre-registered hypotheses

We pre-registered our hypotheses (Cremer & Mazor, 2025a) prior to collecting a new, confirmatory sample. Our hypotheses (listed also in Appendix B) were centred around the reflection of spatial rule discovery in a range of behavioural metrics and informed by the behavioural patterns we observed in reconstructed games of previous pilot participants who had perfect performance on the last, half-revealed board.

We tested the hypothesis that knowledge about the rule is reflected in participant’s number of clicks (a reduction of errors over rounds) and reaction times. We posited that participants who follow the rule will take longer to uncover the first of a symmetric pair of black cards but be faster to uncover the second of the pair. This we interpreted as reflecting the certainty over the second location which can only be had if the symmetric rule is known. We pre-registered to test if this measure correlated with performance on the half-revealed board by regressing the difference in RT against the number of clicks on the last board.

We set out to test the hypothesis that we can identify rule discovery in the choices and timings of choices of participants, by comparing two descriptions of these choices by two different models which operate under different constraints. We predicted that rule-following participants would be best described by a rule-following model in which choices are constrained by symmetry.

We test the hypothesis that rule-followers show more horizontal mouse cursor movement relative to vertical movement over time, reflecting the likelihood to click on the mirror-symmetric twin of a newly discovered black card as well as a search and compare pattern that identifies symmetric locations along the x-axis rather than the y-axis. The measures and analyses we use to test these hypotheses are described in a following section on behavioural markers.

Behavioural Markers of Discovery

We will report results from a sample of behavioural measures, some of which are traditional measures such as error rates, and some were inspired by observing reconstructions and hypothesising about cognitive markers in mouse cursor movements. The measures detailed here are also used in chapter VI.

Knowledge-State–Dependent Constraints

We construct two models that make choices under different constraints.

The *naive model* was constructed to consider the rules of the game but be ignorant of the spatial symmetry rule. We enumerated all board arrangements which adhere to the following constraints: 10 black cards, x = {2,3} black cards locations already revealed, groups of black cards of size ={2,3,4} and corners of groups may touch. We did this for every unique board (in their initial, partially revealed state) that participants played (see Figure 11 Top Left for an example). Boards were designed as 6x6 Boolean arrays, with 1 representing black and 0 representing white cards. These sets of boards are legal boards according to the rules of the game only. Black cards were permitted to be asymmetrically arranged left and right of the midline.

The *symmetric model* was constructed to consider the rules of the game *and* the spatial symmetry rule. In the process of enumerating all legal board arrangements for the symmetric model, we thus added the constraint that five black cards on the left side must be the mirror-symmetric twins of five black cards on the right side of the vertical midline. We then implemented a procedure for evaluating each click, made by each person, through the lens of each of the two models.

Essentially, the models assume players starts the game with a belief state that assigns equal probability to all possible board configurations: symmetric board states for the symmetric model, or symmetric and asymmetric for the naive model.

From this belief state we then derive probabilities that specific cards are black or white, by taking the average colour across all board configurations that are consistent with the belief state.

The belief state is then updated after each state of the game, ruling out game configurations that are inconsistent with the player’s newly acquired knowledge.

We rank the data of participants as follows. For every click (c) by participant (s) on board (b), we took the most recently chosen location (i, j) (where (i and j = {0, 1, 2, 3, 4, 5})) of an unrevealed card. We obtained a ranking of the choice i, j by each model. This ranking indicates how likely it was (given the constraints by each model) that the location (i, j) contained a black card. For each model (m = {naive, symmetric}), we took the set of all legal boards and summed across all boards per unique, unrevealed location. We then normalised by the number of boards that are consistent with the player's simulated knowledge state to obtain a location-specific probability of a hit for each unrevealed card. After computing the complement of each probability (1 - p), unrevealed cards were ranked. We ensured that all positions with a tied rank returned their average rank. We then extracted the rank of the specific, chosen location (i, j).

Finally, we updated the known board state from the configuration before click c_{ij} to the configuration after c_{ij}, assigning a value of 1 (hit) or 0 (miss) to location (i, j) based on the true underlying board configuration. In a last step, we excluded all boards from each of the two sets of legal boards (as defined by each model) that were inconsistent with this new piece of information (e.g. we excluded all boards that included a 1 in location (i, j) if the ground truth board was just revealed to contain a 0 in location (i, j).

Reaction Time in Symmetric Pairs

Here we analyse the reaction time (RT) between each consecutive click (e.g. $t_1 - t_0$). Participants self-selected when they made their clicks. We analyse how much time passed in between clicks made.

We measure the tendency to uncover the second of a pair of symmetric cards with greater speed than the first. The expectation was that this reflects increased confidence of rule-followers who have deterministic knowledge about the location of the second of a half-revealed pair after they discovered the rule.

First, we selected the second to largest number of clicks (second to longest board played) for each participant and computed the median RT per click (up until the second to largest click number) across all boards (except the last). We used this measure to control for differences in reaction time that depend on click number in the round.

We again excluded the half-revealed board from this analysis. All clicks on this board were second clicks. For each participant (s) and each click (c) we computed a normalised RT measure: the participant's RT for (c) minus the participant's median RT at this click number (e.g. the 9th click made in each round). Note values can be negative because we normalise by median RT on each click number, meaning negative values show faster responses than the median RT.

$$\mathrm{RT}_{\mathrm{norm}}^{(s,c)} = RT^{(s,c)} - \widetilde{RT}^{(s,\cdot)} \qquad (1)$$

We labelled each click by whether it was revealed as the first or second member of its symmetric pair. We computed the median of the normalised RT measure across all clicks (in the class of either 1st or 2nds) per participant and board.

To correlate this measure with performance on the half-revealed board, we took the per-person average difference between the median RT across all 1sts and across all 2nds in the last three boards. For each subject (s) and each board (b), we compute:

$$\Delta^{(s,b)}RT = \bar{RT}_F^{(s,b)} - \bar{RT}_S^{(s,b)} \qquad (2)$$

We then calculate the person-specific average of the RT differences across the last three boards (Equation 3).

$$\bar{\Delta}^{(s)} = \frac{1}{3}\sum_{b=4}^{6} \Delta^{(s,b)RT} \qquad (3)$$

We regress this value against each participant's number of clicks on the half-revealed board.

Motion Direction

The mouse cursor's x and y coordinates across the screen were recorded approximately every 16 milliseconds. After dividing each board's list of coordinates into window slices (of 8 measures, resulting in ca. 128 milliseconds per window), we subtracted the sum of the absolute distance travelled along the x-axis from the sum of the absolute distance travelled along the y-axis (for each window). This is illustrated in Figure 15 (Left).

After dividing each board's coordinate log into 8 equal time windows, we computed the total absolute distance travelled along the x-axis and y-axis within each window. We then subtracted the total y-axis distance from the total x-axis distance to obtain a motion direction score per window. Positive values indicate more horizontal (left-right) movement and negative values indicate more vertical (up-down) movement. Finally, we averaged motion direction scores across all windows for each participant and board, to show a per-board, per-person measure of motion direction as a function of game round in Figure 15.

We fit a linear regression to motion direction against the game round at the participant-level. Coefficients from each participant were regressed against their number of clicks in the 7th round.

Statistical Methods & Python Packages

Data analyses were conducted using *Python* 3.10.13 and the following packages: *NumPy* 1.26.4, *SciPy* 1.13.1, *Pandas* 2.2.2, *Matplotlib* 3.9.2, *Seaborn* 0.13.2 and *Statsmodels* 0.14.2. This chapter was written using Quarto (Allaire et al., 2025). Experiments were coded in JavaScript (using jsPsych (de Leeuw et al., 2015) and p5.js (McCarthy et al., 2024)), and hosted on a Jatos server (Lange et al., 2015).

We computed linear mixed effects models to predict measures based on time spent in the game (round r) and modelled each participant as a random effect using individual intercepts, using the Python *statsmodels* package.

$$Y_{ij} = \beta_0 + \beta_1 \mathrm{r}_{ij} + b_{0j}$$

We use the package SciPy. Stats for t-tests and model rank. We perform Welch's t-test for unpaired samples in accordance with (Zimmerman, 2004) and never assume equal variances.

As pre-registered, we did not include the last, half-revealed or the first practise board in our statistical analyses. This board was used to identify rule discovery (in performance-based manner) and previous rounds were used to identify behavioural patterns that are associated with rule-discovery.

Results

All results from this pre-registered experiment were consistent with results from a previous pilot dataset (n=101), which was used to inform the pre-registration. We start by showing a common measure of performance. The task is designed such that participants who discover the symmetry rule are able to complete each round faster and with fewer errors to retain more points.

Number of Errors

We thus first report performance by showing the number of clicks to completion per round. As noted in our pre-registration, we designated all participants who had no more than 1 error on the half-revealed board (Figure 16 Bottom Left) as participants who are most likely to have discovered the symmetry rule. This resulted in (as expected based on the pilot sample)4 37 *rule-followers* and 62 participants who did not meet our criterion and who we here simply designate as '*other'* participants (as we have no evidence according to our criterion to suggest that they learned the symmetry rule).

---

4 The pre-registered expectation was based on a pilot dataset in which 40/101 participants turned out to be rule-followers according to the same selection method.

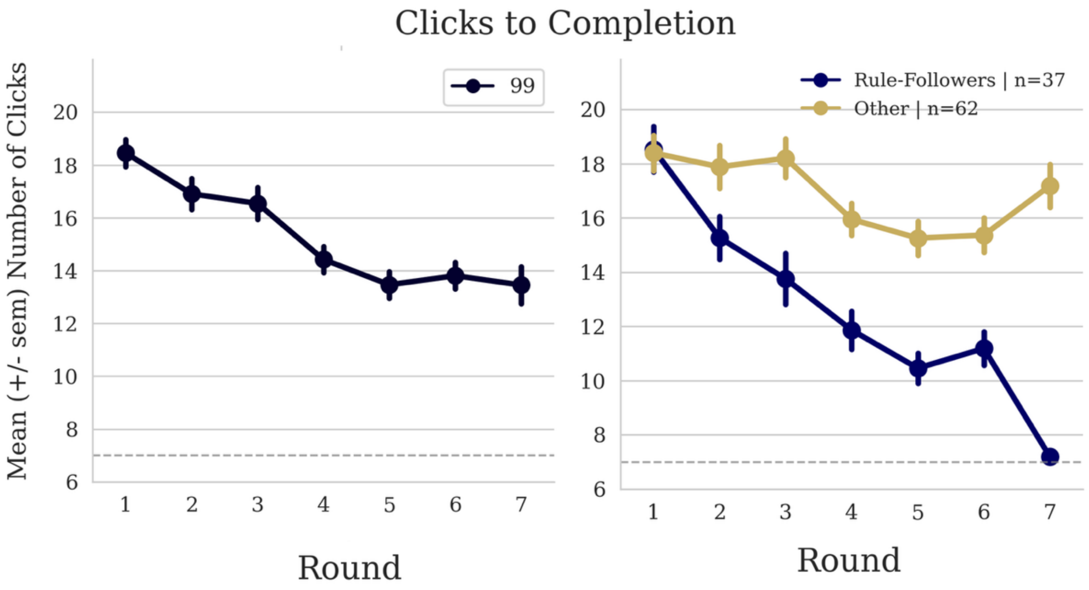

Figure 12: **Clicks to Completion Chapter V Left**: Average number of clicks (and sem) per round across all participants. **Right:** Average number of clicks per round for each subgroup of participants: rule-followers (blue) versus all other participants (gold)

The half-revealed board is used only as a selection criterion to form the subgroup of rule-followers. All the following analyses in which we compare behavioural markers of discovery are based on data from the first 6 boards, as pre-registered.

Participants improved with the number of rounds played. They required fewer clicks (thereby revealing fewer white cards and retaining more points) over time. In Figure 12 (Left) we plot the mean and standard error of mean of clicks to completion across all (n=99) participants. Note that the minimal number of clicks required to complete each board was 6 or 7 depending on the board played (the order of which was counterbalanced across participants).

To capture this improvement quantitatively, we fit a mixed effects model to the number of clicks to completion, where round number was set as a predictor (and each participant was modelled as a random effect). The number of revealed white cards declined significantly: $\beta =$

-0.161, 95% CI [-0.32, -0.01], $p = 0.041$, $df = 790$ as also shown in Figure 12 (Left). A repeated-measures t-test shows a significant drop in errors between round 3 and 4 (t(df = 98) = 2.99, $p = 0.003$).[5]

In Figure 12 (Right) we plot rule-followers and others separately, which highlights that rule-followers make markedly fewer errors than the rest of the population. A Welch test on each participant's average number of clicks across rounds (1 to 6), shows that a separation of the population by errors on the 7th board results in a significant difference in performance between rule-followers and other participants: t(df = 97) = -5.64, p = < .001.

We then applied a separate linear regression to all participant's number of clicks per game round (1-6) and extracted the slope parameter. These coefficients were then regressed against the participant's number of clicks in the 7th round (the half-revealed board). This way we confirmed that the 7th board is an indicator for how well participants played the game. Performance on the half-revealed board predicted a steeper learning rate: $\beta = 0.0348$, 95% CI [0.0024, 0.0671], $p = 0.035$, $R^2 = 0.045$, df = 97. The average slope of participants who applied the rule in the half-revealed board was -1.52, whereas the learning rate of all remaining participants was -0.72.

Note that the subset of participants who had more errors than our designated class of rule-followers may still have noticed the symmetric arrangement. 29 of 37 rule-followers described the pattern of black card arrangements as 'symmetric' or 'mirrored' in response to a debrief question after the game. As did 9 out of 62 participants who landed in our group 'other' and who did not meet our pre-registered performance criterion for board 7.

---

[5] This is also the case if we correct for (bonferroni) multiple comparisons across 7 rounds.

Participants with few errors on the last board may have been lucky, and participants who did not meet our low-error threshold could still name the rule, which suggests that the number of clicks on the last board is best considered to be a reflection of the probability that a participant discovered the rule, with fewer errors indicating a higher probability of rule discovery.

Participants who were not able to notice or verbalise the rule, put forward a range of alternative hypotheses, ranging from letter shapes (L or Ts), to “a skull” and “zigzag” or “the helmet of Samus Aran from Super Metroid”. Minimal clicks on the half-revealed board are indicative of having had the right hypothesis about the spatial pattern and of having accepted this hypothesis as true. All responses to the debrief question are listed in Appendix B.

Model-Based Decision Evaluation

We next assess whether the identity and timepoint of participants’ card selections were best described by a selection mechanism that takes the symmetric arrangement of black cards into account, or whether they were better described by a selection mechanism that only considers the rules of the game, but not the symmetric arrangement. For this, we employed a two-pronged modelling approach. Both models are described in detail in the Methods section.

Model fit was assessed by first, generating a set of all possible board configuration for each unique board in its initial, partially revealed configuration. The number of possible board configuration (i.e. the boards that could be true, given the available information) were constrained by the rules that each model took into account ($LegalBoards_m$). The naive model only took the rules of the game (10 black cards, group size $=\{2, 3, 4\}$, group corners may touch) into account, whereas the symmetric model additionally also took the symmetry rule into account. The set of possible boards according to the symmetric model were all symmetric along the vertical midline.

For each click (c), per person and board, we computed the probability that the chosen location (i, j) contained a black card according to each model m = {naive, symmetric}. We did this by counting how many boards, out of the set of $LegalBoards_m$ at the time of (c), contained a black card on the location (i, j). The complement of this probability (1 - p) was then converted into a rank, which we used to retrieve the favourability of the choice (i, j) according to m. Was the choice more likely made with or without the consideration of symmetry?

After the rank was retrieved, we updated the visible board with the newly obtained information of the card colour in location (i, j) and excluded all boards whose entry in (i, j) conflicted with the newly obtained information (e.g. boards which had a black card on (i, j) even though (i, j) was just revealed to be white were now no longer legal).

The symmetric rule constraint meant that the initial set of possible arrangements for the symmetric model was substantially smaller than that of the naive model and that newly obtained information would more drastically narrow down the number of possible boards in the symmetric model and do so more quickly than in the asymmetric model. This is a reflection of that rule discovery, which leads to a more complete understanding of the constraints by which the board arrangements are generated, should permit more efficient inference. The options space is smaller if a consistent pattern is taken into account during search.

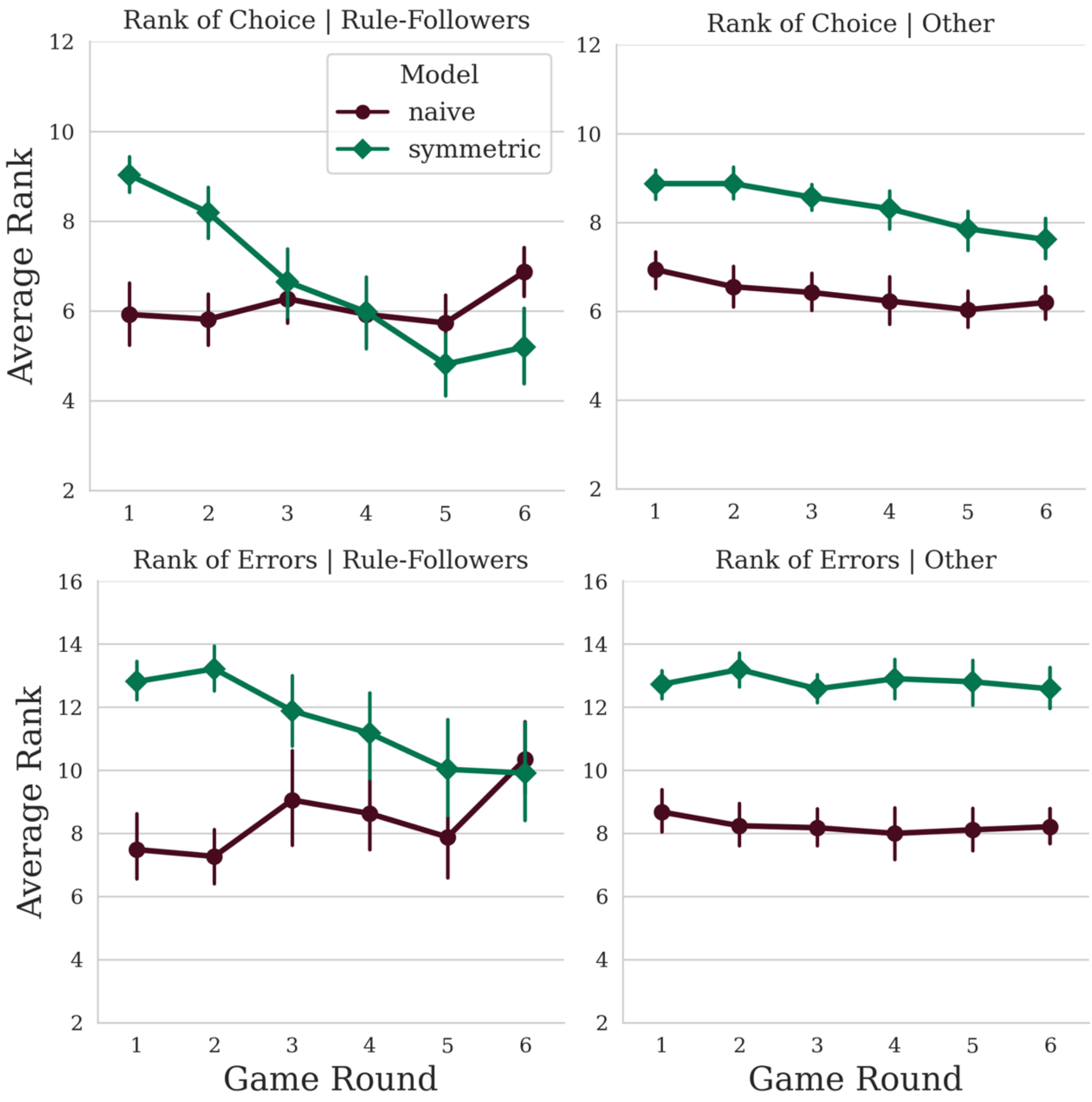

Figure 13: **Model Ranks Chapter V Top Left**: Choices (errors and hits) of Rule-Followers ranked by the symmetric (green) model and the naive (red) model respectively. The average rank of the symmetric model declines and provides a better description of choices made by participants over rounds. **Bottom Left**: Errors of Rule-Followers ranked by the symmetric (green) model and the naive (red) model respectively. The average rank of the symmetric model declines and shows that errors are more in line with the symmetric model over time. **Top Right**: Choices (errors and hits) of Others ranked by the symmetric

model and the naive model. **Bottom Right**: Errors of Others ranked by the symmetric model and the naive model.

In Figure 13 we plot model rank as a function of game round, for choices that yielded both white and black cards (Top). We show model dynamics for each subgroup (rule-followers and other participants) but also analyse how model fit varies across the entire population. First, Figure 13 (Top), shows the mean and standard error of rankings by each model for each group. It reveals that choices by rule-followers (Figure 13 Top Left) have an average fit that is distinct from participants who likely did not discover the rule. While both groups are better described by the naive model at first (dark red lines in Figure 13), rule-followers make choices that are ranked better by the symmetric model (dark green) after round 4. The rank the symmetric model surpasses the rank of the naive model *and* the naive model rank increases in round 5 and 6, showing that choices made by rule-followers are favoured *less* by a naive model.

A mixed effects model, fit to all rule-followers, in which we predicted the rank by the symmetric model by game round (and modelled each participant as a random effect), finds a significant improvement (decrease) in rank over time (β = -0.839, 95% CI [-0.99, -0.69], p = < .001, df = 220).

A mixed effects model, fit to all rule-followers, in which we predicted the rank by the naive model by game round (and modelled each participant as a random effect), finds an increase (but not a significant one) in rank over time (β = 0.107, 95% CI [-0.02, 0.23], p = 0.095, df = 220).

This trend is not apparent when we fit these models to participants who did not follow the rule (Figure 13 Top Right). The average symmetric model rank remains above the average naive model rank even in later rounds.

A mixed effects model fit to all participants' choices who did not follow the rule, in which we predicted the rank by the symmetric model per game round (and modelled each participant as a random effect), still finds a significant decline in rank with a smaller effect size: $\beta = -0.255$, 95% CI [-0.34, -0.17], $p = < .001$, $df = 370$.

In our design, performance (the number of clicks to completion) and rule-discovery are correlated: participants who know the symmetry rule can reveal black cards with fewer attempts and point losses. A favourable ranking of clicks (and their timepoint) in which rule-followers revealed black cards is thus partially a reflection of their ability to uncover black cards with fewer tries, since the symmetric model narrows down the options for black card locations more quickly. We already showed in Figure 12 (Top) that rule-followers find black cards with fewer tries and thus the favourable ranking by the symmetric model of clicks in which black *and* white cards were revealed is perhaps not surprising. We therefore turned to assess model-fit using only errors. The models allow us to ask whether the errors of our subgroup of rule-followers also became more rule-like over time.

The contrast between subgroups is indeed stronger when we restrict our analysis to errors (see Figure 13 Bottom Row). We now fit the models to white clicks by participants who were later classified as rule-followers. A mixed effects model was fit to the rank of the symmetric model and confirmed that their errors also become more consistent with knowledge of the symmetry rule: $\beta = -0.708$, 95% CI [-0.97, -0.44], $p = < .001$, $df = 210$. Rule-followers later tended to make unavoidable errors that reflected the true remaining uncertainty of the board, even if symmetry was considered.

This also leads to a significant *increase* in the rank by the naive model: $\beta = 0.441$, 95% CI [0.16, 0.72], $p = 0.002$, $df = 210$, reflecting that errors became less likely under model constraints that included asymmetric boards.

When we fit the same mixed effects model structure to the rank by the symmetric model, also now using *errors* of other participants (Figure 13 Bottom Right), we no longer saw an effect: $\beta = -0.053$, 95% CI [-0.19, 0.08], $p = 0.450$, $df = 369$.

We additionally wanted to confirm the relationship between model fit and the likelihood of rule discovery without subsetting. For this, we computed an average difference in model rank (Naive - Symmetric Model) at the participant-level, across the last three boards. Positive values indicated that symmetric model described the data better than the naive model. We then regressed this measure against the number of clicks on the half-revealed board at the population level.

A linear regression shows a significant negative trend ($\beta = -0.1394$, SE = 0.0305, $t = -4.57$, 95% CI [-0.2, -0.0788], $p = <.001$) confirming that the symmetric model is more likely to overtake the naive model for participants who completed the half-revealed board with fewer clicks. This trend persisted when we ran the same analysis on errors only: $\beta = -0.1517$, SE = 0.0561, $t = -2.71$, 95% CI [-0.263, -0.0405], $p = 0.008$.

Rule-Discovery in Reaction Times

Every black card had a mirror symmetric twin on the other side of the vertical midline. This is illustrated in Figure 16 (Top Middle), where the numbers illustrate the first (1) and second (2) of a single pair. A participant who discovered the rule has the advantage that they only need to uncover a single black card (e.g. 1 in Figure 16 Top Middle) under uncertainty for each of the pairs. The location of the second of the pair is always certain. Uncertainty and confidence are often thought to be reflected (and measurable) in participants' reaction times (RT) (Kiani et al., 2014). This suggests that rule-followers will show faster responses in revealing the mirror-symmetric card of an already revealed black card.

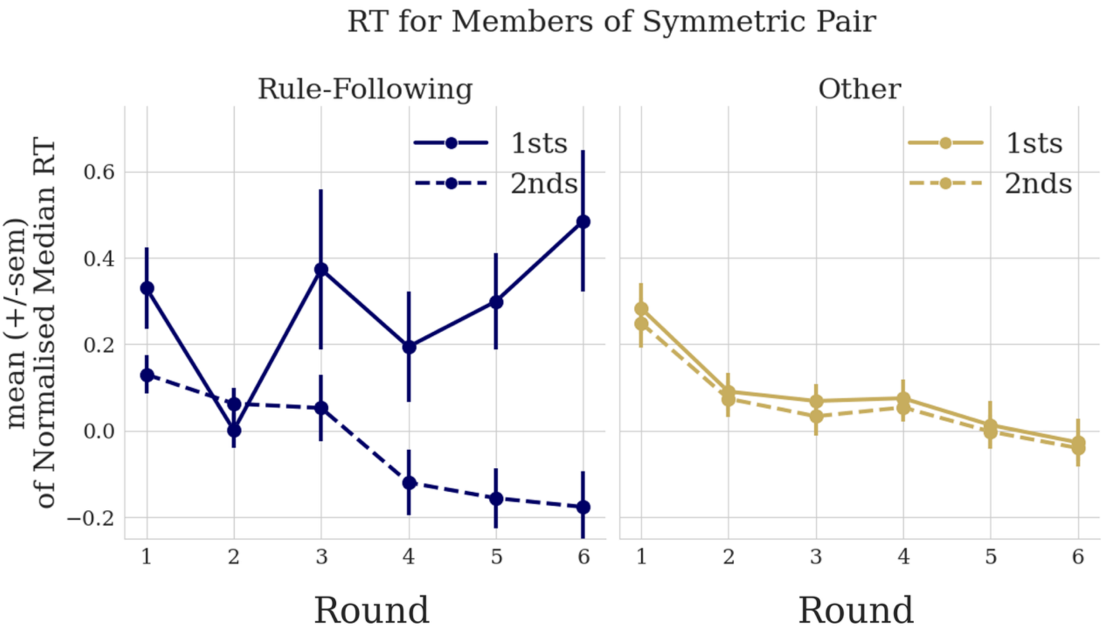

Figure 14: **Reaction Time Divergence Chapter V A**: Rule-Followers' (n=37) diverging reaction times (mean +/-sem) for black cards revealed as 1st (purple) of the symmetric pair vs 2nd (blue) of the symmetric pair. **B**: Others' (n=62) non-diverging reaction times (mean +/-sem) for black cards revealed as 1st (purple) of the symmetric pair vs 2nd (gold) of the symmetric pair.

We tested this hypothesis by measuring each participant's median RT across all clicks which revealed the 1st of a pair, while controlling for their median RT at each click number across the game (see Methods). We also do this for all clicks which revealed the 2nd of a pair. In Figure 14 we plot the mean and standard error of this RT measure for rule-followers (Left) and other participants (Right), separated by all 1st and 2nd (dashed line) clicked in each pair.

We assess the dynamics of Firsts in each subgroup {rule-followers, others}. To do this, we applied a linear mixed effects model to the median RT of all Firsts to compare them across subgroups (comparing the two full lines in Figure 14 across panel Left and Right). We add round and subgroup as a predictor. This revealed a significant interaction between subgroup and round, showing that rule-followers show a significant increase in Firsts (see increase in full line in Figure 14 Left) relative to the Firsts of others (β = 0.0934, SE = 0.0285, z = 3.28, 95% CI [0.0375, 0.1492], p = 0.001). We also find a negative main effect of round (β = -0.0508, SE = 0.0174, z = -2.92, 95% CI [-0.085, -0.0167], p = 0.004), owed to the significant decline in RT of Firsts in the subgroup Others (Figure 14 Right), and no main effect of subgroup (β = -0.1293, SE = 0.1113, z = -1.16, 95% CI [-0.3474, 0.0887], p = 0.245). We then apply the same analysis to Seconds, comparing the two subgroups and find no effect of subgroup (β = -0.0252, SE = 0.07, z = -0.36, 95% CI [-0.1623, 0.1119], p = 0.718). and no significant interaction (β = -0.0202, SE = 0.0179, z = -1.13, 95% CI [-0.0554, 0.015], p = 0.261) but a main, negative effect of round (β = -0.0473, SE = 0.011, z = -4.31, 95% CI [-0.0688, -0.0258], p = <.001), showing that both rule-followers and others show an increase in speed for seconds as they play the game for longer. Only for rule-followers however does this increase in speed diverge from a slowing down when revealing the first of a pair.

We also measured RT divergence by computing each participant's average difference in median RT of 1sts minus 2nds for the last three boards (see Methods - RT Difference).

This indicator of RT divergence was regressed against the participant's number of clicks on the last board and verified a significant association between RT divergence and near optimal behaviour on the last board: β = -0.0187, SE = 0.0072, t = -2.58, 95% CI [-0.033, -0.0043], p = 0.011.

The difference in time taken to make a choice between firsts versus seconds of a symmetric pair is predicted by rule-following. The reduction in uncertainty that is gained by discovering the rule was potentially what was reflected in faster responses for certain locations.

Mouse Movement Direction

Rule-followers can theoretically be more certain about the location of a second member of a pair after revealing the first. If rule-followers tend to ‘cash-in’ the new, certain black card after discovering the first of a new pair, we would expect more horizontal mouse movement after rule-discovery. A second reason to expect more horizontal movement, is that the search pattern of the mouse cursor might move to guide or reflect movement of their eyes and attention, as participants compare the position of a revealed card on the left to the same mirrored position on the right (or vice versa).

We computed the motion direction of each participant’s mouse cursor to test this (Figure 15 Left for illustration). We subtracted the per-participant sum of the absolute distance (right and left, in both directions) travelled along the x-axis from the sum of the distance travelled along the y-axis, per round.

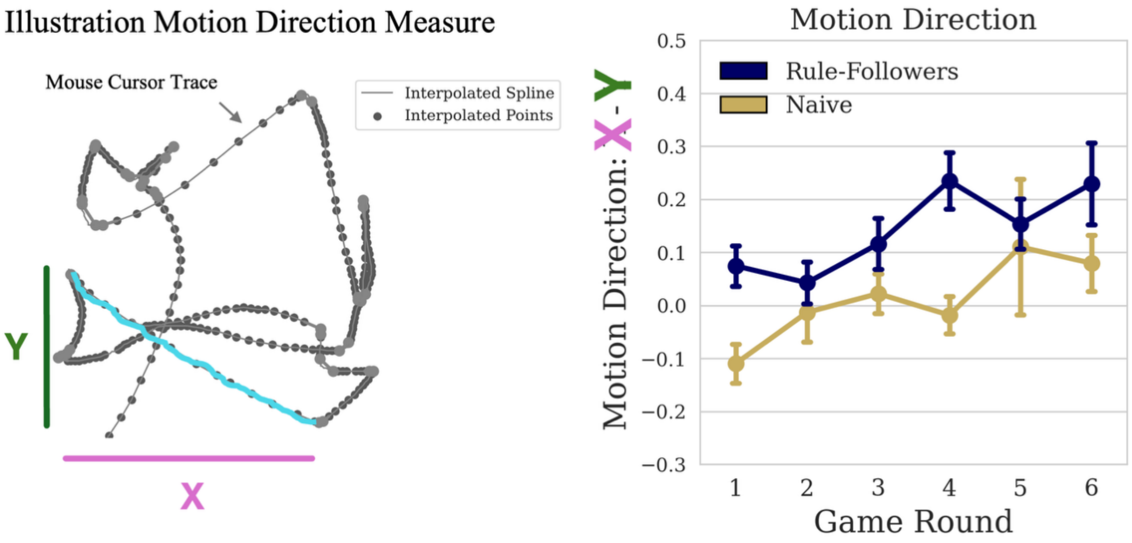

Figure 15: **Motion Direction Left:** Illustration of motion direction measurement. Grey connected dots show the trace of the recorded mouse cursor positions across one round by one participant. Summed absolute movement along the horizontal plane (X) is computed separately from the total absolute distance travelled along the vertical plane (Y)

Right: Mouse Movement Direction: for each subgroup we show means (+/- standard error of the mean) of horizontal - vertical (X-Y) movement per game round. Higher values on the Y axis of this figure indicate more horizontal movement than vertical movement.

In Figure 15 right we show the population mean (+/- sem) motion direction (absolute movement along x-axis minus absolute movement along y-axis) across participants, divided into two groups: those who likely discovered the rule and those who did not. A higher value on the y-axis indicates that participants showed more horizontal movement than vertical movement. A mixed effects model in which we predict the rule-followers' motion value by the numerical predictor of game round (and participant ID as random effect), finds the slope to be significantly positive: $\beta = 0.04$, SE = 0.0107, $z = 3.75$, 95% CI [0.0191, 0.0609], $p = <.001$. This is also still the case for the other subgroup, with a smaller coefficient: $\beta = 0.0317$, SE = 0.0101, $z = 3.14$, 95% CI [0.0119, 0.0515], $p = 0.002$. A Welch test between the two groups, using individuals' average motion energy across all rounds, however, testifies that rule-followers made overall significantly more horizontal cursor motions than naive participants: $t = 3.807$, $p = < .001$. We also regress the motion direction against the board number (per person) and retrieve a coefficient for each participant. A linear regression of these coefficients against the person-specific number of clicks on the last board showed a negative, but not statistically significant trend $\beta = -0.0017$, SE = 0.002, $t = -0.84$, 95% CI [-0.0056, 0.0023], $p = 0.402$.

Discussion

We presented a task in which participants had the opportunity to discover a spatial rule. We showcased analyses that were in parts inspired and validated by dynamic reconstructions of participants' game play and will extend this list of measures of discovery behaviour in chapter VI.

About 40% of participants made the discovery and as a result, became significantly better at the task. We introduced a last, half-revealed board on which the optimal number of clicks was indicative of having had the right hypothesis about the spatial pattern and of participants using this rule to guide their action choices. We showed measures both within pre-registered subsets of the sample and in relation to a continuous measure of performance across the full sample.

A model that took the symmetric constraint into account provided a better rank of rule-followers choices (hits and misses) and mere errors, showing that the error distribution shifted away from choices that were inconsistent with the symmetry rule. This was not the case for the remainder of the population (other). Reaction time analyses showed that rule-discovery was reflected in the certainty which rule-followers gained after they discovered one of an unrevealed symmetric pair. This was not the case for participants in the other group.

Note that suboptimal performance on the half-revealed board does not show that participants did not discover the rule. Participants who made multiple errors on the half-revealed board, may have had the right hypothesis in mind, but might not have accepted it or might not have been able to apply it.

Indeed, a decline in rank of the symmetric model when considering only naive participants indicates that some participants in our naive subgroup probably did notice the symmetric pattern and incorporated some of this knowledge into their choices, despite making too many errors to meet our pre-registered exclusion criterion for rule discovery. Nearly all participants who ended up in our pre-registered subselection of rule-followers could name the symmetry rule explicitly, but a handful of participants in the remaining group also named it correctly. Our pre-registered selection criterion for rule-discovery misses a few participants whose performance was uncorrelated with their discovery of the rule.

This interpretation is consistent with our finding that participants who were not in our rule-following subgroup still showed a significant increase in left-right movement over time, even though this was steeper and generally higher in rule-followers. It is possible that participants who do not discover the rule explicitly, still learned implicitly about the distribution of black cards and adjusted their search behaviour accordingly. A regression between clicks on the last board and the regression coefficient of individual increases in horizontal motion also showed the expected relationship, though this did not reach significance, again indicating that both the number of clicks and the motion metric are no direct measures of discovery, but that they do correlate, both with each other and an explicit discovery of the rule.

Unlike LeGris et al. (2024), we did not query (via self-report) if participants discovered the rule via a sudden insight. We thus make no claim that rule-followers discovered the rule in a discontinuous fashion. The task does however make it possible to discover the rule from one moment to the next.

Continuous measures in a rich task environment provides rich data open to a range of analysis options, of which we have only shown a few. As viewers of reconstructed games may notice, dynamic reconstructions make it apparent that many recorded behaviours remain unexplained and underexplored, particularly in the mouse movement. Our focus in the use of this task was to explore measures of metacognition in chapter VI. But future studies can extend the measures we have presented here to identify fine-grained behavioural predictors of cognitive discovery that discriminate rule-followers from those who remain oblivious.

The examples of LeGris et al. (2024) and our own, show that observations of dynamic reconstructions can inform hypotheses and measurement, especially in empirical domains where quantitative measures are just being developed as is the case in e.g. fine-grained behavioural markers of discovery or models of the self.

We suggest that dynamic reconstruction can also give rise to new experimental designs that are tailored to zoom in on behavioural patterns that were only observed in the subset of participants in the first sample. In the present study, we for instance observed what we called ‘freezing’ behaviours, in which, at a moment of high uncertainty, participants would move the mouse cursor outside of the response space and suddenly stop moving, before moving again with a straight line at rapid pace to a location which they then clicked without any further search, which gave the appearance of ‘stepping back to think’, akin to a person stepping back from a painting to try and notice any patterns they might have missed. One interpretation is of course that our theory of mind overinterprets such behaviours, but another possibility is that this behaviour is an *epistemic action* as introduced in chapter I, i.e. an action that aids a cognitive process. The analysis of this action could thus help to understand the cognitive process or help to know whether and how such an action does in fact lead to improved performance (probability of correct clicks after the freeze). We quantified and captured this behaviour, moving back and forth between the measure the dynamic reconstructions. Scaling a functional measure to the whole sample however showed that only a handful of participants showed this behaviour. This is the point at which the experimenter has two choices: discard the analysis (this was our choice in this case) or, if the researchers still have a hunch they might be onto something, design a new experiment which zooms in on that behaviour (Figure 10 (t_3))

Dynamic reconstructions cannot be used to provide statistical support for hypotheses, but we suggest it could be used to generate them. How can experimenters ensure that they not only exclude known hypotheses by inquiry, but continuously remain open to uncovering other plausible new hypotheses and measures and behavioural markers? We suggest we make use of our very own interpersonal inference machinery.

Chapter VI

Abstract

We study peoples' metacognitive model of their own learning dynamics in a spatial rule discovery task. Participants who made a discovery are later asked to re-enact their own behaviour of how they transitioned from an epistemic state of ignorance to knowledge. Through a comparison between participants' behavioural markers of playing the game for real, and their behaviour during re-enactment, we were able to discern systematic successes, failures and biases in participants' simulations of their former selves.

Introduction

"The Empirical Self of each of us is all that he is tempted to call by the name of me. [...] We are dealing with a fluctuating material" (p.291, James, 1901)

In the last chapter, we showed that we can measure how people transition from a state of ignorance about a rule to a state of knowledge about this rule. Behavioural markers in reaction times, number of errors and types of errors, mouse movements and self-reports, indicated a representational shift between epistemic states.

In chapter VI, we now want to propose that we can use that same task to study peoples' models of themselves during epistemic state transitions. We will do so by asking them to re-enact the behaviours associated with the transition, i.e. ask them to *act* as if their cognitive state was different than it actually is. Acting in accordance with a *counterfactual* epistemic state we assume to require participants to access the cognitive processes that gave rise to the behaviours that are particular to that counterfactual state. We posit that those behaviours which are accurately re-instated, are retrieved from cognitive processes that are *included* in peoples' models of themselves, whereas systematic failures or biases in acting behaviour might indicate cognitive processes that are out of model reach.

We present an exploratory experimental design which leans heavily on Mazor, Firestone, et al. (2024), who used an *epistemic pretence* condition, during which participants made a series of choices while pretending *not* to know something that they did in fact know. This was a condition intended to elicit self-simulation, or in other words: a condition which asked participants to run their self-model forwards. The utility of forward-runs is thought to be one of the reasons we have self-models at all: compressed and structured representations of our (past) actions, preferences and goals allows us to make predictions about how we will act in the future and fill in the memory gaps of how we acted in the past (Mazor, 2025). Self-predictability (which necessitates some self-consistency) has implications for the actions we take to accomplish long-term goals, commit to tasks or to relationships (Bodner & Prelec, 2003). Such model(s) of ourselves likely rely heavily on our memories of how we have acted and felt in the past, as well as on cultural expectations and the labels we are assigned to by others.

Self-knowledge of our consistent tendencies and preferences (e.g. "I have always enjoyed gardening, so I am likely to enjoy the job in the botanical gardens") is as useful as models of how we change: of how we 'change our mind', undergo a 'change of heart' and how we learn or 'update' our self-model. Sometimes, it turns out that we are not the people we thought we were. According to McCoy et al. (2019), 75% of participants said that they learned something new about themselves, when they had to make personal choices over fantastical scenarios that they were presented with (e.g. 'becoming a vampire'), showing that self-simulation of counterfactuals and even impossibilities (through what the authors call 'modal prospection'), is both possible and experienced as informative. As the introductory quote by William James suggests, models, including those of the self, are subject to revision and usefully so: Paul et al. (2023) propose that this adaptiveness is what provides a cohesive reference (or indexing) function to an agent in shifting contexts. The flexibility of the construct is also stressed in Dennett (1992), who suggests that narration is the primary force driving the experience of selfhood.

We were particularly interested in peoples' ability to model the effects of inner change and tackle the question at a lower level of abstraction. We asked to what extent people have access to the cognitive processes that underpin a shift from ignorance to knowledge. There are a few reasons to study self-simulations of counterfactual knowledge states. Consider just two: simulating counterfactuals is both common-place (consider for instance, the case of lying (Oey et al., 2023), where people might pretend that they are in a counterfactual state of knowledge, hiding secrets and lying that they do not know something); and often plays a central role in questions about ethical conduct (consider for instance, the case of spy-craft (Fabre, 2022), where an act such as a lying might be considered morally permissible if, counterfactually, truth-telling would put informants in danger).

The concealment of one's true inner state plays some role in most interactions people have (consider the case of 'putting on a brave face'). Second, knowledge states can easily be manipulated experimentally, by showing or telling participants information that they previously did not have access to. Third, Mazor, Firestone, et al. (2024) already showed that quantitative behavioural measures of participants pretending not to know something, can provide insight into the boundary of what is and is not included in peoples self-model and furthermore what biases these models might be subject to.

As mentioned in the previous chapter, many results from the metacognitive research community show that metacognitive access is incomplete. Consider just two examples. Levin & Angelone (2008), who asked participants to predict and assess their perceptual abilities in tasks that are known to showcase perceptual inaccuracies (such as change blindness). The authors found large overestimations in peoples' self-evaluation. In an absent/presence detection task Schipper & Mazor (2025) show that metacognitive judgement about absence was modulated by an incorrect belief about what feature made stimuli harder to detect. Participants' belief about perception was not true to their actual perception. With the method of acting, Mazor, Firestone, et al. (2024) too find inaccuracies, but also shows that participants are able to put on a convincing act.

In Mazor, Firestone, et al. (2024), participants played games (such as Battleship), in which they knew the correct answer, but were incentivised to *act as if* they did not know the solution. In a within-subject comparison between actions taken in 'pretend-games' versus real games, Mazor, Firestone, et al. (2024) found that participants who acted made choices that were more rational than chance, indicating that they did not just behave randomly. Participants also accurately reflected the varying difficulty of different game rounds, by varying their number of guesses. The spatial distribution of choices was significantly correlated across playing and acting games and actors also emulated reaction time differences between hits and misses, as well as the uncertainty of the game state. Moreover, in the behavioural deviation from real games Mazor, Firestone, et al. (2024) found evidence for that acting choices were retrieved from a generative, model-based self-representation: participants showed a regression to the mean (underestimation of difficulty in hard game rounds and overestimation of difficulty in easy rounds) and their choices had lower entropy. This is in line with expectations if one assumes that action choices are guided by a generative, compressed model of memory for actions, biased by distortions towards prototypically (Spens & Burgess, 2024).

We want to extend this work. To show how, we now briefly refresh the task that we introduced in chapter V. Participants turned cards on a grid, aiming to find the remainder of a set of black cards, whilst avoiding white cards. They could discover that a secret rule determines the locations of black cards: blacks are arranged symmetrically left and right of midline.

In the present study, participants who do discover the rule (i.e. we select the class of rule-followers and discard the class ‘other’), will repeat the same task once more, except that now, just as in Mazor, Firestone, et al. (2024), the correct locations are marked with a visible cross, so that participants have access to two kinds of information at all times: they know exactly where the black cards are hidden *and* they know about the symmetric arrangement. When participants in Mazor, Firestone, et al. (2024) acted, they were in a state of knowledge and acted as if they were ignorant at all times. In our study, we will instead require that *actors pretend to transition between a state of ignorance to a state of knowledge*. Our design also differs in that we increase the temporal distance between playing the game for real and acting. This ensures that we have access to groundtruth behaviour (playing for real) within the same participant, but we can also exclude the possibility that participants are merely recalling their action sequences by explicit memory, rather than eliciting their self-model. Akin to Boger et al. (2025) who found consistency between random sequences generated by participants across sessions with a long temporal delay, we also intend to elicit a generative model of choices, which, albeit likely based on memories of previous experiences of discovery, should not be based on direct memory of choices participants made when playing the game for real.

In chapter V, we presented a task in which an immeasurable (at least directly) cognitive transition from ignorance to knowledge is accompanied by shifts in behaviours. In chapter VI we now present another sample in this task. We then elicit participants inferences about how their own actions reflect an epistemic state transition. We do this by asking participants to reinstate their own behaviours without them being able to remember what they did exactly. A within-subject comparison between playing and acting along several behavioural markers of discovery, is expected to show some structure that is maintained, but also some systematic failures to replicate actions from real games faithfully. This is a feature of the design: we interpret behaviours that are not re-instantiated as resulting from cognitive processes that are not accessible to the person at that time, or indeed that the behaviours are not part of the person's model of how they behave when they make a discovery.

Our hypotheses about failures to re-instate behaviour faithfully, mostly concern the period before discovery. An intuitive reason as to why was introduced in chapter V: discovery remains an elusive phenomenon, one which we not only struggle to predict experimentally (Popp et al., 2024), but which we also often fail to find an explanation for in our personal lives. Academics cannot, for the most part, *plan* to have an idea or a new hypothesis on any given morning. This inquiry about where hypotheses come from (Dasgupta et al., 2017) and how we might shape what and how many hypotheses come to mind (Bear et al., 2020), by using metacognitive access to this process, is important (it could e.g. reduce prejudice, alleviate depression or facilitate scientific progress). Humans probably know more about how they select between explicit hypotheses (Klayman & Ha, 1989), than how the hypotheses came about (Johansson et al., 2006; Nisbett & Wilson, 1977), which is why we suspect that participants in our task will struggle to know what actions led them to make a discovery during play. We expect participants are less likely to re-enact the errors and search patterns that led them to notice the symmetry pattern.

We also expect that subtle behavioural markers of the ignorant knowledge state will not be part of the re-enactment. Failure to truly 'delete' known information from the behavioural markers is intimately tied to knowing what actions were taken prior to discovery. The actions that facilitated discovery were taken in a state of ignorance, but fully adhering to ignorant actions might be hard when one is knowledgeable. Intuitively speaking, in the theatre, the actor knows that they are not really Hamlet, which is why they need years of training to convince anyone that they in fact are. Empirically, Mazor, Firestone, et al. (2024) found systematic inabilities to ignore knowledge, as was shown in a letter guessing game, in which a word was partially revealed and in which players, who did *not* know the true, infrequent word, had biases towards guessing particular letters as common words sprung to mind more easily. Actors, who knew the true, infrequent word, however showed a skewed distribution of letter guesses towards the true words. They guessed letters that an ignorant participant would rarely guess and could not prevent the leakage of secret information into the action sequence or hypothesis generation process. We expect similar effects in our study. We pre-registered the assessment of differences across conditions and knowledge states along several measures which are detailed in the methods sections: the number of clicks to completion, the rank of choices and errors by models of different constraints, two different reaction time patterns, horizontal mouse movements and hypothesising. We expected to find evidence of distortions in behaviour, such as evidence for more drastic behaviours than during game play, which would be evidence in line with Mazor, Firestone, et al. (2024), who also found 'over-acting' and systematic distortions towards prototypicality (Hemmer & Steyvers, 2009; Kerrén et al., 2024), suggesting that participants retrieve a generative model when they act.

Re-enactment is a powerful metacognitive research design framework, because it can in principle be applied to any behaviour of interest and does not assume that a metacognitive process need be explicit. Metacognitive access (or improvement) can be examined by asking people to run the model of that process forwards. We here introduce a high-resolution, within-subject design which provides us with a ground truth of how each participant did in fact behave and allows us to analyse differences in subtle behavioural markers of an epistemic shift.

Methods

Participants

The research complied with all relevant ethical regulations and was approved by the University of Oxford Central University Research Ethics Committee (code R90986/RE002). Participants were recruited via Prolific and gave informed consent prior to their participation.

To be eligible to take part in this study, participants' Prolific approval rate had to be 95% or higher and their reported first language had to be English. We sample within the USA and the UK. Participants were instructed that they could only enter the game if they were not using a tablet/smartphone.

Participants were paid £9/hour. We paid £1 bonus payment to all rule discovers in part 1, and for part 2 we paid an additional £1 to the top 30%, who were ranked in terms of their difference in total number of clicks in round 2 versus 1. We then paid an additional £4 to the top 3 participants in that ranking.

Exclusions

We collected 218 participants in part 1, 14 of which were excluded based on having used a tabled (exclusion method as described in Chapter V). As pre-registered (Cremer & Mazor, 2025b), we invited only participants who met our criterion for rule discovery to join part 2. 119 participants had an accuracy of 100% on the 2AFC test. 47 of them named the rule explicitly ('symmetric','mirrored'). 14 additional participants did not name the rule but also had 100% on the 2AFC test and less than a single error on the half-revealed board. Due to the low rate of participants who explicitly named the rule, we amended our pre-registered plan and included those 14 participants in the set of participants who were re-invited.

52 out of 61 participants we invited to participate in part 2, returned to the study. We excluded all participants who did not make any errors on the first board played or who gave

clearly incorrect responses in our comprehension check in which they were asked to describe their task (descriptions did not however need to be complete). As per our exclusion criteria, we excluded 8 participants, six of which did not make an error in the first round. This resulted in a sample of 44 participants. One participant was excluded (only) from the RT analysis because their median RT on the first click was more than four times larger than the average median RT of participants in the cohort.

Design

We constructed a within-subject, two-parts experiment. We will briefly describe the structure of the two experiments and how they relate, before diving into the details of each part in the next section.

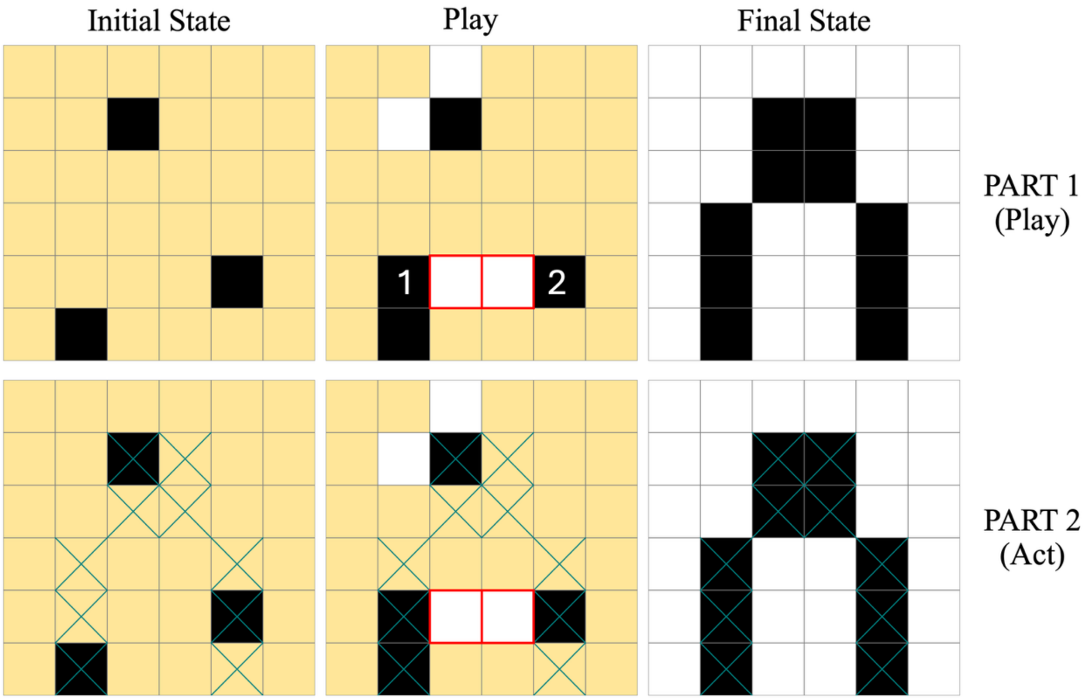

Figure 16: **Illustration of Design Chapter VI** Three exemplary board states in Part 1 (top) and Part 2 (bottom). Yellow indicates unrevealed grid locations. White indicates that a revealed location did not contain a target. Black indicates that a target was revealed. **Top Left:** The initial state of a board (each round starts with a new board) as seen by a

participant playing Part 1. Their goal is to reveal the rest of the 10 black cards that are hidden underneath the yellow cards. Participants may discover a hidden spatial rule: black cards are symmetrically arranged along the vertical midline. At first sight of each board, one black card from a distinct group (a connected set of black cards of size 2-4) is already revealed. **Top Middle:** Participants play by revealing cards via a mouse click. If they reveal a white card, they lose 1 point. We show errors (two white cards framed in red left and right of the vertical midline) which reveal that the participant does not play this round in accordance with the spatial symmetry rule. A rule-follower knows that the symmetric twin of an already revealed white card is predictably going to be white. **Top Right:** The final state of the board once all black cards have been revealed. The remaining yellow cards fade into white. **Bottom Left:** The initial state of the same board as seen by the same participant in Part 2. The cross hints now show all locations of black cards at all times. Participants in Part 2 now know the symmetry rule and the locations of black cards during the full game. **Bottom Middle:** Their goal is now to play this board as they did when they only knew what they knew in that round of the game in Part 1. Here we show an ideal case in which the participant in Part 2 replicates the exact errors they made in Part 1, including the symmetric errors that are a result of not knowing the rule. **Bottom Right:** Final state of the board once all black cards have been revealed.

In Part 1, participants played the same game as in chapter V with minor variations (Figure 16 Top Row). To recap the main structure of the game: participants were presented with a grid of black and white cards, hidden behind a yellow cover. Upon a mouse-click, the hidden card was revealed. Participants were asked to uncover all black cards while revealing as few white cards as possible. Each white card subtracted one point from their total initial game score. During the game, participants had a chance to discover that target locations are predictably arranged according to a spatial rule: mirror-symmetry along the x-axis.

After two weeks, participants were invited to play Part 2. They then completed the exact same game again, except that this time they were told that their goal is to act exactly as they did previously. This required them to perform in accordance to a counterfactual (pre-discovery) knowledge state in the first rounds, because they now both knew the spatial symmetry rule *and* the locations of black cards (which we marked with a cross; Figure 16 Bottom Row).

We will compare Part 2 against Part 1 across our behavioural markers of learning to assess whether there are systematic biases and failures in retrieving the cognitive processes which led to participants' discovery behaviour in Part 1.

Since part 2 required participant to *act as if* they do not know the location of black cards, we will refer to part 2 as 'Act' in distinction to part 1, which we refer to as 'Play'.

Procedure

Part 1 / Play

Where this part of the study deviates from the design described in chapter V, we show the description in **bold**. The game we describe in this section can be accessed via the following link: https://carlacremer.github.io/discovery_play_memory_2/.

Participants were told that the experiment is part of a two-part experiment and that they would be paid after completing both parts. They were then instructed about the rules of the game.

Their goal was to reveal all 10 black cards while revealing as few white cards as possible. Participants started the game with **200 points** and lost 1 point for every white card uncovered. A scorecard on the screen showed participants' remaining points. Participants were informed that black cards are arranged in groups of size 2, 3, or 4 and that group corners may touch. At the start, each board had a single black card per group already revealed (Figure 16 Top Left). The black cards that were already revealed were never a mirror-symmetric pair along the vertical midline.

A participant **played 4 boards**, each of which presented a new arrangement of 10 black cards (see Figure 17 A; we shortened the full game to 4 boards + one special, half-revealed board, to reduce the cognitive load for participants in part 2.) Each person played the same unique boards, and their order was randomised across participants. During the instructions, participants passed comprehension checks and were shown explanatory images. They also played one practice board in which they uncovered a single group (a two-row horizontal bar) of black cards (which was symmetric along both the x and y-axis).

Before they begun, **participants were told that black cards are arranged according to a simple, secret rule**, which, if discovered, will help them retain points (we introduced this change to increase the number of people who discovery the rule and can be reinvited to part 2). The rule was that all black cards in all boards are arranged symmetrically (Figure 17 C). Participants who discovered the symmetric spatial arrangement could retain more points (avoid revealing white cards) than those who did not. For each new black card revealed, there was one more correct location (the symmetric twin) that could be known to the rule-follower immediately.

After the last black card of every board was revealed, an animation showed the gradual fading of all cards except all black cards, which momentarily highlights the symmetric arrangement (Figure 16 Top Right). After 4 rounds, participants were given specific instructions to play one more board. They were told that this board contains 14 black cards, exactly half of which are already revealed (Figure 17). Unbeknownst to them, the hidden black cards on this board were always the mirror-symmetric cards of cards that are already revealed. This board was the same for all participants and was always shown last. A participant who understood that the rule does not, in theory, need to incur the cost of revealing any white card on this board.

Participants were given no time constraint. In a short pause between each board played, participants were told about the loss they incurred in the last round. They could then take a self-paced break. They were shown an image of a coloured number that marked the identity of the upcoming round (Figure 17 A). They clicked a button to commence the next round.

In between rounds 3 and 4, participants were **asked if they had already discovered the rule** and if they could **name the rule** by typing a single-word description into a response box.

After all boards have been completed, participants were asked again if they discovered the rule, if they can name the rule, and **at what point they discovered it**. For this last question they were shown all coloured numbers (which marked the rounds of the game) and selected one of the given response options (a round number or “I didn’t discover it”). They then completed **five 2AFC trials** in which we tested their ability to distinguish between boards that follow the rule versus that did not. On every trial, participants were shown two boards (Figure 17 D) that followed the rules of the game (group size between 2-4, 10 black cards, corners of groups may touch). One of the boards (in randomised order) was symmetrically arranged along the x-axis (Figure 17 D Top) and the other was either asymmetrically arranged (Figure 17 D Bottom) or symmetric along the y-axis. Participants were instructed to click on the board which followed the secret rule. They got no feedback.

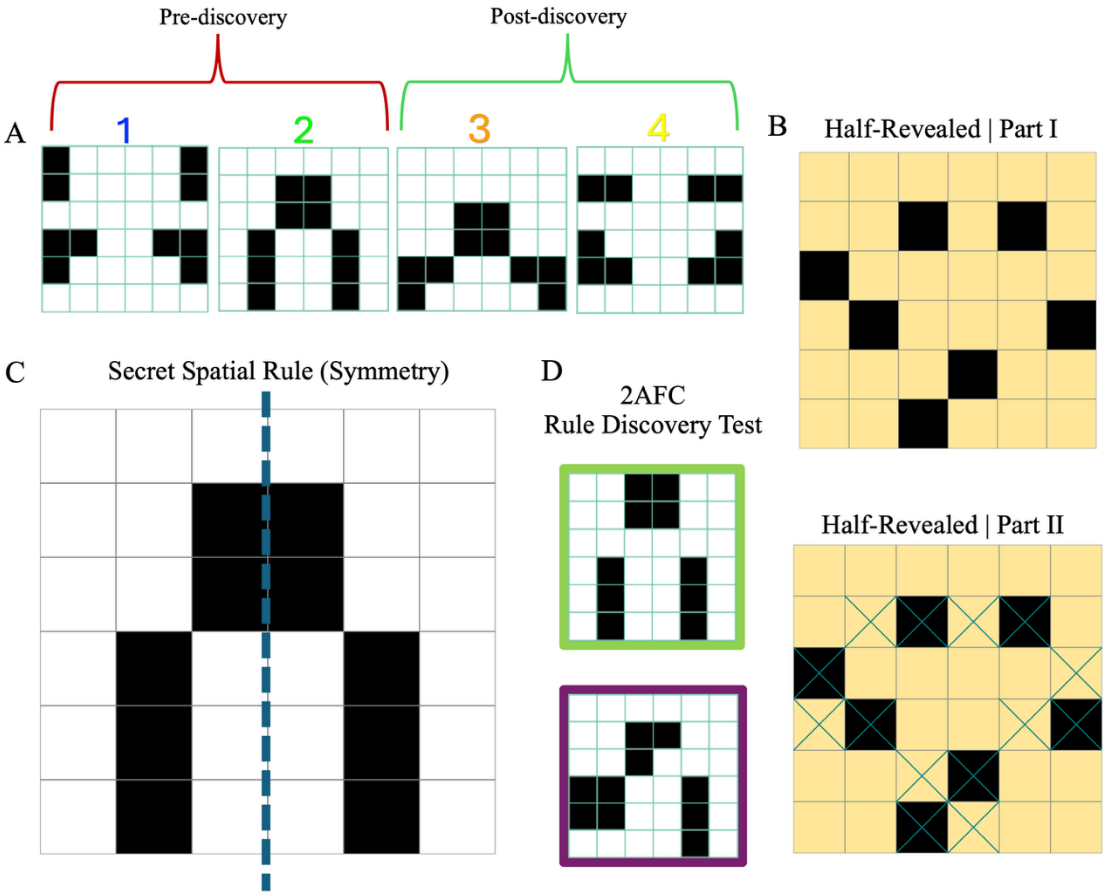

Figure 17: **Second Illustration of Design Chapter VI A:** All unique boards (showing final state) played during the first 4 rounds. Each board is symmetrically arranged along the vertical midline and associated with a coloured number, indicating the round (timepoint) in which the board is played. The order in which the first four boards were seen was randomised across participants. **B:** Top: The last round always showed this half-revealed board (number five), in which exactly half of all (14) black cards were already revealed. One of each mirror symmetric twin was already revealed. Participants who know the symmetric rule can complete this board without revealing any white cards. Bottom: We showed the same half-revealed board in Part 2. All black cards are marked with a cross. **D:** Example of one of five 2AFC trials which were played after the game. Participants were shown two boards and were asked to select the board in which the arrangement of black

cards follows the rule. One of the boards (framed in green) was an arrangement which they had not previously seen but which followed the symmetry rule (along the x-axis). The other board (framed in red) was an asymmetric board which otherwise followed the rules of the game or the same board as the correct board but flipped to be symmetric along the y-axis. The coloured frame is added here for illustrative purposes only and was not shown to participants.

Part 2 / Act

Two weeks later, participants who discovered the rule (see section exclusions) were invited to join Part 2 (Act). The game we describe in this section can be accessed here: https://carlacremer.github.io/discovery_act_memory_2/.

First, participants' memory of the game rules was refreshed. Participants were then informed about a change in how the game is to be played. An illustrative image showed participants how cross hints now indicate the locations of all unrevealed black cards at all times (see Figure 16 Bottom). The instructions now also revealed the secret rule: "Black cards are always arranged symmetrically to the left and right of the midline."

Participants were then informed that their goal in Part 2 is to reconstruct their own behaviour from Part 1.[6] Rather than trying to remember their exact behaviour, participants were asked to "think about how you would behave if you only knew what you knew then".

---

[6] We briefly list the exact wording of the instructions we used, because the experiment is instructions sensitive. Our phrasing included relevant sentences such as: 'Can you play part 2 so that it resembles how you played part 1?', 'In each round, just try to think about how you would behave if you only knew what you knew then.', 'Can you play the game like you did when you didn't know the locations of black cards?', 'this experiment is about whether you can behave like you would if you didn't know something that you actually already know.',

The instructions reminded participants that in Part 1, they did not know the symmetry rule at first. We added a comprehension check with multiple options to choose from (comprehension check 1) to test if participants understood that they must only use their knowledge of the symmetry rule from the timepoint at which they in fact came to know it in Part 1.

They were informed that their behaviour in Part 1 will serve as the ground truth for their behaviour in the counterfactual knowledge state and that their clicks, reaction times, and mouse movements will be recorded and compared. They were reminded that their bonus payment now only depended on how closely their Part 2 gameplay resembled their Part 1 gameplay (including their errors). They were informed that the maximum bonus for Part 2 is $5. Before participants began, we asked them to describe the task in detail by typing a description into a response box (comprehension check 2).

Before each new round, participants responded to the question “When you played round [round number] last time, had you already discovered that black cards are arranged symmetrically?” with either yes or no. All else remained as it was in Part 1. Each participant played the exact same order of boards that they played previously. Before the half-revealed board, participants were reminded of the instructions they got in part 1 before this board but were told again that their goal remained as specified in part 2: to act as if they only had the knowledge they had at the time. At the end participants answered two debrief questions: 1) if they acted as if they had discovered the rule, and if so, how; and 2) what they thought of the experiment. In a previous pilot study, we had added a short memory test before part 2 began,

---

‘Your bonus now only depends on how closely today’s game play (with knowledge) resembles your previous game play (without knowledge).’

to measure if participants remembered which board arrangements they had played. We found no above chance memory recall after two weeks.

Optimal behaviour in the task demanded that participants simulate a counterfactual state of mind which depends on the timepoint of the game: in early rounds they must act as if they neither know the symmetry rule nor the location of black cards, while in later rounds they must act as if they have discovered the symmetry rule but are still unaware of the exact location of black cards. Can participants re-enact the discovery of a rule even though they already know it?

Dynamic Reconstructions

We used generated and used reconstructions of all 44 participants as described in chapter V. We make several of these reconstructions available to the reader via the following link: https://carlacremer.github.io/reconstruction_VI/. A sample of participants' game play can be observed in each round in both Play and Act. We show a static reconstruction in Figure 18 A and illustrate that observations of dynamic reconstructions were an input to the range of measures we used (Figure 18 C).

Reconstructions were implemented as described in chapter V, methods section.

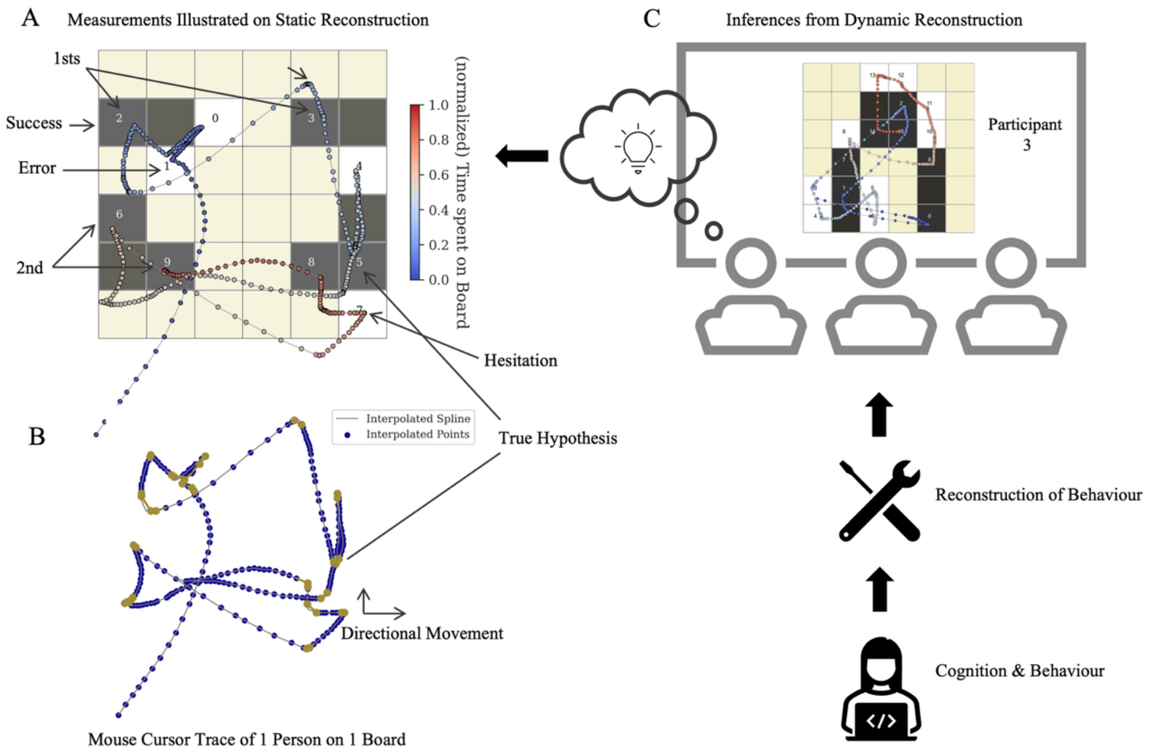

Figure 18: **Illustration of Behavioural Measures of Knowledge States & Method of Reconstruction A:** Static Reconstruction of Board Played by 1 Participant. Dot motion traces mouse movement over time (earlier = blue, later = red). Numbers indicate click order. Black cards with no number were already revealed. Click on black cards = successful click. Click on white card = Error. 1st are the set of black cards that were revealed as the first of a symmetric pair. 2nds are the set of black bards that were revealed as the second of a symmetric pair. Correct hypotheses are hovers over unrevealed black cards but not clicking on them **B:** Interpolation points tracing the recorded mouse cursor locations (from A). Sharp angles were analysed and here visualised to label timepoints of potential decision events. Cursor locations were used to compute directional movement along the horizontal and vertical direction of the screen. **C:** Measures were generated, refined and revised by observing reconstructions of games played.

Behaviour Markers of Discovery

The model-based analysis, RT divergence and mouse motion direction analysis are as described in chapter V. We add three additional markers: hypothesising behaviour, hesitation behaviour and self-reports. All new methods are described below.

Knowledge-State–Dependent Constraints

The naive and symmetric models were constructed just as described in chapter V, methods section.

RT Divergence

As described in chapter V, methods section. See illustration of measure in Figure 18 A: 1st and 2nd.

Motion Direction

As described in chapter V, methods section. See illustration of measure in Figure 18 B: Directional Movement.

Participants' Hypotheses

In this chapter we add an analysis of hypothesising, which we define to be actions that indicate that a participant is considering a particular location, but does choose to reveal that location at the time.

We capture hypotheses by analysing participants' mouse cursor movement. A cursor position is recorded throughout the experiment. We trace the trajectory across each board (Figure 18 B) and mark angles (between 30-180 degrees) as potential decision points (see marked, golden areas in Figure 18 B). Within this range of redirecting movements, we select all phases in which a participant hovered over the same unclicked card for more than 500 milliseconds. If they did not click on this card, we consider this event a hypothesis, i.e. a click that might have been committed to, but which the participant decided against after considering it. If the location was clicked on during a later, different hover period (the mouse must have left the location once), it is still counted as a hypothesis at that time. In part 1, these hypotheses were therefore falsely rejected hypotheses about the true location of a black card. In part 2 these actions were based on insider knowledge.

Hesitation

We define hesitations as the time passed between the selection of a location by mouse position and the click event (Figure 18 A: Hesitation). We previously defined RT as the time between the previous click and the current click. Hesitation is the time between the last timepoint in which the cursor was outside the selected card and the current click. For instance, a hesitation of 200 ms means that the player waited 200 ms from the moment of entering the card with the mouse before clicking.

For every row of x and y coordinates which recorded mouse cursor position on screen, we identified whether the card over which the cursor was located, was the same or different. For all (rows) recorded entries in which the cursor remained over the same card, we added a distinct ID number (per participant, per board). We identified all IDs in which a click occurred. This click moment (row) was labelled as either black (1) or white (0). We took each click made and calculated the difference (in ms) of time passed between the start of the uninterrupted hovering over the card that was clicked (identified by it’s unique ID), up to the moment at which the click was made. This difference in time (converted into seconds) is what we referred to as a hesitation.

Self-reports of Knowledge State

8 participants did not enter a timepoint of discovery (they had the options of choosing ‘I did not discover it.’) in Play or Act, despite being able to name the rule. This also means that

some participants who did provide a round number, maybe did not remember or were unsure, but gave a (therefore unreliable) response anyways. We chose, in violation of our pre-registered plans (Cremer & Mazor, 2025b and Appendix B), not to use the individual's self-reported discovery round to split the data per knowledge state, to avoid having to exclude those participants, based on an indicator that likely is only an approximation of the discovery moment. We instead chose to label pre and post-discovery rounds in accordance with the most frequent reported round of discovery (see Results Section, which also matched our three previous pilot experiments in which we had found a marked drop in errors in that round.

Statistical Analyses and Software Packages

Data analyses were conducted using *Python* 3.10.13, in particular *NumPy* 1.26.4 and *Pandas* 2.2.2, figures were generated using *Matplotlib* 3.9.2 and *Seaborn* 0.13.2. Experiments were coded in Jatos (Lange et al., 2015). Rank was computed using *SciPy* 1.13.1. Mixed effects models were computed using *Statsmodels* 0.14.2.

We deviated from the pre-registered plan to mostly run ANOVAs using knowledge state and part, since not all participants showed the behaviours, we were testing in all factor combinations which thus reduced our sample size. To ensure we made use of all the data we collected, we mainly used simple linear mixed effects models in which we added each participant as a random factor (Wilson & Collins, 2019). When we used two predictors, we report interaction effects. Where appropriate, we verified the results with t-test, Welch test and ANOVAs, but do not report them all here. All our following analyses on behavioural markers of discovery are based on data from the first 4 boards, as per our pre-registration.

Results

Self-Reported Timepoint of Discovery

13 out of all 44 participants indicated the same round as their round of discovery across both part 1 (play) and 2 (act).

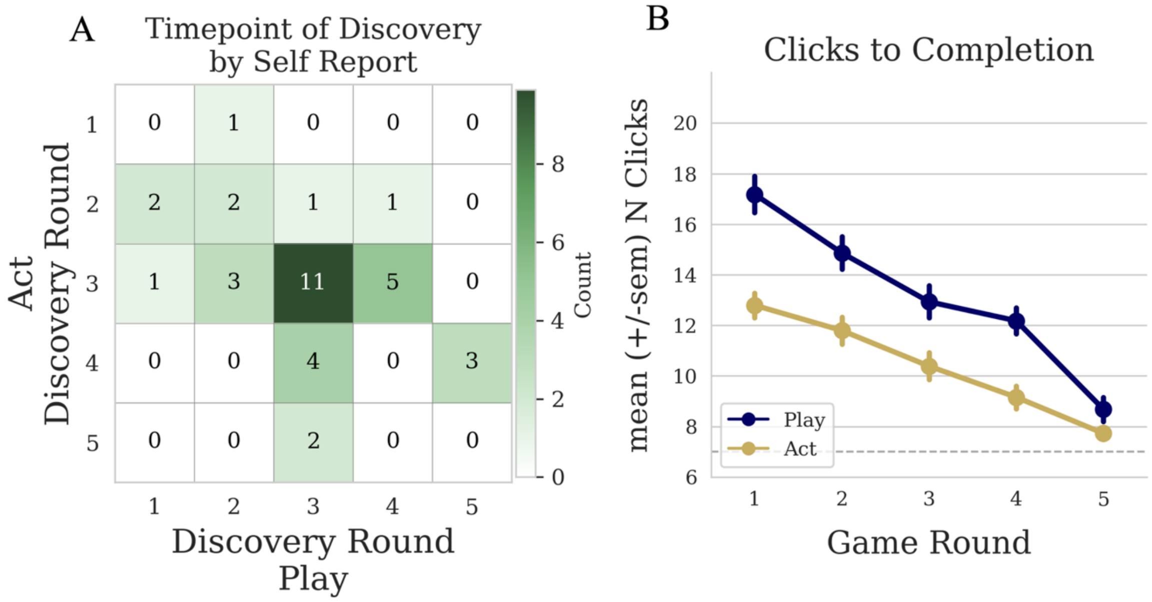

Figure 19: **Discovery in Reports and Clicks to Completion Chapter VI A:** Game round in which discovery of symmetry was made, according to self-report: Count of self-reported discovery round in Play (columns) vs Act (rows). Diagonal entries and their values show number of participants (13) whose self-report was consistent across Act and Play. **B:** Performance measured in clicks to completions (ideal performance is 7 clicks), showing average number of clicks per game round in Play vs Act.

Figure 19 A shows how many participants reported each round = {1,2,3,4,5} as their discovery round in both Play (columns) and Act (rows). Most participants did not remember exactly when they discovered (or reported discovery) by the time that two weeks had passed. Participants tended to report discovery in round 3 during both play and act. In part 2, participants were not far off the judgement they gave in part 1. For instance, no participant who said they discovered the rule in round 1 during Play, said they discovered it in round 5 during Act. A chi-square test does not consider the distributions of responses to be independent ($\chi^2(16) = 30.2333$, $p = 0.0168$).

To analyse the effect of knowledge state (ignorant versus informed about the symmetry rule) on the measures to follow, we split the boards in half and label the first two rounds as 'pre-discovery' and the last two as post-discovery. This is our categorical label of knowledge state, based on the most frequently reported round of discovery. We illustrate this in Figure 17 A: the first two rounds are labelled pre-discovery, i.e. as the naive knowledge state (knowledge state = 0). The last two rounds are labelled post-discovery (knowledge state = 1).

Number of Errors

We measured performance by the number of clicks to completion. Participants showed improvement with the number of rounds played. This was the case for both Play and Act. They used fewer clicks (thereby revealing fewer white cards and retaining more points) over time.

In Figure 19 **B** we plot the mean and standard error of clicks to completion across participants in each condition. The minimal number of clicks required to complete each round was 7.

Based on pilot studies, we expected a main effect of acting, with fewer clicks (in part 2) showing an overestimation of their abilities to uncover black cards.

We apply a mixed effect linear regression in which both condition {Play, Act} and knowledge state {0,1} were categorical predictors of the number of clicks. Individual participant IDs were modelled with individual intercepts, as random effects.

When acting, participants made significantly fewer errors, as is shown by a negative, significant main effect of Act relative to Play: $\beta = -3.7273$, SE = 0.5415, $z = -6.88$, 95% CI [-4.7886, -2.666], $p = <.001$.

The task was constructed such that the discovery of symmetry allowed participants to retain points by avoiding white cards and completing each round with fewer clicks. We thus expected a main effect of knowledge state with fewer clicks after knowledge is acquired, or participants act as if knowledge was acquired.

The effect of knowledge state was significant, yielding a reduction in the number of clicks as is expected by the knowledge about symmetry: $\beta = -3.4659$, SE = 0.5415, $z = -6.4$, 95% CI [-4.5272, -2.4046], $p = <.001$. The interaction effect was insignificant ($\beta = 0.9432$, SE = 0.7658, $z = 1.23$, 95% CI [-0.5577, 2.4441], $p = 0.218$), showing no deviation for how much knowledge acquisition reduced the number of clicks when participants acted as if they gained knowledge.

The analysis on the number of clicks shows that actors understood that they had to make more errors at the start of the game than at the end of the game. They underestimated the number of errors they actually made when they played the game in the ignorant state during the first rounds. They did however reduce the number of errors, consistent with the behaviour of rule discovery in real games and ended up with near optimal performance on the last, half-revealed board (Figure 19 round 5). In terms of the *reduction* of the number of clicks, they thus behaved as if they had not discovered the rule at the start and behaved as if they discovered the rule by the end of the game.

Model Fits

In Figure 20 we plot model rank as a function of game round, for choices that yielded both white and black cards (Top Row) and for chosen white cards only (Bottom Row). We show model evaluation for participants during Play (Left Column) and Act (Right Column). The symmetric model rank is plotted in green.

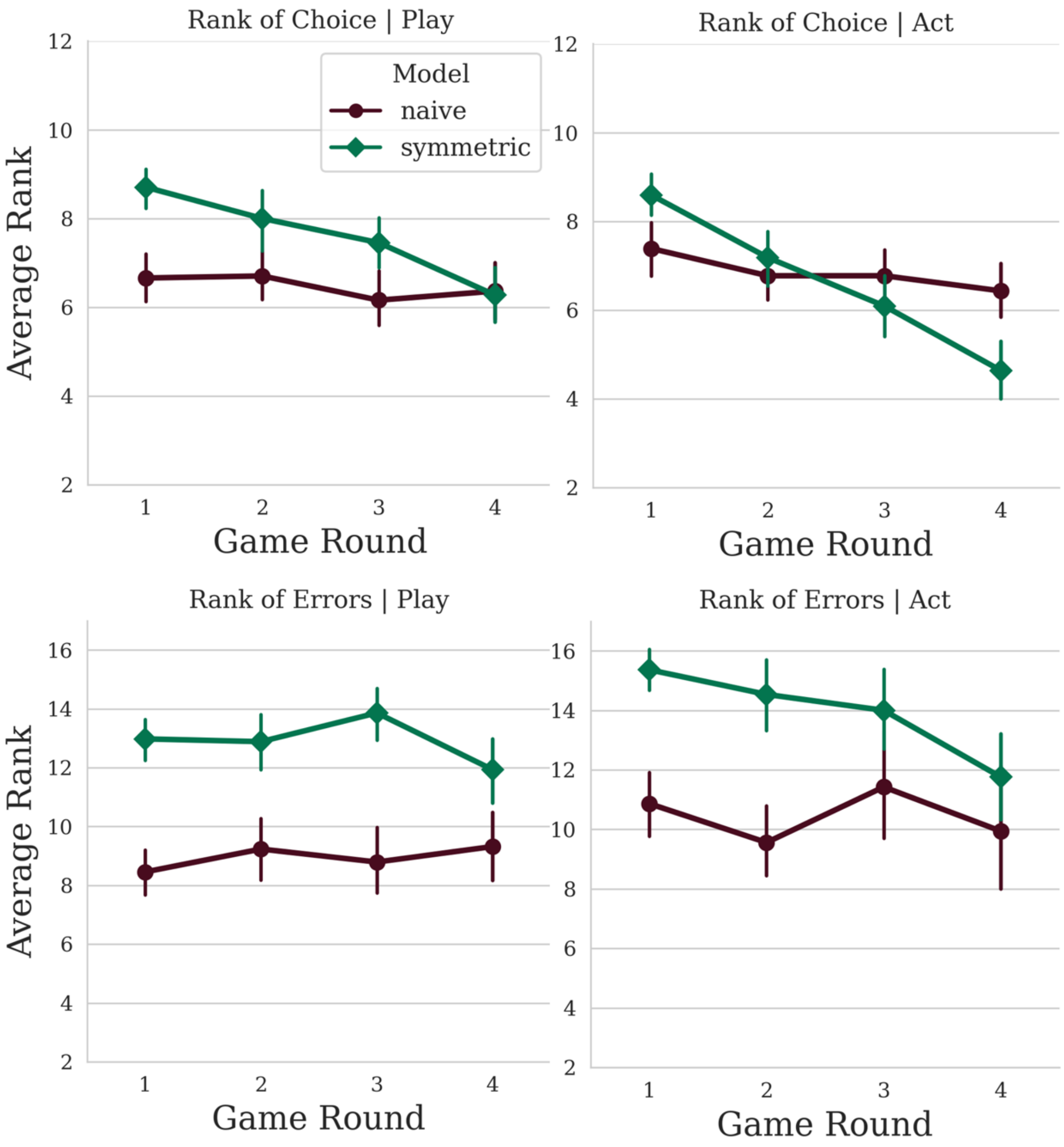

Figure 20: **Model Ranks Chapter VI Top Left:** Choices while playing (hits and misses) ranked by the naive (red) and symmetric (green) model. We plot the average rank (+/-sem) across all participants for each game round. **Top Right:** Choices while acting ranked by the same models. **Bottom Left:** Rank of errors while playing. **Bottom Right:** Rank of errors while acting.

We predicted the rank of the symmetric model (choices) using a mixed effect model with condition ={Play,Act} and knowledge state ={0,1} as predictors. We find no significant effect of condition (β = -0.3903, SE = 0.3139, z = -1.24, 95% CI [-1.0056, 0.225], p = 0.214), a significant effect of knowledge state (β = -1.4769, SE = 0.3139, z = -4.7, 95% CI [-2.0921, -0.8616], p = <.001) and an interaction effect: β = -1.0375, SE = 0.444, z = -2.34, 95% CI [-1.9077, -0.1674], p = 0.019. The rank of choices as evaluated by the symmetric model declined with a transition into a knowledgeable state and this decline was stronger when participants acted.

When we apply the same statistical test to the rank of errors, we find a significant effect of condition (β = 1.866, SE = 0.5275, z = 3.54, 95% CI [0.8321, 2.8999], p = <.001), no significant effect of knowledge state (β = -0.2023, SE = 0.5275, z = -0.38, 95% CI [-1.2362, 0.8317], p = 0.701), and an interaction effect that approaches but does not reach significance: β = -1.3248, SE = 0.746, z = -1.78, 95% CI [-2.787, 0.1374], p = 0.076. Errors by actors are assessed as being more in line with the symmetry rule constraint in later rounds.

In chapter V, we found that rule-followers showed a significant decline in rank of errors as assessed by the symmetric model. Here, we do not replicate this result: a mixed effects model fit to errors made during Play (Figure 20 Bottom Left), in which we predicted the rank of error as assessed by the symmetric model by game round (and modelled each participant as a random effect), finds no significant improvement (decrease) in rank over time (β = -0.211, SE = 0.205, z = -1.03, 95% CI [-0.61, 0.19], p = 0.303).

The decline of the rank of choices (Figure 20 Top Left), under the same statistical test, however, remains significant (consistent with chapter V): β = -0.788, SE = 0.123, z = -6.38, 95% CI [-1.03, -0.55], p = < .001. When participants acted, their choices *and* errors both showed a significant decline in rank (Choices: β = -1.271, SE = 0.125, z = -10.16, 95% CI [-

1.52, -1.03], p = < .001; Errors: β = -1.009, SE = 0.262, z = -3.85, 95% CI [-1.52, -0.49], p = < .001). Decisions by participants in the acting condition showed a more drastic tendency to be in line with the symmetric rule as rounds progressed (Figure 20 Right Column).

Qualitative inspection of the rank of errors during Play (Figure 23 Bottom Left), also suggests that the naive model rank increases over time (which would be in line with expectations if errors become less naive), but this increase was not significant (β = 0.226, SE = 0.219, z = 1.03, 95% CI [-0.20, 0.66], p = 0.302).

It is of note that the naive *and* symmetric model provided a tendentially higher (worse) rank of errors made by actors. The symmetric model first indicates a high rank, but its assessment declines to the same level as it does during Play, while the naive model ranks actor errors about half a rank higher on average than it did for players.

Temporally Ordered vs Disordered Clicks

We tested the hypothesis that participants in part 2 make choices that are consistent with the state of the revealed board at each timestep. We did this by computing the rank, i.e. the suitability of choosing a location according to the naive model, which takes the rules of the game, but not the symmetry rule into account. In Figure 21 we show the rank of the real order of clicks versus the rank of a shuffled order of clicks. Shuffling was only done within either black or white clicks, such that all click IDs that were errors remained errors, and all click timepoints that led to success, also uncovered black cards. Within each group, we shuffled which card was revealed at what timepoint within each round.

If choices and errors made sense to a game-playing (but not rule-following) model, given the state of the board at the time of the click made, the naive model fit should show a lower rank for the real data than the shuffled data. In Figure 21 we see this in all conditions and click combinations, bar the naive model fit to errors made by actors, which shows a much smaller difference.

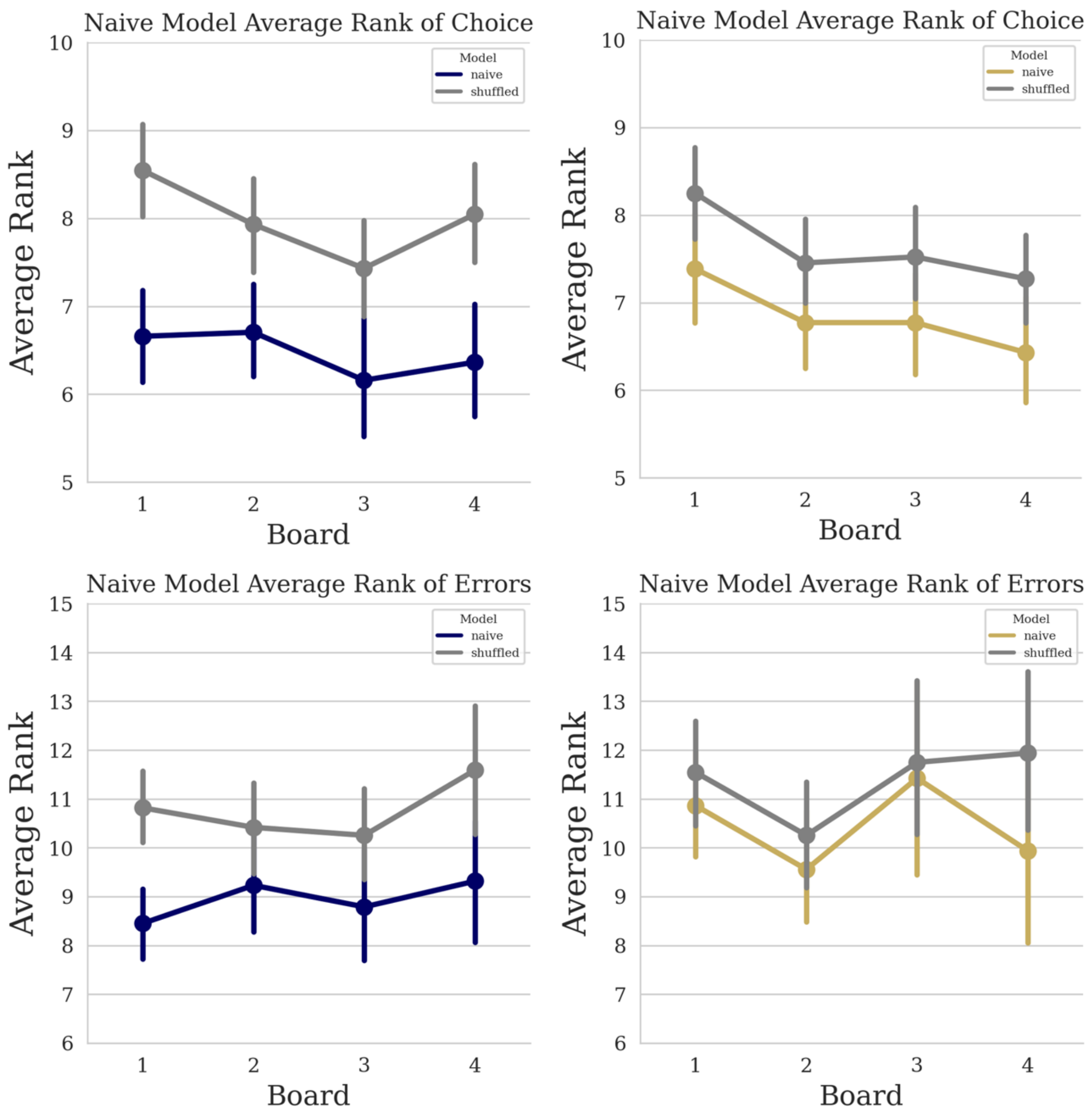

Figure 21: **Naïve Model Rank of Temporally Ordered vs Disordered Clicks Top Left:** Choices (hits and misses) by players ranked by the naive model. Blue shows real data, grey shows shuffled data. We plot the average rank (+/-sem) across all participants for each game round. **Top Right:** Choices by actors ranked by the naive model. Golden shows real data, grey shows shuffled data. **Bottom Left:** Errors only, by players ranked by the naive model. Blue shows real data, grey shows shuffled data. **Bottom Right:** Errors by actors ranked by the naive model. Golden shows real data, grey shows shuffled data.

A paired t-test between ranks of all shuffled versus real choices of players Figure 21 Top Left verifies the visible difference in rank ($t(43) = -11.435, p = <.001$) and the same advantage for real data is found in the same test applied to choices of actors, see Figure 21 Top Right: ($t(43) = -5.688, p = <.001$).

We now turn to evaluate errors. As visible in Figure 21 (Bottom Left), the rank value difference between real and shuffled data across boards is significant (paired t-test $t(43) = -7.792, p = <.001$) for errors of players. This is also still the case for errors of actors in Figure 21 (Bottom Right) where the shuffled (grey) data is ranked worse than the real errors ($t(43) = -2.721, p = 0.009$).

When we compare (Figure 21 blue line in Bottom Left to golden line in Bottom Right) the average rank of real errors by players to the average rank of real errors by actors (as assessed by the naive model), we find that the rank of actors is significantly higher than that of player: ($t(43) = 3.04, p = 0.004$). This shows that participants still made sensible (non-random) errors according to the state of the board, but less so than players.

The difference between actors and players is not found in the same analysis on choices (hits and misses: (Figure 21 blue line in Top Left to golden line in Top Right) , where a t-test finds that that actors are not significantly worse in rank than players: $t(43) = 1.555, p = 0.127$. We will return to this point in the discussion.

Reaction Time Divergence

We next assess whether participants re-enacted the reaction time divergence between firsts and seconds clicked during play (see Methods Section). In Figure 18 A we illustrate the measure. We extracted the median RT between all black cards that were the first of a symmetric pair (click number 8) and between all black cards that were the second of a pair (click number 9). We controlled for each person's average RT at each click made.

In chapter V, this measure reflected the epistemic difference between rule-followers versus those who had not discovered the rule. It is a measure of rule-followers reacting to the uncertain state of the board when the next pair is fully unrevealed, versus the deterministic certainty that is gained once one of the pairs is already revealed. How good are people at simulating the gain in certainty that derives from knowledge about symmetry in later rounds?

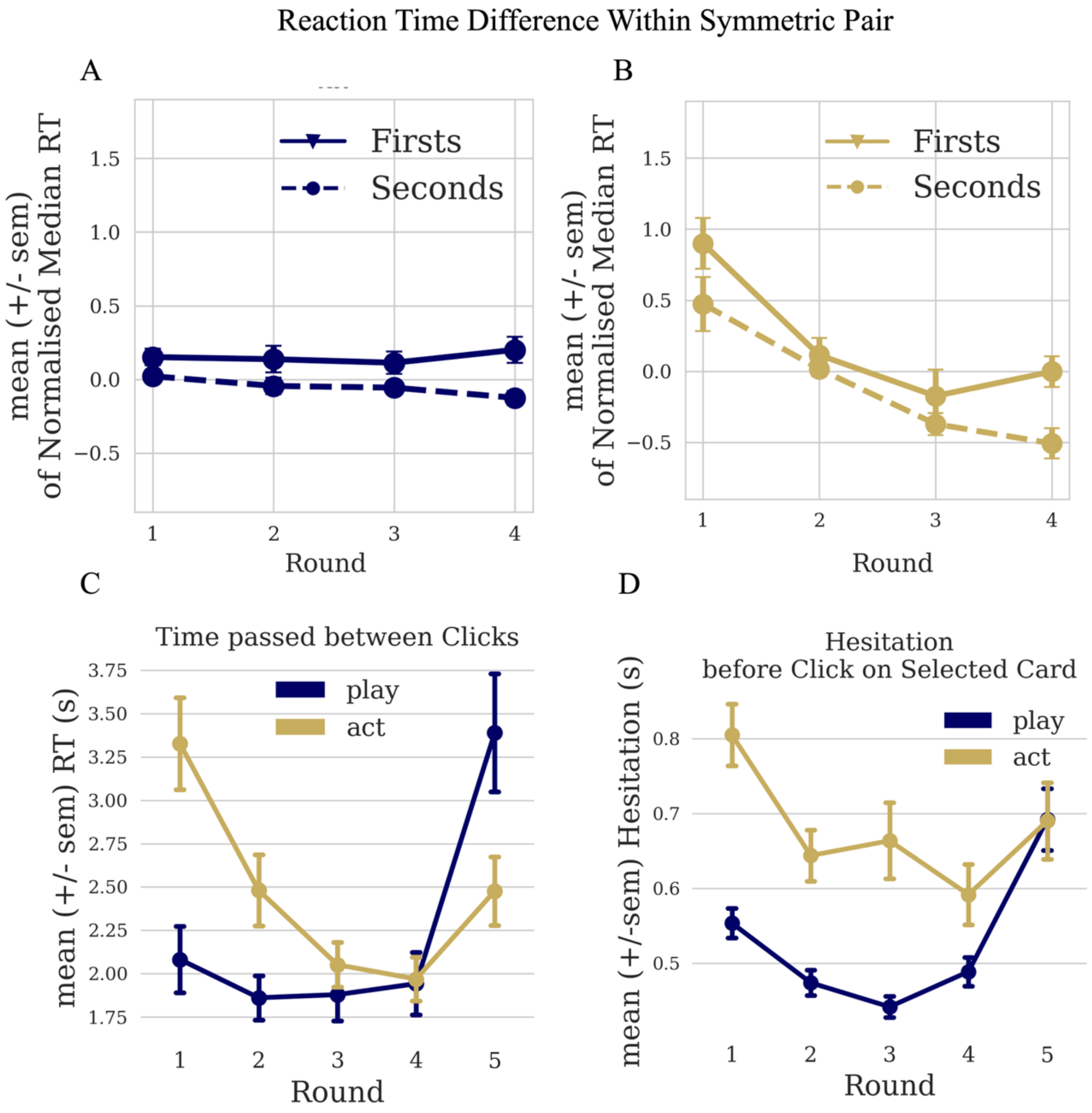

Figure 22: **Results Reaction Time Act versus Play Chapter VI A:** Play RT Divergence: Firsts mean (+/-sem) shown in full blue line. Seconds shown in dashed blue

line. RT of revealing the first of one symmetric pair versus that revealed the second of a pair. **B:** Act RT Divergence: Seconds mean(+/-sem) shown as dashed line. **C:** RT (time between consecutive clicks), mean(+/-sem) across population in play (blue) versus in act (gold) **D:** Hesitations in Play versus Act: the time passed between the moment the mouse cursor enters a card that is later clicked, to the moment of that click. Again, showing mean(+/-sem) per Play versus Act.

As we observe in Figure 22 A and B, RT of Firsts were always slower than Seconds (dashed lines), in both Play and Act. We ran a mixed effects model with condition {Play, Act} and knowledge state {0,1} as predictors of RT divergence (again, participants are added as random effects). RT divergence is calculated as the per person, per board median RT of all Firsts minus the median RT of all Seconds. We neither find a main effect of condition ($\beta = 0.1112$, SE = 0.1493, $z = 0.75$, 95% CI [-0.1813, 0.4038], $p = 0.456$) or knowledge state (0.0915, SE = 0.1488, $z = 0.61$, 95% CI [-0.2002, 0.3831], $p = 0.539$), nor an interaction effect (-0.0073, SE = 0.2108, $z = -0.03$, 95% CI [-0.4204, 0.4059], $p = 0.972$).

The difference values between firsts and seconds in the last board appear qualitatively larger for actor than players, but a paired t-test of those difference values between Play and Act showed the divergence between Firsts and Seconds in Act was not significantly greater than in Play ($t(42) = 1.0119$, $p = 0.3174$). We do confirm quantitatively that Firsts have higher RT values (on average across all rounds) than Seconds and that this is case in both Play ($t(42) = 3.7331$, $p = 0.0006$) and Act ($t(42) = 3.8142$, $p = 0.0005$).

We then looked at the dynamics of Firsts. We applied a linear mixed effects model to the median RT of all Firsts to compare them in Act versus Play (the two full lines in Figure 22 A,B). We looked at the interaction term between the categorical variable of condition

{Play,Act} and the numerical variable of round {1,2,3,4} to predict RT across cards that were revealed first.

We found a main effect of condition (β = 0.8382, SE = 0.2139, z = 3.92, 95% CI [0.419, 1.2574], p = <.001) and no main effect of game round (β = 0.0123, SE = 0.0552, z = 0.22, 95% CI [-0.0959, 0.1206], p = 0.823). But we found an interaction effect, showing that the RT across Firsts changes across rounds in Act relative to Play. The slope was significantly steeper and negative in Act relative to Play: β = -0.3117, SE = 0.0781, z = -3.99, 95% CI [-0.4647, -0.1586], p = <.001, owed to the high RT that actors showed at the start of the game (Figure 22 C).

We applied the same analysis to Seconds (comparing Figure 22 A blue dashed line against Figure 22 B golden dashed line) and found a main effect of condition (β = 0.6721, SE = 0.1541, z = 4.36, 95% CI [0.37, 0.9743], p = <.001), again no main effect of round (β = -0.0451, SE = 0.0396, z = -1.14, 95% CI [-0.1228, 0.0326], p = 0.255), but once again an interaction effect, showing that the RT across Seconds changes across rounds in Act relative to Play. The slope was significantly steeper and negative in Act relative to Play: β = -0.2875, SE = 0.0562, z = -5.11, 95% CI [-0.3976, -0.1773], p = <.001.

Reaction Time

In Figure 22 C, we show another measure of reaction times: each line shows how much time tended to pass between clicks (e.g. click 2 at t_2 minus click 1 at t_1). Here we consider all clicks, regardless of whether they were black or white, firsts or seconds of pairs. In Figure 18 A we illustrated click event by numbers written on the cards that were clicked. Our reaction time measure is the difference in time between consecutive clicks ([1, 2, 3,…8,9] in the illustration), averaged across them.

For Figure 22 C, we computed a per-person average (in seconds) over all clicks committed per round and in Figure 22 C we show the mean (+/- sem) across the population, separated by the condition in which participants played for real versus when they acted. Visibly, acting resulted in higher reaction times and a steep decline (increase in speed) as more rounds were played.

A paired t-test of per-person average RT values on the first four rounds, comparing playing versus acting shows this quantitatively: $t(42) = -4.225$, $p = 0.0001$. Actors were on average slower between clicks than players. Figure 22 C however also shows a marked increase in players reaction time on the half-revealed (5th) board. A paired t-test of average RT values on the 5th round, comparing playing versus acting shows that players were much slower between clicks on the 5th board than actors ($t(42) = 2.5952$, $p = 0.013$).

We selected the first four rounds and ran a mixed effects model in which we predict individual average RT values based on condition and knowledge state. We found a main effect condition ($\beta = 0.9322$, SE = 0.1226, $z = 7.6$, 95% CI [0.6919, 1.1725], $p = <.001$), no effect of knowledge state ($\beta = -0.0607$, SE = 0.1226, $z = -0.5$, 95% CI [-0.3011, 0.1796], $p = 0.620$), but an interaction effect ($\beta = -0.8327$, SE = 0.1734, $z = -4.8$, 95% CI [-1.1726, -0.4929], $p = <.001$), showing that in the first four rounds, participants' time spent between clicks did not change when they played for real. Actors however showed a drastic change: from taking longer in earlier rounds (more than a second longer on average than when they played for real), their reaction times sped up considerably up to the 4th round, in which they were as fast as when they played. On the 5th board however, they underestimated the reaction times of playing for real and completed the board with clicks that were on average a second closer to one another than during play.

This RT measure captures time spent on any activity within the time between one click and another. We now turn towards a particular behavioural component of the time spent between clicks, which we noticed by observing dynamic reconstructions and which we can measure using the mouse movement measure.

Hesitation

Recording of the mouse cursor position allowed us to analyse a second, more precise measure of participants' uncertainty. We consider a specific period of the time between clicks: the time between the moment (t) in which the mouse cursor has landed on the selected card and the timepoint (t_c) at which the click is committed to. Figure 18 A illustrates this measurement which we called a 'hesitation' in the bottom right corner of the grid, in which the participant committed to a click: the timepoint from the moment the mouse cursor is recorded to hover over the card to the moment when the click was made (here visible in the sharp movement of the cursor trajectory) is what we call a hesitation. We compute hesitations for all clicks.

Hesitations are a second indicator of uncertainty, specific to the uncertainty over *committing* to a particular hypothesis. In the previous section, we saw that reaction time dynamics by actors were faithful to the dynamics of playing for real. Are participants able to emulate hesitation lengths too?

Figure 22 D shows the average fraction of a second waited before all clicks made in each round, per condition. Actors were more hesitant before making a click. A per-person paired t-test between individual average hesitation in play versus act verified the difference ($t(43) = 3.3267$, $p = 0.0018$). Qualitatively, we observe a decline in hesitations in both real games and acting games and an uptick in hesitation during the half-revealed board, which is similar across playing and acting. A mixed effects model in which we use condition and rounds as predictors, finds a significant reduction in hesitation during play relative to acting ($\beta = -0.2881$, SE = 0.064, $z = -4.5$, 95% CI [-0.4134, -0.1627], $p = <.001$), a significant negative effect of board ($\beta = -0.062$, SE = 0.0165, $z = -3.76$, 95% CI [-0.0944, -0.0297], $p = <.001$), and an interaction term which approaches significance ($\beta = 0.0416$, SE = 0.0234, $z = 1.78$, 95% CI [-0.0042, 0.0874], $p = 0.075$).

Correct Hypotheses

We also use the mouse cursor recording to identify hypothesising behaviour, which we defined as considering a card by hovering over the card, but choosing not to click it at that time. When no click occurred, we consider it a rejected hypothesis which can either be falsely reject (the card underneath was black) or correctly rejected (the card underneath was white). Here we showcase results from an analysis of correct hypothesising across playing and acting (see Appendix B to see an analysis of false hypotheses).

Fundamentally, when acting, the actors need to ‘delete’ (from behaviour) the available information (cross hints) about the location of black cards and instead behave as if the search would occur while they are ignorant about the solutions. Can participants do this?

Consider first that participants on average made the same number of hypotheses in the real game and when they acted ($t(43) = -0.5945$, $p = 0.553$), showing that acting did not make them change the frequency with which they hypothesised.

In Figure 23 A we show the average proportion of correct hypotheses out all hypothesis events across participants in each condition. These are longer dwellings over black cards that do not result in a click. The figure shows that actors posit hypotheses over correct locations more often than players in the first three rounds.

As before, we ran a mixed effects model with condition and knowledge state as predictors, in this case predicting the ratio of correct hypotheses to all hypotheses.

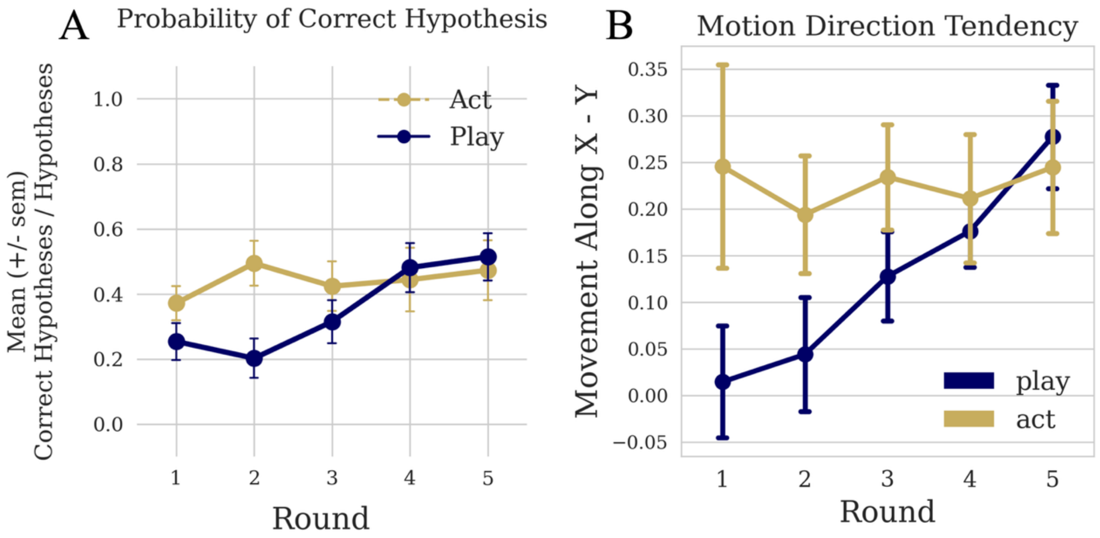

Figure 23: **Results Motion Analysis Chapter VI A:** Proportion of hypotheses (locations considered) that were correct in Play vs Act. Showing average across individual ratios per game round. **B:** Mouse cursor behaviour along the horizontal and vertical axis. Larger values on the y-axis indicate more left-right mouse movement. Figure shows mean (+/-sem) direction tendency across participants total absolute horizontal minus vertical mouse movement in round.

This yields a main effect of condition, showing that during Act, participants are significantly more *tempted* to click on unrevealed black cards: $\beta = 0.1953$, SE = 0.0624, $z = 3.13$, 95% CI [0.0729, 0.3176], $p = 0.002$. We also found a main effect of knowledge state ($\beta = 0.1555$, SE = 0.0641, $z = 2.42$, 95% CI [0.0298, 0.2812], $p = 0.015$), but no interaction effect ($\beta = -0.1424$, SE = 0.0942, $z = -1.51$, 95% CI [-0.3271, 0.0422], $p = 0.131$). We then also ran a mixed model on the increase in hovers over black cards during play only, using round as predictor to show that players indeed shifted their hypotheses towards correct locations when they learned to apply the symmetry rule: $\beta = 0.1589$, SE = 0.0645, $z = 2.46$, 95% CI [0.0325, 0.2852], $p = 0.014$.

Actors thus showed the same rate of hypothesising behaviour, but the distribution over where those hypotheses landed was skewed by the knowledge, they had available to them, and they were unable to prevent this shift from occurring even when they pretended to be ignorant.

Motion Direction

We additionally observed a systematic failure to emulate the ignorant knowledge state in an analysis which considered the directional tendency of mouse movements. In chapter V, we observed that rule-followers tended to show more horizontal movement than vertical movement as the game progressed. In Figure 23 B we now show the difference between total horizontal and vertical distance traversed by the cursor across a game, per person and then averaged across participants for each condition.

A mixed effects model shows a significant increase in left-right motion for players (as predicted by round, participant as random effect): $\beta = 0.0568$, SE = 0.0177, $z = 3.21$, 95% CI [0.0222, 0.0915], $p = 0.001$. It does not show this for actors: $\beta = -0.0063$, SE = 0.0246, $z = -0.25$, 95% CI [-0.0545, 0.042], $p = 0.799$. We used a mixed effects model using both condition and knowledge state as predictors (modelling participants as random effects) and found a main effect of acting ($\beta = 0.1905$, SE = 0.0516, $z = 3.69$, 95% CI [0.0893, 0.2916], $p = <.001$), a main effect of knowledge state ($\beta = 0.1226$, SE = 0.0516, $z = 2.37$, 95% CI [0.0214, 0.2238], $p = 0.018$) and no interaction effect ($\beta = -0.1196$, SE = 0.073, $z = -1.64$, 95% CI [-0.2627, 0.0234], $p = 0.101$).[7]

---

[7] An ANOVA two factors condition and knowledge state showed a significant interaction (F(1,43)= 4.28, p = 0.045), a main effect of condition (F(1,43)= 7.01, p = 0.011) and no effect of knowledge state (F(1,43)= 2.02, p = 0.163).

Discussion

People can re-enact behavioural markers of discovery, suggesting that they have partial access to, and inference capacity over, how discovery of a rule affects their actions.

Model fits to all choices (revealing black and white cards) showed that actors were able to obtain the same qualitative dynamic description of their choices as reflected in model rank (Figure 20 Top Row). Actors however showed a steeper decline in rank as assessed by the symmetric model, which indicates that they made choices in accordance with the symmetry rule earlier than they did during real games and that they adhere to the symmetry rule more closely in later rounds. This could be a sign of overacting, but also a reflection of not being able to ignore their true knowledge state.

We were unable to replicate the previously found decline in rank of errors during play as assessed by the symmetric model. This could be because we used a selection mechanism of discovery which was independent of performance on the last board. Note that making a discovery and being able to name the rule does not necessitate that participants use this information to adapt their choices, even if they are incentivised to do so. This analysis should be repeated in replication studies.

We showed an analysis in which we compared the rank of the naive model of real data against the rank of shuffled (across time) data (see Figure 21). This shuffled data thus maintains the same identity of selections made but has no temporal structure. In errors made by actors, the shuffled data was ranked nearly as low as the true order of errors made. The difference was still significant, showing that actors did not behave randomly, even when they made errors. The errors made by actors were more sensible than if they did not consider the arrangement of the board at that time at all. They attempted to make errors that made sense according to the state of the board, the rules of the game and the knowledge state they were meant to emulate, but they were worse at doing this in their errors than in their hits.

We did not observe this difference between choices and errors during play: the real decisions were always ranked markedly lower than the shuffled decisions, showing that the errors that players made, were similarly sensible according to the rules of the game and the likelihood distribution of black cards, without the assumption of symmetry. Players' errors were adjusted to the state of the board in a way that actors could not quite re-instantiate: their errors were ranked higher than the errors of players. This is an interesting result, because it shows that actors could model how the state of the board would have helped them make successful (black) choices, but that they could not model a false state of belief that depended on both ignorance over the symmetry rule and the state of the board. In other words, actors noticed when the state of the board could give rise to true hypotheses, but not how it might have made them accept false hypotheses (click on white) when they played in ignorance.[8]

We found that (correct) hypothesising behaviour was higher in the early rounds of Act relative to Play, showing that participants were tempted to click (but able to supress clicking) on black cards when they were shown where black cards were located via cross hints, and/or that they were unable to adjust their search and hypothesising behaviour to a state of epistemic ignorance. Their correct 'hunches' of where the black cards were located, were consistently at the frequency level which rule-followers only ended up with during play once they fully applied the rule in round 4 and 5. This shows that even when actors are deliberately committed to making errors, their mouse movement still reveals their true, rather than acted state of knowledge, despite them being made aware that the mouse movement would also count towards their bonus and that it had to emulate, as close as possible, their previous movements.

---

[8] Or, how they might have formed beliefs under a different assumption (e.g. symmetry along the y-axis).

A similar inability to 'cover up' their true epistemic state was found in higher horizontal mouse movements in early rounds of acting compared to playing. Participants were not aware or unable to replicate the association between an ignorant epistemic state and its effect on subtle behavioural markers in the mouse movement. These two measures suggest that continuous mouse movement is a more difficult behavioural dimension to access metacognitively or control counterfactually.

Actors' reaction times from one click to the next started out much slower in round one and ended up faster than players in later rounds. Their average reaction time across all click types was significantly higher and hesitation before a click, once hovering over the location, was also higher during acting, probably reflecting the generally high cognitive demand of the task. Actors had to reconsider whether they are 'allowed' to make a specific click, given the knowledge state they had at the time. Hesitations also followed a decline over rounds, suggesting that actors had an easier time once they were allowing themselves to make clicks according to the symmetry rule. Specifically in reaction time between clicks they ended up matching their RT during play in round 4, but significantly underestimated RT in round 5. Actors on the half-revealed board imitated the reaction time (but not the hesitations) to be much faster than they actually did during play, potentially suggesting that they did not adequately model the search behaviour on that board and overestimated how easy it was to follow the rule on that board, even if one was knowledgeable about the symmetry rule.

Interestingly, the difference in median, click-number corrected RT that we found during play (in the study presented in chapter V as well as players shown here) was maintained as participants acted. There were able to reflect the epistemic gain in certainty that was derived from rule discovery. This shows an awareness of how rule-ignorance versus knowledge affects action certainty and an ability to re-instate this association between rules and reaction times.

We observed an underestimation of the number of errors when participants acted. It is possible that this was due to a real bias or actors wanting to complete the game. Acting behaviour is more complex behaviour than is typically required in online games and participants responses in debriefs suggest they found it difficult, albeit enjoyable, to play this game. Even though we monetarily incentivised compliance with accurate acting attempts, it is possible that participants in later rounds used their knowledge of where the black cards are hidden, because they wanted to complete the game. This would explain the underestimation of errors. We suggest a future variant of the game should ensure that completion incentives are less likely to interfere with acting precision, to test if underestimation of errors throughout the game persist.

This design provides a rich dataset, and many analyses remain to be explored. Some of the analyses provided new ideas about how to make useful amendments to the design. In Appendix B we show two additional analyses showing that participants appeared to make different types of errors across conditions. The analysis shown in Figure 25 E considered changes in *avoidable* errors. Avoidable errors are all symmetric twins of already revealed white cards (see cards framed in red Figure 16 Midde Column). A rule-follower who behaves optimally knows that these locations cannot contain a black card. This is a promising analysis to pursue because actors who are able to inhabit the state of mind of an ignorant player should not show a bias to avoid these locations and, consequently, have the same probability of revealing avoidable white cards as ignorant players. A failure to re-instantiate the same rate of avoidable false errors (and false hypotheses, see Figure 25 F), would indicate a failure of inhabiting the ignorant knowledge state.

While we observed stark differences in both avoidable errors and hypotheses, suggesting that actors may not be able to recover the errors distribution of an ignorant state, the design of our experiments does not control the number of errors and we thereby could not de-correlate the higher error rate (which would explain a higher rate of avoidable errors even if errors were randomly distributed) during play, from the result shown in Appendix B. An amended design, in which actors are for instance told about the exact number of errors which they committed during play, would allow us to examine if actors avoid certain error types, thereby revealing a metacognitive insensitivity to how epistemic states translate to distributions over error types.

So far, we have only shown analyses which look at the within-subject differences to understand which dimensions tend to be easier and harder for participants to re-instantiate when they pretended to play the game. We can distinguish between e.g. coarse measurement types (reduction in clicks) which participants act out well and subtle measures (e.g. the mouse movement) which actors found harder to approximate. But we have not yet shown that these behaviours are truly *self*-simulations (even though we asked participants to emulate their own behaviour) or whether they might be forward models of an average mind or other minds (Gopnik, 1993). This requires analyses that show that participants more accurately re-enact their own, rather than another participant's behaviour, or that models which are fit on individual data best describe participant-matched held-out data.

The re-enactment method has a few blind spots. For instance, we cannot distinguish between: a behaviour is not known by the participant as being part of the discovery process (fully inaccessible), a behaviour that is known but not thought to be part of the discovery process (accessible, but rejected) or a behaviour that is known, part of their model of the discovery process, but participants fail to retrieve the behaviour or re-instantiate it faithfully. Similarities across play and act might also derive from common factors in the task which the participant responds to in the same way in both studies. A failure to show a behaviour while acting, also does not mean that the cognitive process which underpins the behaviour is in general not subject to metacognitive modulation. We of course also cannot directly measure what cognitive processes participants retrieve when they act. In principle, it is possible that the same behaviour (b) is derived from a different process if participants play versus act. The method of re-enactment has shown promise but is best combined with other methods and designs which allow experimenters to address some of the above critiques.

Nevertheless, we believe re-enactment has a future as a method of inquiry. Future studies could for instance, zoom in on the granularity with which participants can emulate different types of learning, such as discontinuous insight moments (which we do not measure here directly, even if it is possible to have an insight in our task) versus gradual learning. Of particular interest might be whether re-enactment with feedback can lead to measurable and generalisable improvements in metacognition.

The analyses we showcase here, especially those involving mouse movement, were inspired by, developed and verified with the dynamic reconstruction approach. We want to suggest that studies with multi-dimensional and high-resolution data and underexplored questions such as those in discovery or metacognition, will benefit from adding yet another meta-loop (thinking about thinking about thinking), through which researchers can apply their full inference abilities and hopefully find joy in exercising them (Chu et al., 2024).

Overall, we find people were able to emulate a transition of error-prone ignorance to rule-informed behaviour, and we found an awareness about how the discovery of a rule translates into errors and reaction times. As seen in the continuous measure of mouse movements, a return to ignorant behaviour whilst being knowledgeable, was difficult. Participants could not quite act like they did when they knew less. We provide evidence suggesting that people may have cognitive limitations in stepping into their former epistemic selves.

Chapter VII General Summary & Discussion

My aim in writing this thesis was to contribute to the scientific understanding of how human cognition is shaped by the interplay between actions and memories. This chapter will summarise the extent of my accomplishments and their limitations.

Prima facie, actions and memories appear intertwined. Memories guide what action we take, which in turn affects what memories we make. This interaction has both personal and philosophical relevance. Memories of rewarding or painful experiences are essential in shaping what future experiences we choose to have. Our experiences, and our memories thereof, shape our understanding of ourselves, such as personal preferences or identity. Memories affect our actions and our model of ourselves, but our model of ourselves and our actions also affect the future memories we are likely to make.

This thesis is my attempt to narrow this intuition into a set of precise scientific hypotheses and experiments which could attest to effects of *actions in memory* formation as well as to effects of reconstructive *memory in action* sequences during self-simulation. This work thereby intentionally spans two distinct approaches and perspectives on the interplay between actions and memories.

I first presented a series of experiments (chapter II-IV), which suggest that a measurable and generalisable effect of actions on memory can, surprisingly, already be found at a fundamental level of encoding. Action selections appear to reinforce the selected item in memory. I then presented (chapter V, VI) a paradigm in which participants scaffold their actions on memories of past experiences during a self-simulation task. I now first summarise and then criticise the work on action in memory, before summarising and noting the limitations of the work on memory in action.

Summary Action in Memory

In chapter II, I introduced the action alignment effect: in an object-in-place associative memory task, participants benefitted from a training curriculum in which their action selections were aligned with the feedback locations, compared to a training curriculum in which their actions selected a distractor location that was incorrect and thereby misaligned with the supervision signal. In test trials, during which participants recalled what location is associated with the visual test stimulus, participants were more accurate after being exposed to an action aligned curriculum. The curriculum influenced what type of errors participants made during their test phases: action misalignment led participants to confuse the correct location with the incorrect location which they had selected during training. Interestingly, in the action misaligned condition, participants were also significantly more tempted to select the correct location during training (as shown in the direction and efficiency of their navigation paths) once they had learned the correct locations, despite receiving no reward for this. This finding suggests that participants may have had some metacognitive awareness of the benefits of action selection on memory encoding. The design also tested whether action selections are learned and transferred to test trials separately from the object-location associations, which could have explained why action alignment (where the actions at training are the same as the actions at test) confers a benefit over action misalignment (where the actions at training are not the same as the actions at test) *if* the action sequences across training and test were identical. Actions across training and test in the experimental design were however not identical, showing that the action sequence itself is not likely to be the origin of the action alignment benefit.

In chapter III, I listed several experiments which replicated the action alignment effect and deepened our understanding of it by controlling for an effect of attention. Several baseline conditions allowed for a comparison of participants' memory performance at test after training with action aligned, action misaligned compared to observation conditions. Action alignment conferred a benefit relative to a condition in which participants took an unrelated action and action alignment showed no more confusion errors than this baseline condition, which clarified that the distractor location was fully ignored as long as actions were aligned.

In a following experiment, a mere observation of the correct and distractor locations (visually highlighted as per condition of alignment or misalignment) produced a reduced difference. When a selection of a location occurred via attentional selection, the difference between alignment and misalignment was smaller compared to when the target location was selected with an additional action sequence. Conversely, attention manipulation experiments produced no modulation of the action alignment effect. No significant modulation of accuracy (i.e. the action alignment effect) was found by modulating the attention drawn to the distractor location. In the rate of confusion errors, these experiments expectedly found that attention plays some role in the effect, but they produced no evidence that attention was the main culprit underlying what I proposed to be an effect of *action* alignment.

In chapter IV, I tested the generalisability and applicability of the action alignment effect. It was found that, surprisingly, at least two experimental settings appeared to show an advantage of the action alignment effect over the test-enhanced retrieval effect, which suggests that there may be learning context in which action alignment might be on par with, or preferable to, the classic testing effect, which has been introduced into schooling curricula and is part of an accepted canon of knowledge of how people learn best. Experiments demonstrated that action alignment can be used to learn a subset of memory items by preference, suggesting that it can be used to sequence, chunk and structure learning, which may be useful and necessary for more complex and advanced learning tasks. Finally, I related the action alignment effect to the self-reference effect and extended its potential utility by finding that the effect holds even if the action sequence is *not* executed by the learner themselves. This result opens the possibility of applying the action alignment effect in settings in which teachers need to showcase or execute the actions instead of the students themselves.

Limitations Action in Memory

The main limitation of this work on the action alignment effect is that this thesis does not propose a mechanistic account of why the effect might hold. What is the neurocomputational basis of memory modulation through the alignment between action and feedback? This question is particularly interesting, because the action alignment effect holds under minimal motor actions (such as a mouse click) and the mere observation of a virtual game partner who takes navigation actions. Motor *execution* itself does not appear to underly the action alignment effect. Instead, a cognitive representation of an action appears to suffice. My primary interest was to demonstrate the action alignment effect and rule out alternative explanations behaviourally as far as possible, and to place the effect in relation to similar effects in the literature, rather than to propose a mechanistic account which might prematurely narrow the range of plausible frameworks and theories. Future work should continue to test the effect behaviourally but might also begin to capture hypotheses about plausible mechanistic accounts in computational models.

Future experiments should aim to further pinpoint the origin of the effect. What contribution to the effect is attributable to reward or to the success of the action? In the task presented here, a selection of the correct location was not rewarded, and the selection of an incorrect location was not punished. However, training trials in which the feedback appeared in the location that was selected by the participant were also the correct location and therefore the success of an action was directly tied to the feedback signal. This did not allow for a teasing apart of whether the effect stems from an internally generated sense of reward or a representation of success by the participant, or whether it stems from the action being aligned with feedback. I chose to describe the effect as stemming from an action being aligned with feedback, because the studies presented here provide no access to whether participants perceived the action aligned condition as rewarding or as an action success. Describing the effect as an effect of (mis)alignment between action and supervision, was thus the most direct and surface-level description of the difference between the two conditions. This description is therefore preferable in the absence of a neurobiological explanation which may yet rule out this interpretation in the future.

Summary Memory in Action

In chapter V, I first presented a new task environment for the study of how humans discover spatial rules. I introduced a variety of measures, both in choices and subtle mouse movements, which reflected a cognitive transition between two epistemic states (pre-discovery and post-discovery), and showed that by use of correlations between performance and behavioural markers it was possible to distinguish between participants who discovered the spatial rule and those who did not. Participants who discovered the rule made fewer errors over time and were (over time) better described by a model that took the rule constraints into account, compared to a model that did not. Participants who discovered the rule showed a distinct pattern of reaction times that were matched to the spatial arrangement of the solution and also adjusted their tendency to make horizontal mouse movements in accordance with this arrangement after learning the rule. These behavioural markers were not found in the subset of participants who did not discover the rule. This task was introduced as the basis of studying metacognitive access to spatial rule discovery in the subsequent chapter.

The chapter showcased and discussed the method of dynamic reconstruction that was used across chapter V and VI and made the claim that experimenters' inference over reconstructions of participants game play can aid the generation of hypothesis and analysis pipelines.

In chapter VI, I presented a study of how peoples use their memory of their own actions in an effort to model their own behaviours associated with counterfactual cognitive states. I introduced an ambitious experimental pipeline, in which participants were asked to re-run the cognitive processes that underpinned previously experienced epistemic states. The experiment used the pre-discovery and post-discovery states experienced in the task described in chapter V. Participants who did discover the rule while playing the game, were reinvited after their memory had faded, to re-enact their exact behaviour in terms of clicks, reaction times and mouse movements, despite now knowing the solution to the game. In essence, this experiment was aimed at a condition in which participants run a model of themselves forward and imagine how they would act (or would have acted) if they did not know what they in fact already knew. The assumption was that participants' compressed memory model of how they discovered the rule when they first played the game (and how they transitioned between epistemic states), was used to simulate, or construct a replication of, their own counterfactual epistemic state. The experiment was constructed to elicit a self-model in respect to cognitive discovery, in pursuit of the question of what behavioural markers are and are not part of participants' self-model or metacognitive access.

Participants were surprisingly good at re-enacting the transition between ignorance and knowledge in terms of error rate decline but also showed a consistent overestimate of their own ability to complete each round with fewer errors than they made when they played for the first time. Both the selection of errors and successful choices were over time better described by a model that takes the rule into account, showing that participants dynamically and correctly adjust their actions in accordance with a new epistemic state. The data also interestingly revealed systematic inabilities, particularly in emulating the ignorant epistemic state, by actors that are knowledgeable. Analyses found that actors had greater difficulty in replicating the temporal order of errors, slower reaction times in early rounds and an underestimation of reaction times on the last round (which could be optimally solved if the rule was known), as well as an overestimation of hesitation before committing to a hypothesis in all rounds but the last. An analysis showed shifts in the distribution of correct hypotheses across play and act, as well as actors' inability to model the ignorant state in terms of directional tendencies in mouse movements.

Limitations Memory in Action

The advantages of the method of re-enactment, in the way it was implemented here, are also its complications. The experiment was aimed at recording and quantifying self-simulation in a way that tended towards a naturalistic replica of how self-simulation occurs outside the lab. Actors on a stage or spies at court must *act as if* across all behavioural dimensions: their choices, facial expressions and subtle bodily movements must match those of someone who they are not or of someone who is ignorant of some secret which they actually know. The experimental task demanded something quite similar, by requesting that participants emulate a complex sequence of behaviours of first being ignorant, making a discovery and then being knowledgeable. Each of these states, as well as the dynamic and time-dependent transition between them, was associated with coarse (e.g. choices of cards to reveal) and fine-grained, subtle (mouse movement) behavioural markers that reflect the cognitive state. The experiment demanded that participants run a complex and unusual cognitive computation.

This can make it harder for experimenters to tease apart what failures in self-simulation are due to exhaustion or the difficulty of the task (e.g. keeping track of one's own mouse movement while thinking about which error to commit) as opposed to an incapacity to self-simulate along a particular behavioural dimension (e.g. as discussed earlier in relation to the number of errors made). Would the bias towards fewer errors by actors persist if the number of errors was the only behaviour the actor was required to self-simulate? One limitation of the experiment was thus its level of ambition, but usefully, a complex experiment such as this, records a range of behaviours that the experimenters cannot foresee. It may thus be harder to conclude definitives from this data, but this data provides fertile ground to generate hypotheses about which exact behaviours are subject to distortions during self-simulations, which can then inform more narrow and specific tests of those exact behaviours during re-enactment. Follow-up experiments should narrow the range of behaviours that the participant needs to keep track of and investigate each behavioural dimension at a time. For instance, future experiments could focus on participants emulating only the ignorant, pre-discovery knowledge state and investigate whether the distribution over the type of errors committed, when the number of errors is kept constant, are shifted between playing and acting.

Conclusion

Memories and actions are closely intertwined in human cognition. I have empirically shown that the cognitive representation of an action sequence can modulate the strength of associative memory encoding. I have shown an experimental approach to exploring how memories of cognitive states shape our metacognitive models of self and guide action sequences during self-simulation. I thereby offer this thesis as a contribution, by way of novel methods and results, to the academic study of cognitive science and to all who learn.

Appendix A Action in Memory

Confirmatory Results Chapter III & IV

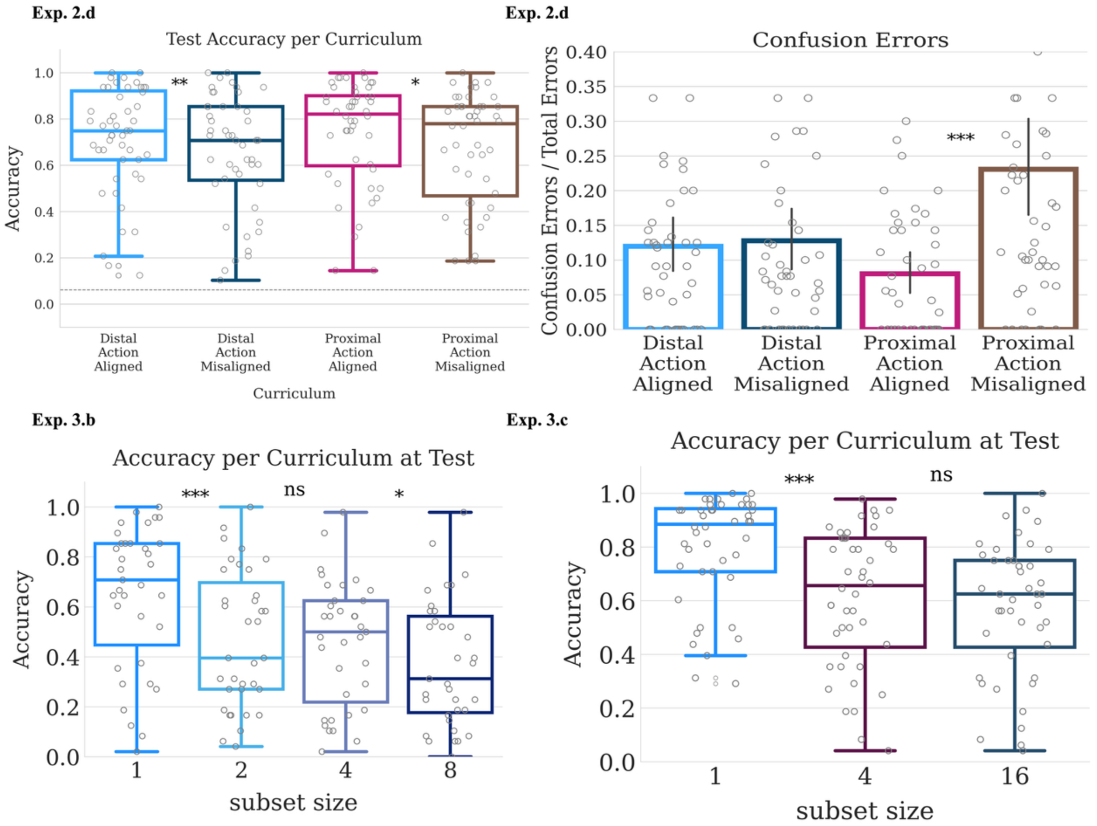

Figure 24: **Confirmatory Results from Studies Chapter III & IV Exp.2.d** Test accuracy per condition which are combinations of {distal, proximal} and {aligned, misaligned}. Each datapoint represents the average accuracy across all test trials per person. Coloured boxplots show population summary statistic. **Exp.2.d:** Ratio of confusion errors at test per condition. Probability of erroneously selecting the other square that was highlighted in a training trial during a test trial trying to recall that same stimulus. **Exp.3.b**: Test accuracy after training with k highlighted options. **Exp.3.c**: Test accuracy after training with k highlighted options.

Minor Design Variations of Studies in Chapter II, III, IV

We implemented minor variations in instructions and stimulus and feedback presentations across the studies to ensure the robustness of our result. Since all results were robust against those variations, we detail them here rather than in the main section. We varied the feedback and response presentation to ensure that our result does not only arise due to a particular supervision setting.

In Exp.3.a-c a green square appeared in the correct location after 900 milliseconds after the target was selected. The stimulus remained in place. After 1.5 seconds, the next trials began. In all other experiments, the stimulus instead moved into the correct location, framed in green. After the target was selected, the remaining grey square and the avatar remains on screen, immovable in the location of the target. The yellow frame disappeared. After 900ms, the stimulus shrunk into the correct locations (covering either the grey square or avatar) and is seen with a green frame for 1.3 seconds. All features on the response grid disappear and two grey squares reappear after a 500 ms delay. This delay was introduced to ensure that the change in squares was not interpreted as a movement of grey squares.

In Exp1. a, b and c, grid locations were divided into 8 pairs of 2 diagonally positioned squares *a* and *b*, which were a Manhattan distance of exactly 4 apart. These pairs remained constant. We did this because findings from our previous studies (not reported here) found that highlighted squares arranged in columns (or rows) facilitated chunking of highlighted locations and confusions between locations, which may have added confounded effects in our studies of the action alignment effect. We therefore chose diagonal pairs, which we had found no chunking effects for.

In one range of experiments (Exp.1.a, Exp.2.b,c,d and Exp.3.b,c,e) we kept the same pair of grey squares highlighted for two consecutive trials. Exp.3.b and c this was adjusted to fit the size of the subset highlighted. A condition where k = 8 resulted in two subsets, whereas when k=4, there are four subsets etc. Each subset of size k would stay selected for k trials in a row, sampling all members of the set at random without replacement. This made it easier to predict the correct location in the misaligned condition, since if *a* was the correction location in trial t, it was predictable that *b* was the correct location in trial t+1. It however also increased the risk of confusion between distractor and target, which may influence the error rate at test. In all remaining experiments, we therefore presented a new pair of highlighted locations on every trial, thereby reducing the risk of facilitating confusions between the target and distractor. Our results were consistent across all studies, where we found the same main effect of action alignment in error rate and error type.

Exp.3.b and c included the variation that training trials in the first block were repeated once after 4 trials. This led to 64 instead of 48 training trials in each condition. This made no difference to our main effect, which expectedly simply resulted in slightly better performance across conditions, but which did not affect the comparison between conditions that we were interested in. The results are also in line with Exp.3.a during which participants were exposed to 48 training trials as usual.

At test, the blue square reappeared in the clicked location and remained on screen until the next trial. This simply indicated to participants the choice they had made. In Exp.3.b and Exp.3.e, a yellow cross instead appeared at the exact location they had clicked. In Exp.2.c the response was visualised by a yellow frame around the clicked square. None of these variations noticeably influenced the effect of interest.

Appendix B Memory in Action

Pre-Registered Hypotheses for Chapter V

The description of the task which we included in the pre-registration is omitted here, but can be found in Cremer & Mazor (2025a), from which we copy the following:

Motivation

This experiment is aimed at capturing rich behavioural markers of rule discovery. It will be used to study how hypotheses about rules come to mind and how they are processed (Bear et al., 2020).

Participants & Procedure

The research complies with all relevant ethical regulations and was approved by University of Oxford Central University Research Ethics Committee (code R90986/RE002). Participants will be recruited via Prolific and will give informed consent prior to their participation.

To be eligible to take part in this study, participants' Prolific approval rate will need to be 90% or higher and their reported first language must be English. Participants are instructed that they can only enter the game if they're not using an ipad/phone (this ensures we will have data about the cursor behaviour from all participants).

We will collect data until we reach 100 participants. The median time to complete the experiment is 10 minutes. Participants will be paid an hourly wage of £9 (and up to +$1 bonus if their game is within the top 30% of participants).

Data analysis

We will analyse reaction times, number of clicks, position of clicks, error patterns and dynamic cursor movements. The location of the mouse cursor will be recorded at all times to later reconstruct a dynamic simulation of participants' behaviour throughout the game (Mazor, Firestone, et al., 2024).

We will analyse commonly-used indicators of cognitive processes (such as accuracy and reaction times) both to characterise broad behavioural patterns across the sample, and in order to distinguish participants who discover the spatial rule from those who do not. We will then correlate these measures with quantitative descriptions of the cursor dynamics.

Performance will be measured by the number of clicks per person per board. Based on pilot data we expect that around 40% of participants will discover and apply the rule, as reflected by near perfect performance on the last board. The last board allows us to identify and compare rule-followers against participants who do not end up learning the rule. We will define participants who discovered the rule as those who completed the 7th board with not more than 7+1 clicks.

Hypotheses

H1) We will test the hypothesis that naive participants will be best described by a naive model at first and that a symmetric model will overtake the naive model for participants who we classify as those who discover the rule. Each model counts all arrangements of black cards that could be the true arrangement, given the number and position of black cards that are already revealed. The symmetric model thus has much tighter constraints on which locations could plausibly hide black cards. The set of possible arrangements on the board is much smaller than the set considered by the naive model. After each click (on which new information is gained about one location), each model updates the set of viable arrangements by excluding all arrangements that are inconsistent with the newly gained information. Each location on the grid is then ranked according to the likelihood of hosting a black card, which predicts the likelihood of choosing that location according to the model. We will test the hypothesis that this is not a merely a reflection of performance.

H2) We will test the hypothesis that knowledge about the rule will be reflected in participant's reaction times in that participants will take longer to uncover the first of a symmetric pair of black cards but will be faster to uncover the second of the pair. We will control for each person's average RT at each click made, since RT for later clicks is likely smaller. We will extract the difference in median RT between all black cards that were the first of a pair vs the second of a pair. We will test the hypothesis that the RT difference increases as a function of game number (1-6).

H2a) We will test the hypothesis that the beta coefficients, obtained by a regression of the RT difference against the board number, are significantly and negatively correlated with the number of clicks on the 7th board.

H2b) We test the hypothesis that the RT difference on the 6th board is negatively correlated with the number of clicks made on the last board, showing that participants who know the rule are more likely to have faster responses for the symmetric twin of an already uncovered black card. In addition to the number, locations, and timings of clicks, we will record and analyse the movement of the participants' cursor throughout the game.

H3) We will test the hypothesis that players show more horizontal movement relative to vertical movement over time, reflecting the likelihood to click on the mirror-symmetric twin of a newly discovered black card as well as a search and compare pattern that identifies symmetric locations along the x-axis rather than the y-axis.

H3a) We hypothesise that this should be the case among participants who discovered the rule more than among participants who did not discover the rule.

Rejection criteria

Participants will be excluded if: they cannot report a completion code, they have missing data, they used an ipad rather than a desktop and mouse

Sample size justification

The sample size is based on pilot experiments. We aim to have a dataset that contains both participants that discovered the rule as well as those that do not.

Debrief Responses Chapter V

Debrief responses of n = 37 who were rule-followers on half-revealed board:

- Symmetry
- After a few boards, I noticed that the pattern was symmetrical around the vertical centre line of the board
- Smiley face?
- Symmetry
- Symmetrical in vertical axis but I don’t know what else
- yyes there was a pattern, it looked like a skull or a heart or the helmet of Samus Aran from Super Metroid
- it was mirrored
- yes shapes
- Yes, the patterns were reflected across the board. So on one end of the block, a set of black cards were arranged the same way as another set at the opposite end of the block.
- It’s mirrored across the vertical centre line
- mirror symmytry
- The black blocks were symetrical
- yes they were symmetrical, i was not expectng that
- The patter was semetrical
- The patterns were symmetrical from one another on each board.
- Yes, the board was symmetrical
- they were symmetrical
- Horn Pattern?
- Yes, they were in mirror image
- I don’t know the name of the pattern, They seemed to be a mirror image. I clicked on around board .
- It seemed almost symmetrical as if mirroring the pattern from the opposite side.
- They were symmetrical
- They are a mirror image , right & left
- they were symmetrical along the vertical axis
- the cards were mirrored
- a skull?
- Yes, but I can’t really name the pattern.
- symmetrical
- Yes it was a mirror image of the one side
- They’re mirrored
- I think it was mirrored
- I noticed the black squares were arranged symmetrically by a vertical axis.
- Symmetrical
- No I didn’t notice the pattern although I realised there was some kind of symmetry
- The squares were left-right symmetrical.
- They were symmetrical patterns, as each half of the board mirrored each other. If a square was in a different spot on one side it was easy to pick the same equivalent if mirrored on the other, and for other less obvious ones, testing to see if it was white or not first then matching that on the other side afterwards.
- Debrief responses of participants who did not have less than 2 errors on the half-revealed board: mask pattern
- yes they were arranged in triangular patterns,rectangular patterns etc.
- Some of them had symetary
- no
- They were arranged in T-SHAPE and SQUARE pattern

- i did not notice a pattern
- I definitely noticed a pattern, not certain what the name is though…
- no
- some of them appeared to be reflections of the other side of the board
- yes symetrical
- no
- hexagon
- no
- symmtery
- Yes, some were in blocks of four’s and three’s
- L PATTERN
- symetrical
- the black squares were arranged in an L SHAPE
- i did not note the parten of the cards
- Some of them looked like they were mirrored on eachother but without all of it being revealed I found it difficult to find a pattern in most of them until the end.
- No, I didn’t notice any pattern.
- Yes they were mirrored but it took me way to long to figure out
- to be honest it was somehow difficult for me to notice
- No
- a square pattern
- No
- blocks of 2 and 4 joined together
- symmetrical for the most part
- I didn’t nottice a pattern
- I can name few patterns
- no
- I did not notice a pattern.
- I didnt notice a pattern
- zigzag
- No I did not.
- It seemed to be the letter “T”- although I am not sure.
- No, I didn’t notice any pattern at all. To me they look a little bit random from time to time.
- A Tetris pattern
- brick
- Looked like an individual was lifting weights.
- L- shape and square
- Several of the board had symmetrical patterns
- unable to identify
- Yes.I did notice a pattern. To me some patterns looked like L shapes and inverted L shapes together with boxes.
- YES I did, i cant name the partten but they had a similar arrangement on both sides of the box
- Yes i noticed. Yes I can name the partten.
- The black cards were either on the left/right or top/button depending on where the black card is positioned.
- The block squares were arranged in groups of twos, fours or staright.
- I noticed a pattern but it was too late. I cant name it.
- I thought there was definately some kind of pattern but i don’t know what it was.

- I noticed that the black squares were arranged in groups. A group of black squares had at least two or at most four squares.
- Hidden black squares tended to be next to other black squares.
- Four, three, three, four
- No,I did not
- I didn’t notice a pattern so sorry can’t name it!
- Sorry, I didn’t see the pattern

Pre-Registered Hypotheses for Chapter VI

Exclusion criteria

Participants will only be invited to partake in Part 2 if they answered correctly in 4 out of 5 trials in the rule discovery quiz of Part 1 and were able to correctly name the rule after round 5.

Participants will be excluded if - they have missing data - they used an ipad rather than a desktop and mouse - they failed comprehension check 2 in Part 2 - In Part 2, they responded ‘yes’ to the question of whether they already knew the rule in round 1 in Part 1. - they made no errors in round 1 of Part 2.

Sample size

In Part 1, we will collect data until we reach either 200 participants or 150 participants who have discovered the rule (whichever comes first), based on our selection criterion (the final 2AFC task and rule naming). We will analyse all participants who return to participate in Part 2 unless they fail our exclusion criteria. The median time to complete the experiment is 10 minutes. Participants will be paid an hourly wage of £9 (and in Part 1 up to +$1 bonus and in Part 2 up to +$5 bonus).

Data analysis

We will analyse reaction times, number of clicks, positions of clicks, error patterns, and dynamic cursor movements. The location of the mouse cursor will be recorded at all times, to later reconstruct a dynamic simulation of participants’ behaviour throughout the game. We will use repeated-measures 2-way ANOVA collapsed across boards, unless otherwise specified.

The games of each participants will be separated into four sets: Part 1, before rule discovery Part 1, after rule discovery (for this we will use their reported point of rule discovery) Part 2, before rule discovery Part 2, after rule discovery (for this we will use the first round in which participants report having discovered the rule) Unless otherwise specified, each test statistic will be tested as a function of experiment part (1 or 2), knowledge state (before or after rule discovery), and the interaction between the two. A follow up regression analysis will test the effect of board number and experiment part, as well as the interaction between them, on test statistics.

Hypotheses

1. We will run a 2-way ANOVA (factors: study part and knowledge state) on the difference in the number of clicks to completion per board.
2. We will run a 2-way ANOVA (factors: study part and knowledge state) on the rate of errors that are inconsistent with knowledge of the symmetry rule (all symmetric twins of already revealed white cards; see Fig 1. Middle (cards framed in red), as a proportion of all errors (IIa), and separately, as a proportion of legal errors (IIb).
3. We will run a 2-way ANOVA (factors: study part and knowledge state) on the number of long (more than 500 ms) dwellings over cards that are inconsistent with the symmetry rule (as defined above), and which do not result in a click. We trace participants' mouse trajectory across each board and mark sharp angles (between 30-180 degrees) as potential decision points. Within this range of sharp movements, we select all phases in which a participant hovered over the same unclicked card for more than 500 milliseconds. If they did not click on this card,

we consider this event a ‘ghost click’, i.e. a click that could have been committed to, but which the participant decided against after considering it.

4. We will run a 2-way ANOVA (factors: study part and knowledge state) on the number of long (more than 500 ms) dwellings over unrevealed black cards, and which do not result in a click.
5. We will run a 2-way ANOVA (factors: study part and knowledge state) on the difference in decision time between the first and second cards revealed within symmetric pairs. We will control for each person’s average RT at each click made, since RT for later clicks is likely smaller. We will extract the difference in median RT between all black cards that were the first of a pair vs the second of a pair.
6. We will run a 2-way ANOVA (factors: study part and knowledge state) on the difference between total horizontal and vertical distance traversed by the cursor across a game.
7. We will test the hypothesis that participants in Part 2 make choices that are consistent with the state of the revealed board at each timestep. We do this by constructing a model which enumerates all possible board arrangements according to the rules of the game (but not the symmetry rule). After each click, the model excludes all arrangements that are inconsistent with the newly obtained information. Using the remaining set of viable boards, each location on the grid is then ranked by the likelihood of holding a black card. This rank predicts the suitability of choosing the location according to the model. We will compare the rank of the data according to this model, against the rank of a shuffled order of clicks/errors.
8. We will then construct a symmetric model which only takes symmetric arrangements into account. The set of possible arrangements on the board is much

smaller than the set considered by the previously constructed, naive model. The mean rank per board (for all guesses, VIIIa, and for errors only, VIIIb) will be subjected to the 2-way ANOVA with factors study part and knowledge state.

Additional Analyses Chapter VI

Avoidable Errors & Avoidable False Hypotheses

Does the rate of errors that are inconsistent with knowledge of the symmetry rule differ between the conditions? We intended to test the hypothesis that actors are significantly less likely to commit types of errors that are inconsistent with knowledge of the symmetry rule. Avoidable errors are all symmetric twins of already revealed white cards (see cards framed in red Figure 16 Top Midde). Rule-followers would know that these locations cannot contain a black card and would therefore avoid clicking on them. We intended to test the hypothesis that the rate of avoidable errors as a proportion of all errors/all legal errors was significantly lower for actors than for players in early boards.

In Figure 25 A we plot the proportion of avoidable errors for players and actors.

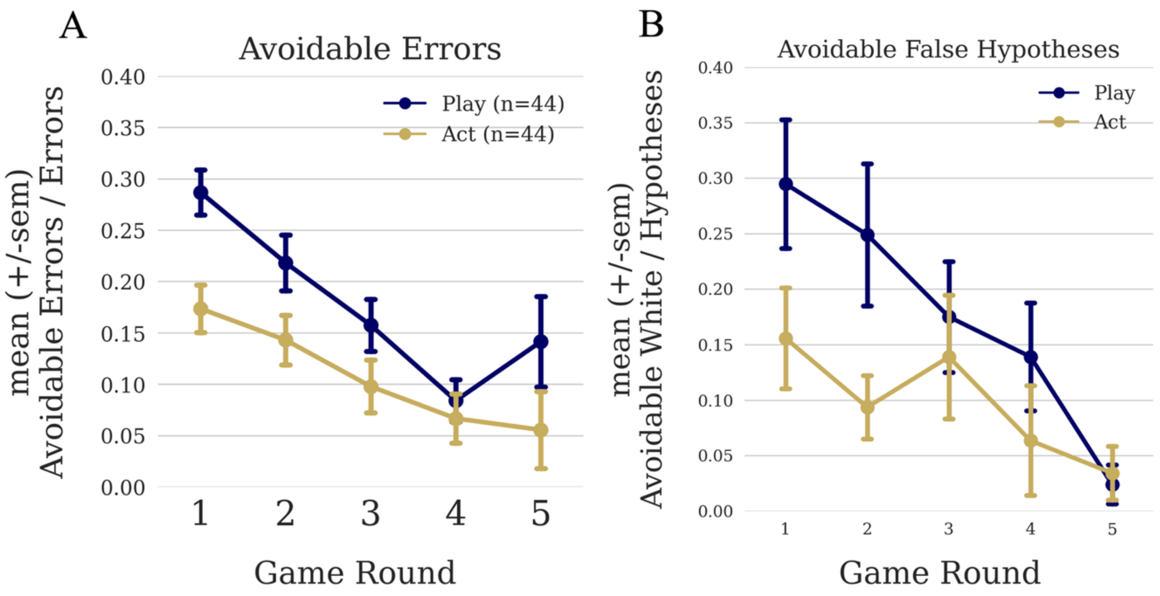

Figure 25: **Additional Analysis of False Hypotheses Chapter VI A)** Avoidable errors as a proportion of all errors made, per condition **B)** Avoidable *false* hypotheses (considering clicking on white) as a proportion of all hypotheses, per condition.

This measure however could not be de-correlated from the number of errors. We do not show the pre-registered ANOVA here (it shows a significant effect) because we realised, after we found a stark difference in the number of errors across conditions, that we need to control for the error rate. We ran a downsampling procedure and found that the difference in the number of avoidable white clicks disappears when we have the same number of white clicks made.

To control for the number of errors made, we ran the following procedure 100 times: for each participant and round, we extracted the smaller number of errors made across part 1 vs part 2. I.e. if a participant revealed 6 white cards in round 3 during part 1 but only 2 white cards in part 2, we downsampled the errors made during play to 2. We select 2 of the 6 errors made in that round at random. After both conditions had a matched number of errors, we determined if any of the remaining errors were mirror-symmetric of one another (which would yield one avoidable error). Lastly, we ran a paired t-test on the per-person ratios of all avoidable errors to all errors made across part 1 vs 2. After repeating this procedure 100x, we find no discernible difference in the rate of avoidable errors given that only 1% of all test statistics yielded a p-value below 0.05. Our analysis does not show that participants make the same types of errors in both parts at the same rate. It shows that we cannot show that there is a difference due to our confound.

In Figure 25 B we show the rate of avoidable false hypotheses over all hypotheses (computed just like correct hypotheses, except for hidden black card). If a false hypothesis occurs over a location for which the symmetric twin has already been revealed to be white, then this hypothesis was 'avoidable', in that a rule-follower (who is aware of the mirror-symmetric arrangement) would not consider this location. We intended to test the hypothesis that actors make fewer avoidable false hypotheses in early rounds, but the analysis is subject to the same confound as above, since the numerator is assessed in respect to already revealed cards. As noted in the discussion section, future versions of this study design can account for this confound.

References

Addis, D. R., Wong, A. T., & Schacter, D. L. (2007). Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*(7), 1363–1377. https://doi.org/10.1016/j.neuropsychologia.2006.10.016

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the Use of Tests: A Meta-Analysis of Practice Testing. *Review of Educational Research*, *87*(3), 659–701. https://doi.org/10.3102/0034654316689306

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval Practice Consistently Benefits Student Learning: A Systematic Review of Applied Research in Schools and

Classrooms. *Educational Psychology Review*, *33*(4), 1409–1453. https://doi.org/10.1007/s10648-021-09595-9

Aggleton, J. P., & Nelson, A. J. D. (2020). Distributed interactive brain circuits for object-in-place memory: A place for time? *Brain and Neuroscience Advances*, *4*, 2398212820933471. https://doi.org/10.1177/2398212820933471

Allaire, J. J., Teague, C., Scheidegger, C., Xie, Y., Dervieux, C., & Woodhull, G. (2025). *Quarto* (Version 1.7) [Computer software]. https://doi.org/10.5281/zenodo.5960048

Alleman, M., Panichello, M., Buschman, T. J., & Johnston, W. J. (2024). The neural basis of swap errors in working memory. *Proceedings of the National Academy of Sciences - PNAS*, *121*(33), Article e2401032121. https://doi.org/10.1073/pnas.2401032121

Allen, K., Brändle, F., Botvinick, M., Fan, J., Gershman, S., Gopnik, A., Griffiths, T., Hartshorne, J., Hauser, T., Ho, M., de Leeuw, J., Ma, W. J., Murayama, K., Nelson, J., van Opheusden, B., Pouncy, H., Rafner, J., Rahwan, I., Rutledge, R., … Schulz, E. (2023). *Using Games to Understand the Mind*. OSF. https://doi.org/10.31234/osf.io/hbsvj

Ash, I. K., Jee, B. D., & Wiley, J. (2012). Investigating Insight as Sudden Learning. *The Journal of Problem Solving*, *4*(2). https://doi.org/10.7771/1932-6246.1123

Ashby, F. G., & Crossley, M. J. (2010). Interactions between declarative and procedural-learning categorization systems. *Neurobiology of Learning and Memory*, *94*(1), 1–12. https://doi.org/10.1016/j.nlm.2010.03.001

Barrett, L., & Stout, D. (2024). Minds in movement: Embodied cognition in the age of artificial intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1911), 20230144. https://doi.org/10.1098/rstb.2023.0144

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660. https://doi.org/10.1017/S0140525X99002149

Barsalou, L. W., Breazeal, C., & Smith, L. B. (2007). Cognition as coordinated non-cognition. *Cognitive Processing*, *8*(2), 79–91. https://doi.org/10.1007/s10339-007-0163-1

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using Lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). *Parsimonious Mixed Models* (No. arXiv:1506.04967). arXiv. https://doi.org/10.48550/arXiv.1506.04967

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is Stronger than Good. *Review of General Psychology*, *5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Bear, A., Bensinger, S., Jara-Ettinger, J., Knobe, J., & Cushman, F. (2020). What comes to mind? *Cognition*, *194*, 104057. https://doi.org/10.1016/j.cognition.2019.104057

Bein, O., & Niv, Y. (2025). Schemas, reinforcement learning and the medial prefrontal cortex. *Nature Reviews Neuroscience*, *26*(3), 141–157. https://doi.org/10.1038/s41583-024-00893-z

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Bengio, Y., Louradour, J., Collobert, R., & Weston, J. (2009). Curriculum learning. *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, 1–8. https://doi.org/10.1145/1553374.1553380

Bennet, R., & Reiner, M. (2022). Shared mechanisms underlie mental imagery and motor planning. *Scientific Reports*, *12*(1), 2947. https://doi.org/10.1038/s41598-022-06800-9

Berckley, J., & Hattie, J. (2023). Making Learning Visible: Observable Correlates of the Aha! Moment when Moving from Surface to Deep Thinking. *The Journal of Creative Behavior*, *57*(3), 439–449. https://doi.org/10.1002/jocb.589

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: Insights from spatial processing. *Nature Reviews Neuroscience*, *9*(3), 182–194. https://doi.org/10.1038/nrn2335

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Bliss, T. V. P., & Collingridge, G. L. (1993). A synaptic model of memory: Long-term potentiation in the hippocampus. *Nature*, *361*(6407), 31–39. https://doi.org/10.1038/361031a0

Bodner, R., & Prelec, D. (2003). Self-Signaling and Diagnostic Utility in Everyday Decision Making. In I. Brocas & J. D. Carrillo (Eds.), *The Psychology of Economic Decisions* (pp. 105–124). Oxford University PressOxford. https://doi.org/10.1093/oso/9780199251063.003.0006

Boger, T., Yousif, S. R., McDougle, S., & Rutledge, R. (2025). *Random behavior is stable across tasks and time*. PsyArXiv. https://doi.org/10.31234/osf.io/mtfdc_v1

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement Learning, Fast and Slow. *Trends in Cognitive Sciences*, *23*(5), 408–422. https://doi.org/10.1016/j.tics.2019.02.006

Bouchacourt, F., Palminteri, S., Koechlin, E., & Ostojic, S. (2020). Temporal chunking as a mechanism for unsupervised learning of task-sets. *eLife*, *9*, e50469. https://doi.org/10.7554/eLife.50469

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 708–731. https://doi.org/10.1037/xlm0000061

Bröker, F., Holt, L. L., Roads, B. D., Dayan, P., & Love, B. C. (2024). Demystifying unsupervised learning: How it helps and hurts. *Trends in Cognitive Sciences*, *28*(11), 974–986. https://doi.org/10.1016/j.tics.2024.09.005

Bröker, F., Love, B. C., & Dayan, P. (2022). When unsupervised training benefits category learning. *Cognition*, *221*, 104984. https://doi.org/10.1016/j.cognition.2021.104984

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 77–165). Halsted Press.

Brown, A. L., Campione, J. C., & Barclay, C. R. (1978). *Training self-checking routines for estimating test readiness: Generalization from list learning to prose recall*.

Brown, J., Lewis, V. J., & Monk, A. F. (1977). Memorability, Word Frequency and Negative Recognition*. *Quarterly Journal of Experimental Psychology*, *29*(3), 461–473. https://doi.org/10.1080/14640747708400622

Brown, R. M., & Robertson, E. M. (2007). Off-Line Processing: Reciprocal Interactions between Declarative and Procedural Memories. *Journal of Neuroscience*, *27*(39), 10468–10475. https://doi.org/10.1523/JNEUROSCI.2799-07.2007

Brzosko, Z., Schultz, W., & Paulsen, O. (2015). Retroactive modulation of spike timing-dependent plasticity by dopamine. *eLife*, *4*, e09685. https://doi.org/10.7554/eLife.09685

Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*(5), 313–323. https://doi.org/10.1037/h0054388

Butz, M. V., & Kurz, E. F. (2017). *How the mind comes into being: Introducing cognitive science from a functional and computational perspective* (First edition). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198739692.001.0001

Canits, I., Pecher, D., & Zeelenberg, R. (2018). Effects of grasp compatibility on long-term memory for objects. *Acta Psychologica*, *182*, 65–74. https://doi.org/10.1016/j.actpsy.2017.11.009

Caya-Bissonnette, L., & Béïque, J.-C. (2024). Half a century legacy of long-term potentiation. *Current Biology*, *34*(13), R640–R662. https://doi.org/10.1016/j.cub.2024.05.008

Cazé, R. D., & van der Meer, M. A. A. (2013). Adaptive properties of differential learning rates for positive and negative outcomes. *Biological Cybernetics*, *107*(6), 711–719. https://doi.org/10.1007/s00422-013-0571-5

Chater, N., Oaksford, M., & Chater, N. (Eds.). (2008). *The probabilistic mind: Prospects for Bayesian cognitive science*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199216093.001.0001

Chronicle, E. P., MacGregor, J. N., & Ormerod, T. C. (2004). What Makes an Insight Problem? The Roles of Heuristics, Goal Conception, and Solution Recoding in Knowledge-Lean Problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 14–27. https://doi.org/10.1037/0278-7393.30.1.14

Chu, J., Tenenbaum, J. B., & Schulz, L. E. (2024). In praise of folly: Flexible goals and human cognition. *Trends in Cognitive Sciences*, *28*(7), 628–642. https://doi.org/10.1016/j.tics.2024.03.006

Chua, E. F., Schacter, D. L., Rand-Giovannetti, E., & Sperling, R. A. (2007). Evidence for a specific role of the anterior hippocampal region in successful associative encoding. *Hippocampus*, *17*(11), 1071–1080. https://doi.org/10.1002/hipo.20340

Clark, G. M., Lum, J. A. G., & Ullman, M. T. (2014). A meta-analysis and meta-regression of serial reaction time task performance in Parkinson's disease. *Neuropsychology*, *28*(6), 945–958. https://doi.org/10.1037/neu0000121

Clarkson, T. R., Cunningham, S. J., Haslam, C., & Kritikos, A. (2022). Is self always prioritised? Attenuating the ownership self-reference effect in memory. *Consciousness and Cognition*, *106*, 103420. https://doi.org/10.1016/j.concog.2022.103420

Clay, V., König, P., Kühnberger, K.-U., & Pipa, G. (2021). Learning sparse and meaningful representations through embodiment. *Neural Networks*, *134*, 23–41. https://doi.org/10.1016/j.neunet.2020.11.004

Coenen, A., Nelson, J. D., & Gureckis, T. M. (2019). Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychonomic Bulletin & Review*, *26*(5), 1548–1587. https://doi.org/10.3758/s13423-018-1470-5

Cohen, N. J., & Squire, L. R. (1980). Preserved Learning and Retention of Pattern-Analyzing Skill in Amnesia: Dissociation of Knowing How and Knowing That. *Science*, *210*(4466), 207–210. https://doi.org/10.1126/science.7414331

Collins, A., Saling, M. M., Wilson, S. J., Jackson, G. D., & Tailby, C. (2022). The Spatial Learning Task of Lhermitte and Signoret (1972): Normative Data in Adults Aged 18–45. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.860982

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058. https://doi.org/10.1016/j.cognition.2007.04.010

Cox, W. T. L., Xie, X., & Devine, P. G. (2022). Untested assumptions perpetuate stereotyping: Learning in the absence of evidence. *Journal of Experimental Social Psychology*, *102*, 104380. https://doi.org/10.1016/j.jesp.2022.104380

Cremer, C. Z., & Mazor, M. (2025a). *Behavioural markers of discovery*. OSF Preprints. https://doi.org/10.17605/OSF.IO/CKMV5

Cremer, C. Z., & Mazor, M. (2025b). *Self-modelling of counterfactual knowledge states*. https://doi.org/10.17605/OSF.IO/ZRNU6

Cremer, C. Z., & Summerfield, C. (2024). *Action bias in self vs other*. OSF Preprints. https://doi.org/10.17605/OSF.IO/59VQ3

Dasgupta, I., & Gershman, S. J. (2021). Memory as a Computational Resource. *Trends in Cognitive Sciences*, *25*(3), 240–251. https://doi.org/10.1016/j.tics.2020.12.008

Dasgupta, I., Schulz, E., & Gershman, S. J. (2017). Where do hypotheses come from? *Cognitive Psychology*, *96*, 1–25. https://doi.org/10.1016/j.cogpsych.2017.05.001

Daprati, E., Sirigu, A., & Nico, D. (2019). Remembering actions without proprioception. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *113*, 29–36. https://doi.org/10.1016/j.cortex.2018.11.020

Dekker, R. B., Otto, F., & Summerfield, C. (2022). Curriculum learning for human compositional generalization. *Proceedings of the National Academy of Sciences*, *119*(41), e2205582119. https://doi.org/10.1073/pnas.2205582119

Dennett, D. C. (1992). The Self as a Center of Narrative Gravity. In F. S. Kessel, P. M. Cole, & D. L. Johnson (Eds.), *[Book Chapter]* (pp. 4–237). Lawrence Erlbaum.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837. https://doi.org/10.1016/0042-6989(95)00294-4 de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods, 47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dijkstra, K., & Post, L. (2015). Mechanisms of embodiment. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01525

Downing-Doucet, F., & Guérard, K. (2014). A motor similarity effect in object memory. *Psychonomic Bulletin & Review*, *21*(4), 1033–1040. https://doi.org/10.3758/s13423-013-0570-5

Druckman, D., Bjork, R. A., & National Research Council (U.S.) (Eds.). (1994). *Learning, remembering, believing: Enhancing human performance*. National Academy Press.

Dudai, Y., & Carruthers, M. (2005). The Janus face of Mnemosyne. *Nature*, *434*(7033), 567–567. https://doi.org/10.1038/434567a

Dunlosky, J., & Tauber, S. (Uma). K. (Eds.). (2016). *The Oxford Handbook of Metamemory*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199336746.001.0001

Ebbinghaus, H. (1885). *Über das Gedächtnis* (1. Auflage). Duncker & Humblot.

Eichenbaum, H. (2004). Hippocampus: Cognitive Processes and Neural Representations that Underlie Declarative Memory. *Neuron*, *44*(1), 109–120. https://doi.org/10.1016/j.neuron.2004.08.028

Eichenbaum, H. (2011a). A Brain System for Declarative Memory. In *The Cognitive Neuroscience of Memory*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199778614.003.0009

Eichenbaum, H. (2011b). A Brain System for Procedural Memory. In H. Eichenbaum (Ed.), *The Cognitive Neuroscience of Memory: An Introduction* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199778614.003.0010

Eichenbaum, H. (2011c). Cellular Mechanisms of Memory: Complex Circuits. In H. Eichenbaum (Ed.), *The Cognitive Neuroscience of Memory: An Introduction* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199778614.003.0003

Eichenbaum, H., & Cohen, N. J. (2001). *From Conditioning to Conscious Recollection: Memory Systems of the Brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195178043.001.0001

Eichenbaum, H., & Cohen, N. J. (2004). Multiple Memory Systems: A Historical Perspective. In H. Eichenbaum & N. J. Cohen (Eds.), *From Conditioning to Conscious Recollection: Memory systems of the brain* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195178043.003.0002

Eichenbaum, H., & Fortin, N. J. (2009). The neurobiology of memory based predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1183–1191. https://doi.org/10.1098/rstb.2008.0306

Epstein, R., Kirshnit, C. E., Lanza, R. P., & Rubin, L. C. (1984). “Insight” in the pigeon: Antecedents and determinants of an intelligent performance. *Nature*, *308*(5954), 61–62. https://doi.org/10.1038/308061a0

Ergo, K., De Loof, E., & Verguts, T. (2020). Reward Prediction Error and Declarative Memory. *Trends in Cognitive Sciences*, *24*(5), 388–397. https://doi.org/10.1016/j.tics.2020.02.009

Eskreis-Winkler, L., & Fishbach, A. (2022). You Think Failure Is Hard? So Is Learning From It. *Perspectives on Psychological Science*, *17*(6), 1511–1524. https://doi.org/10.1177/17456916211059817

Fabre, C. (2022). *Spying Through a Glass Darkly: The Ethics of Espionage and Counter-Intelligence*. Oxford University Press.

Ferreira, C. S., & Wimber, M. (2023). The Testing Effect for Visual Materials Depends on Preexisting Knowledge. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *49*(10), 1557–1571. https://doi.org/10.1037/xlm0001248

Flavell, J. H. (1971). First Discussant's Comments: What is Memory Development the Development of? *Human Development*, *14*(4), 272–278. https://doi.org/10.1159/000271221

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Fleming, S. M. (2024). Metacognition and Confidence: A Review and Synthesis. *Annual Review of Psychology*, *75*(Volume 75, 2024), 241–268. https://doi.org/10.1146/annurev-psych-022423-032425

Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review*, *124*(1), 91–114. https://doi.org/10.1037/rev0000045

Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1280–1286. https://doi.org/10.1098/rstb.2012.0021

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00443

Freedberg, M. V., Reeves, J. A., Fioriti, C. M., Murillo, J., Voss, J. L., & Wassermann, E. M. (2022). A direct test of competitive versus cooperative episodic-procedural network dynamics in human memory. *Cerebral Cortex (New York, N.Y.: 1991)*, *32*(21), 4715–4732. https://doi.org/10.1093/cercor/bhab512

Freedberg, M., Toader, A. C., Wassermann, E. M., & Voss, J. L. (2020). Competitive and cooperative interactions between medial temporal and striatal learning systems. *Neuropsychologia*, *136*, 107257. https://doi.org/10.1016/j.neuropsychologia.2019.107257

Fu, R., & He, H. (2025). The influence of embodied cognition on the perception and representation of numbers: A meta-analysis. *Acta Psychologica*, *258*, 105180. https://doi.org/10.1016/j.actpsy.2025.105180

Gade, M., Druey, M. D., Souza, A. S., & Oberauer, K. (2014). Interference within and between declarative and procedural representations in working memory. *Journal of Memory and Language*, *76*, 174–194. https://doi.org/10.1016/j.jml.2014.07.002

Gagné, N., & Franzen, L. (2023). How to Run Behavioural Experiments Online: Best Practice Suggestions for Cognitive Psychology and Neuroscience. *Swiss Psychology Open: The Official Journal of the Swiss Psychological Society*, *3*(1). https://doi.org/10.5334/spo.34

Ganellen, R. J., & Carver, C. S. (1985). Why does self-reference promote incidental encoding? *Journal of Experimental Social Psychology*, *21*(3), 284–300. https://doi.org/10.1016/0022-1031(85)90021-6

Geist, M., & Pietquin, O. (2010). Kalman temporal differences. *The Journal of Artificial Intelligence Research*, *39*, 483–532.

Gershman, S. J. (2015a). A Unifying Probabilistic View of Associative Learning. *PLOS Computational Biology*, *11*(11), e1004567. https://doi.org/10.1371/journal.pcbi.1004567

Gershman, S. J. (2015b). Do learning rates adapt to the distribution of rewards? *Psychonomic Bulletin & Review*, *22*(5), 1320–1327. https://doi.org/10.3758/s13423-014-0790-3

Gershman, S. J., & Uchida, N. (2019). Believing in dopamine. *Nature Reviews Neuroscience*, *20*(11), 703–714. https://doi.org/10.1038/s41583-019-0220-7

Gerstner, W., Lehmann, M., Liakoni, V., Corneil, D., & Brea, J. (2018). Eligibility Traces and Plasticity on Behavioral Time Scales: Experimental Support of NeoHebbian Three-Factor Learning Rules. *Frontiers in Neural Circuits*, *12*. https://doi.org/10.3389/fncir.2018.00053

Gibson, J. J. (1950). The perception of the visual world. In *The perception of the visual world*. Greenwood.

Gilbert, P. E., & Kesner, R. P. (2002). Role of the rodent hippocampus in paired-associate learning involving associations between a stimulus and a spatial location. *Behavioral Neuroscience*, *116*(1), 63–71. https://doi.org/10.1037//0735-7044.116.1.63

Gilbert, P. E., & Kesner, R. P. (2004). Memory for objects and their locations: The role of the hippocampus in retention of object-place associations. *Neurobiology of Learning and Memory*, *81*(1), 39–45. https://doi.org/10.1016/s1074-7427(03)00069-8

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*(1), 1–19. https://doi.org/10.1017/S0140525X97000010

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*(3), 558–565. https://doi.org/10.3758/BF03196313

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the Revolution to Embodiment: 25 Years of Cognitive Psychology. *Perspectives on Psychological Science*, *8*(5), 573–585. https://doi.org/10.1177/1745691613498098

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences*, *16*(1), 1–14. https://doi.org/10.1017/S0140525X00028636

Graybiel, A. M. (1995). Building action repertoires: Memory and learning functions of the basal ganglia. *Current Opinion in Neurobiology*, *5*(6), 733–741. https://doi.org/10.1016/0959-4388(95)80100-6

Guérard, K., Guerrette, M.-C., & Rowe, V. P. (2015). The role of motor affordances in immediate and long-term retention of objects. *Acta Psychologica*, *162*, 69–75. https://doi.org/10.1016/j.actpsy.2015.10.008

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences*, *18*(4), 194–202. https://doi.org/10.1016/j.tics.2014.01.003

Haazebroek, P., Raffone, A., & Hommel, B. (2017). HiTEC: A connectionist model of the interaction between perception and action planning. *Psychological Research*, *81*(6), 1085–1109. https://doi.org/10.1007/s00426-016-0803-0

Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196–207. https://doi.org/10.1038/nrn.2017.14

Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage*, *67*, 283–297. https://doi.org/10.1016/j.neuroimage.2012.11.020

Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, *56*(1), 51–65. https://doi.org/10.1037/h0062474

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1), 335–346. https://doi.org/10.1016/0167-2789(90)90087-6

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Hart, G., Bradfield, L. A., Fok, S. Y., Chieng, B., & Balleine, B. W. (2018). The Bilateral Prefronto-striatal Pathway Is Necessary for Learning New Goal-Directed Actions. *Current Biology*, *28*(14), 2218–2229.e7. https://doi.org/10.1016/j.cub.2018.05.028

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-Inspired Artificial Intelligence. *Neuron*, *95*(2), 245–258. https://doi.org/10.1016/j.neuron.2017.06.011

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, *104*(5), 1726–1731. https://doi.org/10.1073/pnas.0610561104

Hattori, R., Hedrick, N. G., Jain, A., Chen, S., You, H., Hattori, M., Choi, J.-H., Lim, B. K., Yasuda, R., & Komiyama, T. (2023). Meta-reinforcement learning via orbitofrontal cortex. *Nature Neuroscience*, *26*(12), 2182–2191. https://doi.org/10.1038/s41593-023-01485-3

Hebart, M. N., Dickter, A. H., Kidder, A., Kwok, W. Y., Corriveau, A., Van Wicklin, C., & Baker, C. I. (2019). THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images. *PLoS ONE*, *14*(10), e0223792. https://doi.org/10.1371/journal.pone.0223792

Hebb, D. O. (2002). The first stage of perception: Growth of the assembly. In *The organization of behavior*. Taylor & Francis Group.

Hemmer, P., & Steyvers, M. (2009). A Bayesian Account of Reconstructive Memory. *Topics in Cognitive Science*, *1*(1), 189–202. https://doi.org/10.1111/j.1756-8765.2008.01010.x

Henke, K., Buck, A., Weber, B., & Wieser, H. G. (1997). Human hippocampus establishes associations in memory. *Hippocampus*, *7*(3), 249–256. https://doi.org/10.1002/(SICI)1098-1063(1997)7:3<249::AID-HIPO1>3.0.CO;2-G

Henke, K., Weber, B., Kneifel, S., Wieser, H. G., & Buck, A. (1999). Human hippocampus associates information in memory. *Proceedings of the National Academy of Sciences*, *96*(10), 5884–5889. https://doi.org/10.1073/pnas.96.10.5884

Heuer, A., Crawford, J. D., & Schubö, A. (2017). Action relevance induces an attentional weighting of representations in visual working memory. *Memory & Cognition*, *45*(3), 413–427. https://doi.org/10.3758/s13421-016-0670-3

Hilgard, E. R., & Marquis, D. G. (1940). *Conditioning and Learning*. D. Appleton-Century Company.

Hinton, G. E., Dayan, P., Frey, B. J., & Neal, R. M. (1995). The "Wake-Sleep" Algorithm for Unsupervised Neural Networks. *Science*, *268*(5214), 1158–1161. https://doi.org/10.1126/science.7761831

Hommel, B. (2015). The theory of event coding (TEC) as embodied-cognition framework. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01318

Hon, N., & Teo, J. Z. (2024). The salience of an agency experience influences its ability to improve memory. *Psychological Research*, *89*(1), 11. https://doi.org/10.1007/s00426-024-02032-w

Hon, N., & Yeo, N. (2021). Having a sense of agency can improve memory. *Psychonomic Bulletin & Review*, *28*(3), 946–952. https://doi.org/10.3758/s13423-020-01849-x

Hund, A. M., & Plumert, J. M. (2003). Does Information About What Things Are Influence Children's Memory for Where Things Are? *Developmental Psychology*, *39*(6), 939–948. https://doi.org/10.1037/0012-1649.39.6.939

Ianì, F. (2021). Embodied cognition: So flexible as to be “disembodied”? *Consciousness and Cognition*, *88*, 103075. https://doi.org/10.1016/j.concog.2021.103075

Irish, M., & Vatansever, D. (2020). Rethinking the episodic-semantic distinction from a gradient perspective. *Current Opinion in Behavioral Sciences*, *32*, 43–49. https://doi.org/10.1016/j.cobeha.2020.01.016

Jackson, G. M., Jackson, S. R., Harrison, J., Henderson, L., & Kennard, C. (1995). Serial reaction time learning and Parkinson's disease: Evidence for a procedural learning deficit. *Neuropsychologia*, *33*(5), 577–593. https://doi.org/10.1016/0028-3932(95)00010-Z

James, W. (1901). *The principles of psychology* (Vol. 1). Macmillan; Co., Ltd.

Johansson, P., Hall, L., Sikström, S., Tärning, B., & Lind, A. (2006). How something can be said about telling more than we can know: On choice blindness and introspection. *Consciousness and Cognition*, *15*(4), 673–692. https://doi.org/10.1016/j.concog.2006.09.004

Jones, G. (2003). Testing two cognitive theories of insight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(5), 1017–1027. https://doi.org/10.1037/0278-7393.29.5.1017

Jones, M., & Love, B. C. (2011). Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences*, *34*(4), 169–188. https://doi.org/10.1017/S0140525X10003134

Kalra, P. B. (2025). Asking the right questions: Interrogating the logic and assumptions of paradigms used to investigate interactions between procedural and declarative memory in category learning. *Frontiers in Cognition*, *3*. https://doi.org/10.3389/fcogn.2024.1505513

Kang, P., Tobler, P. N., & Dayan, P. (2024). Bayesian reinforcement learning: A basic overview. *Neurobiology of Learning and Memory*, *211*, 107924. https://doi.org/10.1016/j.nlm.2024.107924

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, *22*(3), 374–419. https://doi.org/10.1016/0010-0285(90)90008-R

Katahira, K. (2018). The statistical structures of reinforcement learning with asymmetric value updates. *Journal of Mathematical Psychology*, *87*, 31–45. https://doi.org/10.1016/j.jmp.2018.09.002

Katyal, S., & Fleming, S. M. (2024). The future of metacognition research: Balancing construct breadth with measurement rigor. *Cortex*, *171*, 223–234. https://doi.org/10.1016/j.cortex.2023.11.002

Kermadi, I., & Joseph, J. P. (1995). Activity in the caudate nucleus of monkey during spatial sequencing. *Journal of Neurophysiology*, *74*(3), 911–933. https://doi.org/10.1152/jn.1995.74.3.911

Kerrén, C., Zhao, Y., & Griffiths, B. J. (2024). A reduction in self-reported confidence accompanies the recall of memories distorted by prototypes. *Communications Psychology*, *2*(1), 58. https://doi.org/10.1038/s44271-024-00108-2

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice Certainty Is Informed by Both Evidence and Decision Time. *Neuron*, *84*(6), 1329–1342. https://doi.org/10.1016/j.neuron.2014.12.015

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PLOS ONE*, *7*(5), e36399. https://doi.org/10.1371/journal.pone.0036399

Kim, S. (2020). Bidirectional competitive interactions between motor memory and declarative memory during interleaved learning. *Scientific Reports*, *10*(1), 6916. https://doi.org/10.1038/s41598-020-64039-8

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*(4), 513–549. https://doi.org/10.1016/0364-0213(94)90007-8

Kizilirmak, J. M., Thuerich, H., Folta-Schoofs, K., Schott, B. H., & Richardson-Klavehn, A. (2016). Neural Correlates of Learning from Induced Insight: A Case for Reward-Based Episodic Encoding. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01693

Klayman, J., & Ha, Y.-W. (1989). *Hypothesis Testing in Rule Discovery: Strategy, Structure, and Content*.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A Neostriatal Habit Learning System in Humans. *Science*, *273*(5280), 1399–1402. https://doi.org/10.1126/science.273.5280.1399

Körner, A., Topolinski, S., & Strack, F. (2015). Routes to embodiment. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00940

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior*, *36*(3), 210–226. https://doi.org/10.3758/LB.36.3.210

Kumaran, D., & Maguire, E. A. (2009). Novelty signals: A window into hippocampal information processing. *Trends in Cognitive Sciences*, *13*(2), 47–54. https://doi.org/10.1016/j.tics.2008.11.004

Lagacé, S., & Guérard, K. (2015). When motor congruency modulates immediate memory for objects. *Acta Psychologica*, *157*, 65–73. https://doi.org/10.1016/j.actpsy.2015.02.009

Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies” (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE*, *10*(6), e0130834. https://doi.org/10.1371/journal.pone.0130834

Lashley, K. S. (1950). In search of the engram. In *Physiological mechanisms in animal behavior. (Society’s Symposium IV.)* (pp. 454–482). Academic Press.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

LeGris, S., Lake, B., & Gureckis, T. M. (2024). Predicting Insight during Physical Reasoning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *46*(0).

Lenth, R. V. (2024). *Emmeans: Estimated marginal means, aka least-squares means*. https://CRAN.R-project.org/package=emmeans

Lerousseau, J. P., & Summerfield, C. (2024). Space as a Scaffold for Rotational Generalisation of Abstract Concepts. *eLife*, *13*. https://doi.org/10.7554/eLife.93636.2

Levin, D. T., & Angelone, B. L. (2008). The Visual Metacognition Questionnaire: A Measure of Intuitions about Vision. *The American Journal of Psychology*, *121*(3), 451–472. https://doi.org/10.2307/20445476

Levin, M. (2024). Self-Improvising Memory: A Perspective on Memories as Agential, Dynamically Reinterpreting Cognitive Glue. *Entropy*, *26*(6), 481. https://doi.org/10.3390/e26060481

Lewicka, M., Czapinski, J., & Peeters, G. (1992). Positive-negative asymmetry or "When the heart needs a reason." *European Journal of Social Psychology*, *22*(5), 425–434. https://doi.org/10.1002/ejsp.2420220502

Lhermitte, F., & Signoret, J.-L. (1972). Analyse neuropsychologique et différenciation des syndromes amnésiques [neuropsychological analysis and differentiation of amnesia syndromes]. *Revue Neurologique*, *126*, 161–178.

Li, A. Y., Yuan, J. Y., Pun, C., & Barense, M. D. (2023). The effect of memory load on object reconstruction: Insights from an online mouse-tracking task. *Attention, Perception, & Psychophysics*, *85*(5), 1612–1630. https://doi.org/10.3758/s13414-022-02650-9

Lin, J. Y., Pype, A. D., Murray, S. O., & Boynton, G. M. (2010). Enhanced Memory for Scenes Presented at Behaviorally Relevant Points in Time. *PLOS Biology*, *8*(3), e1000337. https://doi.org/10.1371/journal.pbio.1000337

Lin, Z., Baptista, S., Gattoni, R., Arnold, J., Flickinger, D., Stringer, C., & Pachitariu, M. (2024). Distinct streams for supervised and unsupervised learning in the visual cortex. *bioRxiv* 2024.02.25.581990; doi: https://doi.org/10.1101/2024.02.25.581990.

Lindley, D. V. (1956). On a Measure of the Information Provided by an Experiment. *The Annals of Mathematical Statistics*, *27*(4), 986–1005. https://doi.org/10.1214/aoms/1177728069

Lindsay, G. W. (2020). Attention in Psychology, Neuroscience, and Machine Learning. *Frontiers in Computational Neuroscience*, *14*, 29. https://doi.org/10.3389/fncom.2020.00029

Lubinus, C., Einhäuser, W., Schiller, F., Kircher, T., Straube, B., & Van Kemenade, B. M. (2022). Action-based predictions affect visual perception, neural processing, and pupil size, regardless of temporal predictability. *NeuroImage*, *263*, 119601. https://doi.org/10.1016/j.neuroimage.2022.119601

Maier, N. (1931). Reasoning in Humans. The Solution of a Problem and its Appearance in Consciousness. *Journal of Comparative Psychology*, *12*(2), 181–194. https://doi.org/10.1037/h0071361

Makovski, T., Jiang, Y. V., & Swallow, K. M. (2013). How do observer's responses affect visual long-term memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1097–1105. https://doi.org/10.1037/a0030908

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94–122. https://doi.org/10.1037/a0032108

Mattfeld, A. T., & Stark, C. E. L. (2015). Functional contributions and interactions between the human hippocampus and subregions of the striatum during arbitrary associative learning and memory. *Hippocampus*, *25*(8), 900–911. https://doi.org/10.1002/hipo.22411

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mazor, M. (2025). Inference About Absence as a Window Into the Mental Self-Model. *Open Mind : Discoveries in Cognitive Science*, *9*, 635–651. https://doi.org/10.1162/opmi_a_00206

Mazor, M., Firestone, C., & Phillips, I. (2024, March 11). *Pretending not to know reveals a capacity for model-based self-simulation*. https://doi.org/10.31234/osf.io/pgxrz

Mazor, M., Moran, R., & Press, C. (2024). The role of counterfactual visibility in inference about absence. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *46*(0). Retrieved from https://escholarship.org/uc/item/4352x5v1

Mazor, M., Siegel, M., & Tenenbaum, J. B. (2023). Prospective search time estimates reveal the strengths and limits of internal models of visual search. *Journal of Experimental Psychology: General*, *152*(7), 1951–1966. https://doi.org/10.1037/xge0001360

McCarthy, L. et al. (2024). *p5.js: A JavaScript library for creative coding*. https://p5js.org.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*(4), 287–330. https://doi.org/10.1037/0033-295X.86.4.287

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

McCoy, J., Paul, L. A., & Ullman, T. (2019). Modal Prospection. In A. I. Goldman & B. P. McLaughlin (Eds.), *Metaphysics and Cognitive Science* (1st ed., pp. 235–267). Oxford University PressNew York. https://doi.org/10.1093/oso/9780190639679.003.0010

McDonnell, J. V., Jew, C. A., & Gureckis, T. M. (2012). Sparse category labels obstruct generalization of category membership. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *34*(34).

Meder, B., Cupri, V., & Nelson, J. (2022). What makes a good query? Prospects for a comprehensive theory of human information acquisition. In *The drive for knowledge: The science of human information-seeking*. Cambridge University Press.

Metcalfe, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition: Knowing about Knowing*. The MIT Press. https://doi.org/10.7551/mitpress/4561.001.0001

Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*, *88*(1), 78–92. https://doi.org/10.1016/j.neuron.2015.09.039

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin*, *117*(3), 363–386. https://doi.org/10.1037/0033-2909.117.3.363

Milner, B. (1965). Physiologie de l’Hippocampe. *Neuropsychologia.*, *3*(3).

Montague, P., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, *16*(5), 1936–1947. https://doi.org/10.1523/jneurosci.16-05-01936.1996

Morishige, K., Hiroe, N., Sato, M., & Kawato, M. (2021). Common cortical areas have different neural mechanisms for covert and overt visual pursuits. *Scientific Reports*, *11*(1), 13933. https://doi.org/10.1038/s41598-021-93259-9

Morris, A. (2025). Invisible Gorillas in the Mind: Internal Inattentional Blindness and the Prospect of Introspection Training. *Open Mind*, *9*, 606–634. https://doi.org/10.1162/opmi_a_00204

Muzzio, I. A., Kentros, C., & Kandel, E. (2009). What is remembered? Role of attention on the encoding and retrieval of hippocampal representations. *The Journal of Physiology*, *587*(12), 2837–2854. https://doi.org/10.1113/jphysiol.2009.172445

Nadel, L., & Hardt, O. (2011). Update on Memory Systems and Processes. *Neuropsychopharmacology*, *36*(1), 251–273. https://doi.org/10.1038/npp.2010.169

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391. https://doi.org/10.1038/nature03390

Nelson, J. D. (2005). Finding Useful Questions: On Bayesian Diagnosticity, Probability, Impact, and Information Gain. *Psychological Review*, *112*(4), 979–999. https://doi.org/10.1037/0033-295X.112.4.979

Nelson, J. D. (2008). Towards a rational theory of human information acquisition. In N. Chater & M. Oaksford (Eds.), *The Probabilistic Mind:Prospects for Bayesian cognitive science* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199216093.003.0007

Newell, A. (1972). *Human problem solving*. Englewood Cliffs, N.J.: Prentice-Hall.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Niv, Y., & Montague, P. R. (2009). Chapter 22 - Theoretical and Empirical Studies of Learning. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics* (pp. 331–351). Academic Press. https://doi.org/10.1016/B978-0-12-374176-9.00022-1

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608–631. https://doi.org/10.1037/0033-295X.101.4.608

Oaksford, M., & Chater, N. (2007). Being economical with the evidence: Collecting data and testing hypotheses. In M. Oaksford & N. Chater (Eds.), *Bayesian Rationality: The probabilistic approach to human reasoning* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524496.003.0006

Oey, L. A., Schachner, A., & Vul, E. (2023). *Designing and Detecting Lies by Reasoning About Other Agents*.

Olivers, C. N. L., & Roelfsema, P. R. (2020). Attention for action in visual working memory. *Cortex*, *131*, 179–194. https://doi.org/10.1016/j.cortex.2020.07.011

Öllinger, M., Jones, G., & Knoblich, G. (2014). The dynamics of search, impasse, and representational change provide a coherent explanation of difficulty in the nine-dot problem. *Psychological Research*, *78*(2), 266–275. https://doi.org/10.1007/s00426-013-0494-8

Ostarek, M., & Bottini, R. (2021). Towards Strong Inference in Research on Embodiment – Possibilities and Limitations of Causal Paradigms. *Journal of Cognition*, *4*(1). https://doi.org/10.5334/joc.139

Ottenheimer, D. J., Bari, B. A., Sutlief, E., Fraser, K. M., Kim, T. H., Richard, J. M., Cohen, J. Y., & Janak, P. H. (2020). A quantitative reward prediction error signal in the ventral pallidum. *Nature Neuroscience*, *23*(10), 1267–1276. https://doi.org/10.1038/s41593-020-0688-5

Packard, M. G., & McGaugh, J. L. (1996). Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning. *Neurobiology of Learning and Memory*, *65*(1), 65–72. https://doi.org/10.1006/nlme.1996.0007

Pan, W.-X., Schmidt, R., Wickens, J. R., & Hyland, B. I. (2005). Dopamine Cells Respond to Predicted Events during Classical Conditioning: Evidence for Eligibility Traces in the Reward-Learning Network. *Journal of Neuroscience*, *25*(26), 6235–6242. https://doi.org/10.1523/JNEUROSCI.1478-05.2005

Patterson, J. D., & Karuza, E. (2022). *Schrödinger's Category: Active Learning in the Face of Label Ambiguity*. PsyArXiv. https://doi.org/10.31234/osf.io/pnzdy

Paul, L. A., Ullman, T. D., De Freitas, J., & Tenenbaum, J. (2023). *Reverse-engineering the self* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/vzwrn

Pavlov, I. P., & Anrep, G. V. (1927). Conditioned reflexes : An investigation of the physiological activity of the cerebral cortex. In *Conditioned reflexes : An investigation of the physiological activity of the cerebral cortex*. Oxford University Press.

Payne, J. D., Schacter, D. L., Propper, R. E., Huang, L.-W., Wamsley, E. J., Tucker, M. A., Walker, M. P., & Stickgold, R. (2009). The role of sleep in false memory formation. *Neurobiology of Learning and Memory*, *92*(3), 327–334. https://doi.org/10.1016/j.nlm.2009.03.007

Pelkey, J. (2023). Embodiment and language. *WIREs Cognitive Science*, *14*(5), e1649. https://doi.org/10.1002/wcs.1649

Petitet, P., Attaallah, B., Manohar, S. G., & Husain, M. (2021). The computational cost of active information sampling before decision-making under uncertainty. *Nature Human Behaviour*, *5*(7), 935–946. https://doi.org/10.1038/s41562-021-01116-6

Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, *414*(6863), 546–550. https://doi.org/10.1038/35107080

Popp, P. J. O., Newell, B., Bartels, D., & Gureckis, T. (2024). *Can cognitive discovery be incentivized with money?* https://doi.org/10.31234/osf.io/phfwx

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Press, C., Thomas, E. R., & Yon, D. (2023a). Cancelling cancellation? Sensorimotor control, agency, and prediction. *Neuroscience & Biobehavioral Reviews*, *145*, 105012. https://doi.org/10.1016/j.neubiorev.2022.105012

Press, C., Thomas, E., & Yon, D. (2023b). The Synergistic Relationship between Perception and Action. In *Expected Experiences* (1st ed., pp. 76–95). Routledge. https://doi.org/10.4324/9781003084082-5

Prince, M. (2004). Does Active Learning Work? A Review of the Research. *Journal of Engineering Education*, *93*(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Prinz, W. (1990). A Common Coding Approach to Perception and Action. In *Relationships Between Perception and Action* (pp. 167–201). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-75348-0_7

Prinz, W. (1997). Perception and Action Planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154. https://doi.org/10.1080/713752551

Proffitt, D. R. (2006). Embodied Perception and the Economy of Action. *Perspectives on Psychological Science*, *1*(2), 110–122. https://doi.org/10.1111/j.1745-6916.2006.00008.x

Prolific. (2024). *Prolific platform for participant recruitment*. https://www.prolific.com

Prull, M. W., Liu, Y., Adhikari, N., Higdon, S. A., Stewart, K. S., & Calo, Z. R. (2023). Can divided attention at retrieval improve memory? Effects of target detection during recognition. *Memory*, *31*(4), 573–587. https://doi.org/10.1080/09658211.2023.2184458

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576–582. https://doi.org/10.1038/nrn1706

Qian, L., Burrell, M., Hennig, J. A., Matias, S., Murthy, V. N., Gershman, S. J., & Uchida, N. (2025). Prospective contingency explains behavior and dopamine signals during associative learning. *Nature Neuroscience*, *28*(6), 1280–1292. https://doi.org/10.1038/s41593-025-01915-4

Quian Quiroga, R. (2023). An integrative view of human hippocampal function: Differences with other species and capacity considerations. *Hippocampus*, *33*(5), 616–634. https://doi.org/10.1002/hipo.23527

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rainforth, T., Foster, A., Ivanova, D. R., & Smith, F. B. (2023). *Modern Bayesian Experimental Design* (arXiv:2302.14545). arXiv. https://doi.org/10.48550/arXiv.2302.14545

Randall, A. (2019). *Should We Be Silent and Not Speak Painting* [Painting, Acrylic on Canvas]. https://www.saatchiart.com/art/Painting-Should-We-Be-Silent-and-Not-Speak/1310929/6652317/view

Ramscar, M., Dye, M., & McCauley, S. M. (2013). ERROR AND EXPECTATION IN LANGUAGE LEARNING: THE CURIOUS ABSENCE OF "MOUSES" IN ADULT SPEECH. *Language (Baltimore)*, *89*(4), 760–793. https://doi.org/10.1353/lan.2013.0068

Rawson, K. A., & Dunlosky, J. (2012). When Is Practice Testing Most Effective for Improving the Durability and Efficiency of Student Learning? *Educational Psychology Review*, *24*(3), 419–435. https://doi.org/10.1007/s10648-012-9203-1

Raz, G., & Saxe, R. (2020). Learning in Infancy Is Active, Endogenously Motivated, and Depends on the Prefrontal Cortices. *Annual Review of Developmental Psychology*, *2*(1), 247–268. https://doi.org/10.1146/annurev-devpsych-121318-084841

Ren, Q., Herbelin, B., Meyer, N. H., Stampacchia, S., Schütz-Bosbach, S., & Blanke, O. (2025). Better Controlled, Better Remembered? Investigating the Impact of Sense of Agency on Episodic Memory. *The Association for the Scientific Study of Consciousness*.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy

(Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Rioult-Pedotti, M.-S., Friedman, D., Hess, G., & Donoghue, J. P. (1998). Strengthening of horizontal cortical connections following skill learning. *Nature Neuroscience*, *1*(3), 230–234. https://doi.org/10.1038/678

Rivers, M. L. (2021). Metacognition About Practice Testing: A Review of Learners' Beliefs, Monitoring, and Control of Test-Enhanced Learning. *Educational Psychology Review*, *33*(3), 823–862. https://doi.org/10.1007/s10648-020-09578-2

Robinson, M. D., & Thomas, L. E. (Eds.). (2021). *Handbook of Embodied Psychology: Thinking, Feeling, and Acting*. Springer International Publishing. https://doi.org/10.1007/978-3-030-78471-3

RoedigerIII, H. L., & Karpicke, J. D. (2006). Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408. https://doi.org/10.1037/h0042519

Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature Communications*, *10*(1), 1141. https://doi.org/10.1038/s41467-019-09075-3

Rouhani, N., & Niv, Y. (2021). Signed and unsigned reward prediction errors dynamically enhance learning and memory. *eLife*, *10*, e61077. https://doi.org/10.7554/eLife.61077

Rouy, M., De Gardelle, V., Reyes, G., Sackur, J., Vergnaud, J. C., Filevich, E., & Faivre, N. (2022). Metacognitive improvement: Disentangling adaptive training from experimental confounds. *Journal of Experimental Psychology: General*, *151*(9), 2083–2091. https://doi.org/10.1037/xge0001185

Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, *143*, 203–216. https://doi.org/10.1016/j.cognition.2015.07.004

Ruggeri, A., Swaboda, N., Sim, Z. L., & Gopnik, A. (2019). Shake it baby, but only when needed: Preschoolers adapt their exploratory strategies to the information structure of the task. *Cognition*, *193*, 104013. https://doi.org/10.1016/j.cognition.2019.104013

Ruitenberg, M. F. L., Duthoo, W., Santens, P., Notebaert, W., & Abrahamse, E. L. (2015). Sequential movement skill in Parkinson's disease: A state-of-the-art. *Cortex*, *65*, 102–112. https://doi.org/10.1016/j.cortex.2015.01.005

Sadeh, T., Shohamy, D., Levy, D. R., Reggev, N., & Maril, A. (2011). Cooperation between the Hippocampus and the Striatum during Episodic Encoding. *Journal of Cognitive Neuroscience*, *23*(7), 1597–1608. https://doi.org/10.1162/jocn.2010.21549

Sajedin, A., Menhaj, M. B., Vahabie, A.-H., Panzeri, S., & Esteky, H. (2019). Cholinergic Modulation Promotes Attentional Modulation in Primary Visual Cortex- A Modeling Study. *Scientific Reports*, *9*(1), 20186. https://doi.org/10.1038/s41598-019-56608-3

Samuelson, L. K., Smith, L. B., Perry, L. K., & Spencer, J. P. (2011). Grounding Word Learning in Space. *PLOS ONE*, *6*(12), e28095. https://doi.org/10.1371/journal.pone.0028095

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The Future of Memory: Remembering, Imagining, and the Brain. *Neuron*, *76*(4), 677–694. https://doi.org/10.1016/j.neuron.2012.11.001

Schacter, D. L., & Slotnick, S. D. (2004). The Cognitive Neuroscience of Memory Distortion. *Neuron*, *44*(1), 149–160. https://doi.org/10.1016/j.neuron.2004.08.017

Schacter, D. L., & Tulving, E. (1994). *Memory systems 1994*. The MIT Press. https://doi.org/10.7551/mitpress/4545.001.0001

Schipper, M., & Mazor, M. (2025). Confidence in absence as confidence in counterfactual visibility. *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Schreiner, M. R., Feustel, S., & Kunde, W. (2025). Linking actions and memories: Probing the interplay of action-effect congruency, agency experience, and recognition memory. *Memory & Cognition*, *53*(4), 1187–1206. https://doi.org/10.3758/s13421-024-01644-2

Schultz, W. (2002). Getting Formal with Dopamine and Reward. *Neuron*, *36*(2), 241–263. https://doi.org/10.1016/S0896-6273(02)00967-4

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Scimeca, J. M., & Badre, D. (2012). Striatal Contributions to Declarative Memory Retrieval. *Neuron*, *75*(3), 380–392. https://doi.org/10.1016/j.neuron.2012.07.014

Seow, T. X. F., Rouault, M., Gillan, C. M., & Fleming, S. M. (2021). How Local and Global Metacognition Shape Mental Health. *Biological Psychiatry*, *90*(7), 436–446. https://doi.org/10.1016/j.biopsych.2021.05.013

Seymour, B., O’Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J., & Frackowiak, R. S. (2004). Temporal difference models describe higher-order learning in humans. *Nature*, *429*(6992), 664–667. https://doi.org/10.1038/nature02581

Shao, R., & Gentner, D. (2019). Symmetry: Low-level visual feature or abstract relation? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *41*.

Sherman, B. E., Turk-Browne, N. B., & Goldfarb, E. V. (2024). Multiple Memory Subsystems: Reconsidering Memory in the Mind and Brain. *Perspectives on Psychological Science*, *19*(1), 103–125. https://doi.org/10.1177/17456916231179146

Shindou, T., Shindou, M., Watanabe, S., & Wickens, J. (2019). A silent eligibility trace enables dopamine-dependent synaptic plasticity for reinforcement learning in the mouse striatum. *European Journal of Neuroscience*, *49*(5), 726–736. https://doi.org/10.1111/ejn.13921

Sinclair, A. H., Manalili, G. M., Brunec, I. K., Adcock, R. A., & Barense, M. D. (2021). Prediction errors disrupt hippocampal representations and update episodic memories. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(51), e2117625118. https://doi.org/10.1073/pnas.2117625118

Smith, M. L. (1988). Recall of spatial location by the amnesic patient H.M. *Brain and Cognition*, *7*(2), 178–183. https://doi.org/10.1016/0278-2626(88)90028-0

Smith, M. L., & Milner, B. (1981). The role of the right hippocampus in the recall of spatial location. *Neuropsychologia*, *19*(6), 781–793. https://doi.org/10.1016/0028-3932(81)90090-7

Soto, F. A., Vogel, E. H., Uribe-Bahamonde, Y. E., & Perez, O. D. (2023). Why is the Rescorla-Wagner model so influential? *Neurobiology of Learning and Memory*, *204*, 107794. https://doi.org/10.1016/j.nlm.2023.107794

Spanò, G., Pizzamiglio, G., McCormick, C., Clark, I. A., De Felice, S., Miller, T. D., Edgin, J. O., Rosenthal, C. R., & Maguire, E. A. (2020). Dreaming with hippocampal damage. *eLife*, *9*, e56211. https://doi.org/10.7554/eLife.56211

Spens, E., & Burgess, N. (2024). A generative model of memory construction and consolidation. *Nature Human Behaviour*, *8*(3), 526–543. https://doi.org/10.1038/s41562-023-01799-z

Spicer, S. G., Mitchell, C. J., Wills, A. J., & Jones, P. M. (2020). Theory protection in associative learning: Humans maintain certain beliefs in a manner that violates prediction error. *Journal of Experimental Psychology: Animal Learning and Cognition*, *46*(2), 151–161. https://doi.org/10.1037/xan0000225

Spivey, M. J. (2023). Cognitive Science Progresses Toward Interactive Frameworks. *Topics in Cognitive Science*, *15*(2), 219–254. https://doi.org/10.1111/tops.12645

Squire, L. R. (1986). Mechanisms of Memory. *Science*, *232*(4758), 1612–1619. https://doi.org/10.1126/science.3086978

Squire, L. R. (1992). Declarative and Nondeclarative Memory: Multiple Brain Systems Supporting Learning and Memory. *Journal of Cognitive Neuroscience*, *4*(3), 232–243. https://doi.org/10.1162/jocn.1992.4.3.232

Squire, L. R., & Zigmond, M. J. (2003). *Fundamental neuroscience.* (2nd ed. / edited by Larry R. Squire ... [et al.]). Academic.

Stephen, D., & Dixon, J. (2009). The Self-Organization of Insight: Entropy and Power Laws in Problem Solving. *The Journal of Problem Solving*, *2*(1). https://doi.org/10.7771/1932-6246.1043

Stillesjö, S., Karlsson Wirebring, L., Andersson, M., Granberg, C., Lithner, J., Jonsson, B., Nyberg, L., & Wiklund-Hörnqvist, C. (2021). Active math and grammar learning engages overlapping brain networks. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(46), e2106520118. https://doi.org/10.1073/pnas.2106520118

Straube, B. (2012). An overview of the neuro-cognitive processes involved in the encoding, consolidation, and retrieval of true and false memories. *Behavioral and Brain Functions*, *8*(1), 35. https://doi.org/10.1186/1744-9081-8-35

Summerfield, C. (2023). *Natural general intelligence: How understanding the brain can help us build AI.* Oxford University Press.

Sun, D., Custers, R., Marien, H., & Aarts, H. (2020). Ideomotor Action: Evidence for Automaticity in Learning, but Not Execution. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00185

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning an introduction. In *Reinforcement learning an introduction*. MIT Press.

Swallow, K. M., Broitman, A. W., Riley, E., & Turker, H. B. (2022). Grounding the Attentional Boost Effect in Events and the Efficient Brain. *Frontiers in Psychology*, *13*, 892416. https://doi.org/10.3389/fpsyg.2022.892416

Swallow, K. M., & Jiang, Y. V. (2013). Attentional Load and Attentional Boost: A Review of Data and Theory. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00274

Symons, C. S., & College, H. (1997). The Self-Reference Effect in Memory: A Meta-Analysis. *Psychonomic Bulletin*, *121*(No. 3), 371–394.

Sziklas, V., Lebel, S., & Petrides, M. (1998). Conditional associative learning and the hippocampal system. *Hippocampus*, *8*(2), 131–137. https://doi.org/10.1002/(SICI)1098-1063(1998)8:2<131::AID-HIPO5>3.0.CO;2-O

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., … Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476–6483. https://doi.org/10.1073/pnas.1916903117

Thomas, A. K., Lee, M., & Hughes, G. (2016). Introspecting on the Elusive: The Uncanny State of the Feeling of Knowing. In J. Dunlosky & S. (Uma). K. Tauber (Eds.), *The Oxford Handbook of Metamemory* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199336746.013.16

Timms, M. A. C., Mandel, D. R., & Nelson, J. D. (2024). Applying Information Theory to Validate Commanders' Critical Information Requirements. In *Handbook of Military and Defense Operations Research* (2nd ed.). Chapman and Hall/CRC.

Toh, Y. N., & Lee, V. G. (2022). Response, rather than target detection, triggers the attentional boost effect in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *48*(1), 77–93. https://doi.org/10.1037/xhp0000977

Tsien, J. Z., Huerta, P. T., & Tonegawa, S. (1996). The essential role of hippocampal CA1 NMDA receptor–dependent synaptic plasticity in spatial memory. *Cell*, *87*(7), 1327–1338.

Tsuji, N., & Imaizumi, S. (2022). Sense of agency may not improve recollection and familiarity in recognition memory. *Scientific Reports*, *12*(1), 21711. https://doi.org/10.1038/s41598-022-26210-1

Tulving, E. (1985a). How many memory systems are there? *American Psychologist*, *40*(4), 385–398. https://doi.org/10.1037/0003-066X.40.4.385

Tulving, E. (1985b). Memory and consciousness. *Canadian Psychology / Psychologie Canadienne*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

Vakil, E., Schwizer Ashkenazi, S., Nevet-Perez, M., & Hassin-Baer, S. (2021). Implicit sequence learning in individuals with Parkinson's disease: The added value of using an ocular version of the serial reaction time (O-SRT) task. *Brain and Cognition*, *147*, 105654. https://doi.org/10.1016/j.bandc.2020.105654

Valentin, S., Kleinegesse, S., Bramley, N. R., Seriès, P., Gutmann, M. U., & Lucas, C. G. (2024). Designing optimal behavioral experiments using machine learning. *eLife*, *13*, e86224. https://doi.org/10.7554/eLife.86224

Van Dam, W. O., Rueschemeyer, S.-A., Bekkering, H., & Lindemann, O. (2013). Embodied Grounding of Memory: Toward the Effects of Motor Execution on Memory consolidation. *Quarterly Journal of Experimental Psychology*, *66*(12), 2310–2328. https://doi.org/10.1080/17470218.2013.777084

Vandist, K., De Schryver, M., & Rosseel, Y. (2009). Semisupervised category learning: The impact of feedback in learning the information-integration task. *Attention, Perception, & Psychophysics*, *71*(2), 328–341. https://doi.org/10.3758/APP.71.2.328

Varela, F. J., Rosch, E., & Thompson, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. The MIT Press. https://doi.org/10.7551/mitpress/6730.001.0001

Wang, J. X., Kurth-Nelson, Z., Tirumala, D., Soyer, H., Leibo, J. Z., Munos, R., Blundell, C., Kumaran, D., & Botvinick, M. (2017). *Learning to reinforcement learn* (arXiv:1611.05763). arXiv. https://doi.org/10.48550/arXiv.1611.05763

Ward, E. J., & Scholl, B. J. (2015). Inattentional blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin & Review*, *22*(3), 722–727. https://doi.org/10.3758/s13423-014-0745-8

Watson, C. E., Cardillo, E. R., Ianni, G. R., & Chatterjee, A. (2013). Action Concepts in the Brain: An Activation Likelihood Estimation Meta-analysis. *Journal of Cognitive Neuroscience*, *25*(8), 1191–1205. https://doi.org/10.1162/jocn_a_00401

Weisberg, R. W. (2015). Toward an integrated theory of insight in problem solving. *Thinking & Reasoning*, *21*(1), 5–39. https://doi.org/10.1080/13546783.2014.886625

Widrow, B., & Hoff, M. E. (1960). Adaptive Switching Circuits. *1960*, *IRE WESCON Convention Record*, pp. 96–104.

Wilcox, E. G., Ding, C., Sachan, M., & Jäger, L. A. (2024). Mouse Tracking for Reading (MoTR): A new naturalistic incremental processing measurement tool. *Journal of Memory and Language*, *138*, 104534. https://doi.org/10.1016/j.jml.2024.104534

Williams, J. M. G., Ellis, N. C., Tyers, C., Healy, H., Rose, G., & Macleod, A. K. (1996). The specificity of autobiographical memory and imageability of the future. *Memory & Cognition*, *24*(1), 116–125. https://doi.org/10.3758/BF03197278

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *eLife*, *8*, e49547. https://doi.org/10.7554/eLife.49547

Yang, Y., Sibert, C. L., & Stocco, A. (2024). Reliance on Episodic vs. Procedural Systems in Decision-Making Depends on Individual Differences in Their Relative Neural Efficiency. *bioRxiv: The Preprint Server for Biology*, 2023.01.10.523458. https://doi.org/10.1101/2023.01.10.523458

Yannier, N., Hudson, S. E., Koedinger, K. R., Hirsh-Pasek, K., Golinkoff, R. M., Munakata, Y., Doebel, S., Schwartz, D. L., Deslauriers, L., McCarty, L., Callaghan, K., Theobald, E. J., Freeman, S., Cooper, K. M., & Brownell, S. E. (2021). Active learning: "Hands-on" meets "minds-on." *Science*, *374*(6563), 26–30. https://doi.org/10.1126/science.abj9957

Yebra, M., Galarza-Vallejo, A., Soto-Leon, V., Gonzalez-Rosa, J. J., De Berker, A. O., Bestmann, S., Oliviero, A., Kroes, M. C. W., & Strange, B. A. (2019). Action boosts episodic memory encoding in humans via engagement of a noradrenergic system. *Nature Communications*, *10*(1), 3534. https://doi.org/10.1038/s41467-019-11358-8

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416

Yon, D., Zainzinger, V., De Lange, F. P., Eimer, M., & Press, C. (2021). Action biases perceptual decisions toward expected outcomes. *Journal of Experimental Psychology: General*, *150*(6), 1225–1236. https://doi.org/10.1037/xge0000826

Yu, Z., Guindani, M., Grieco, S. F., Chen, L., Holmes, T. C., & Xu, X. (2022). Beyond t test and ANOVA: Applications of mixed-effects models for more rigorous statistical analysis in neuroscience research. *Neuron*, *110*(1), 21–35. https://doi.org/10.1016/j.neuron.2021.10.030

Zeithamova, D., & Bowman, C. R. (2020). Generalization and the hippocampus: More than one story? *Neurobiology of Learning and Memory*, *175*, 107317. https://doi.org/10.1016/j.nlm.2020.107317

Zgonnikov, A., Aleni, A., Piiroinen, P. T., O'Hora, D., & Di Bernardo, M. (2017). Decision landscapes: Visualizing mouse-tracking data. *Royal Society Open Science*, *4*(11), 170482. https://doi.org/10.1098/rsos.170482

Zimmerman, D. W. (2004). A note on preliminary tests of equality of variances. *The British Journal of Mathematical and Statistical Psychology*, *57*(Pt 1), 173–181. https://doi.org/10.1348/000711004849222