Temporal Processes in Perceptual and Collaborative Decision-making

by

Xiang Li

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Psychology

New York University

September, 2024

Wei Ji Ma

With courage, press on; then find your rest,

And fulfill the greatness of your life's quest.

Kidney Band, *Late Spring*

Acknowledgements

I would like to extend my deepest gratitude to my advisor, Wei Ji Ma. Before starting graduate school, I didn't fully appreciate the importance of having a great supervisor. When the pandemic struck and made everything more challenging, his patience and care were pivotal to the completion of my PhD. I find myself incredibly fortunate to have him as a brilliant scientist to work with, a considerate mentor who fosters a supportive lab environment, and a role model and teacher advocating for social good both inside and outside of academia. Thank you for everything you have done for me in the six years we have worked together.

I would not have been able to pursue my PhD at NYU without the massive support I received during my undergraduate years. In my sophomore year, when I was curious about research but afraid to actually step into a lab, it was Yan Bao and Ernst Pöppel who opened the door to psychological research for me, providing the extensive training needed for a young scientist. As I became more interested in modeling research, Hang Zhang helped me develop a modeling project on working memory with collaborators and generously guided me in mathematical modeling. In the summer of my junior year, Mehrdad Jazayeri offered me to intern at MIT, where I gained invaluable exposure to research life in the United States.

Then, I would like to acknowledge the members of the Ma lab. First, I want to thank Luigi for collaborating on the project of approximate inference and for being a second mentor to me. I am also grateful to my wonderful labmates: Hsin-hung Li, Heiko Schütt, Dongjae Kim, Aspen Yoo, Andra Mihali, Jennifer Laura Lee, Ionatan Kuperwajs, Peiyuan Zhang, Daisy Lin, Nastaran

Arfaei, Jeroen Olieslagers, Jordan Lei, Yichen Li, Qixiu Fu, Dongqi Bao, Sixuan Chen, and Joshua Calder-Travis. I am privileged to have shared both the academic and social life with you all.

Next, I want to thank the amazing members of my dissertation committee, Brenden Lake, Bob Rehder, Michael Landy, and Natalia Vélez, for their precious guidance and feedback on the work presented here and their care throughout my PhD life.

I am also immensely thankful to my friends who supported me during my graduate years: those from high school (Jiaxin Wei, Huakai Liu, Hao Zhu, Di Tong, Yalun Feng, Mike Li, and Xinyi Fang), my undergraduate friends (Wan Li, Ying Tong, Han Su, Yihan Liu, Haoyu Li, Haoxue Fan, Jingyi Zhou, Jiameng Wang, Letian Chen, Chen Zhao, Yiwen Sun, Yu Wang, Muzhi Wang, Xiangjuan Ren, Tianyuan Teng, Xiaoxiong Lin, Lingyan Wang, Hanying Leng, Xinyu Gao, Ruoming Sun, and Minghao Luo), my colleagues from this department (Fangfang Hong, Hormet Yiltiz, Omri Raccah, Danyang Han, Xiao Ma, Feng Cheng, Shutian Xue, Jingwen Li, and Yunzi Ding), and my poker friends (Junda An, Xiuyuan Zhao, Pan Tan, Siyuan Fang, and Shida Wu). Thank you for enlightening my life during the otherwise monotonous pandemic days. I apologize if I have missed anyone here, as it is an impossible list for me to finish.

Next, I would like to thank Xiang Nan, who has been by my side for the past five years. I am grateful that we have learned how to communicate and express love for each other. Without your unwavering support, especially during the stressful final year of my PhD, my graduate life would have been much harder.

Lastly, I would like to thank my family, especially my parents Xinwen Li and Shuiying Ma. Though neither had the chance to attend college, they did everything to provide me the opportunity to attend the best university in mainland China and pursue graduate school. While they may bot be perfect, I am touched by how they have continuously learned to befriend and listen to me over the years. I am deeply indebted to their kindness and support throughout my graduate life. I love you so much.

Abstract

Humans make decisions every day, and the ability to make accurate decisions is crucial for navigating an ever-changing world. These decisions range from simple perceptual tasks, such as recognizing traffic lights while driving, to complex social decisions like planning a family vacation together. While many studies have explored the outcomes and contributing factors of decision-making, fewer have delved into the underlying temporal processes.

In this dissertation,we address this gap by examining the temporal processes involved in two distinct sub-fields of psychology: perceptual decision-making and collaborative decision-making. Chapter 1 provides a general introduction to approximate inference through active sampling. In Chapter 2, we develop a novel framework of approximate inference through active sampling and apply it to the task of perceptual categorization. Chapter 3 offers a broad introduction to the cognitive mechanisms of collaboration. In Chapter 4, we develop a two-player real-time collaborative game to investigate the temporal processes of collaboration, with a special focus on role coordination. Chapter 5 presents a Bayesian linear extrapolation model to explain the temporal processes of role coordination observed in the previous chapter. Together, my work enhances our understanding of the temporal processes involved in human decision-making.

Contents

List of Figures

List of Tables

1 Approximate Inference through Active Sampling: an Introduction

Imagine you are exploring the basement of your family house when a picture of a young woman catches your eye. Who is she? Could it be your mother, one of your aunts, your grandmother, or someone else entirely? The lighting is not dim, and the picture is clear, yet you still deliberate over the woman's identity because many of your female relatives bear a resemblance to the picture. Gradually, you eliminate some possibilities and eventually decide that it is your mother. You then go upstairs to ask if anyone knows the answer. During this process, you may feel a sense of curiosity mixed with uncertainty, leading you to recall impressions of your female relatives and contextual information about your family members. The setting of the basement and the age of the photo might also influence your reasoning, as older photos might limit the pool of possible individuals.

Now, let's step back from this scenario and abstract the process you just experienced. You received perceptual stimuli from the environment (the old picture) and tried to make a decision based on it (identifying the woman). This process is known as perceptual inference, and your deliberation on the woman's identity illustrates the temporal processes involved in making such a decision. Perceptual inference is crucial for survival and daily functioning, enabling us to navigate complex and dynamic environments. More daily examples include comprehending the sound of a friend in a noisy restaurant, recognizing traffic lights while driving, or smelling leftover food

to determine if it is spoiled.

Just as the examples above illustrates, the state of the world is typically unknown to agents. Therefore, reasoning about hidden states of the world based on noisy or ambiguous observations is central to cognition (al-Haytham, 1989). The leading theory of how humans perform such inference is Bayesian inference, whereby the brain computes probability distributions over world state variables of interest via Bayes' rule (Knill & Richards, 1996).

The hypothesis that the brain implements Bayesian inference accounts for a wide range of cognitive abilities. These abilities range from perception, where the perceptual system combines information with varying reliability levels (Ma et al., 2023); motor control, where the motor system integrates cues from different modalities (Körding & Wolpert, 2004); and language learning, where children infer the meaning of a new word from several examples (Xu & Tenenbaum, 2007). It also extends to everyday reasoning, such as estimating the remaining duration of a movie that has been playing for 40 minutes (Griffiths & Tenenbaum, 2006).

In Bayesian accounts, a key concept is the *likelihood* of a world state, which is equal to the probability of the current observations x given that world state. The theory commonly assumes that the brain calculates the likelihood $p(x|\text{world state})$ based on the learned statistical structure of the world and derives the posterior probability $p(\text{world state}|x)$, which is proportional to the prior $p(\text{world state})$ times the likelihood $p(x|\text{world state})$:

$$p(\text{world state}|x) \propto p(\text{world state}) \times p(x|\text{world state})$$

However, this assumption seems unreasonable, as high-precision neural computations are energetically costly (S. Laughlin, 1981; Lennie, 2003). A more plausible and *resource-rational* (Griffiths et al., 2015) approach is that the brain produces less costly, imprecise, and noisy estimates of the likelihood or posterior probability, refining them only when necessary.

In line with this thought, recent studies have found that Bayesian inference is not always con-

ducted precisely; instead, approximate inference often occurs, particularly when the task structure becomes complex (Acerbi et al., 2014; Beck et al., 2012; Drugowitsch et al., 2016; Findling et al., 2019; Lieder et al., 2018; Vul et al., 2014). Several possibilities might contribute to the approximation. Beck et al. (2012) and Herce Castañón et al. (2019) discuss scenarios where agents use incorrect parameters during inference, leading to biased results. Therefore, the approximate inference might be the camouflage of biased inference. Drugowitsch et al. (2016), Findling et al. (2019), and Lange et al. (2021) suggest that the mental inference process itself (likelihood calculation) may be noisy, even after accounting for perceptual and decision noise (Drugowitsch et al., 2016). Additionally, S. Gershman and Goodman (2014) attribute approximate inference to amortized inference, where humans reuse previous relevant history of inferences for new tasks.

Furthermore, S. J. Gershman et al. (2012), Haefner et al. (2016), Lange et al. (2021), Lieder et al. (2018), Rullán Buxó and Savin (2021), Sanborn and Griffiths (2007), Savin et al. (2014), Shi and Griffiths (2009), Shivkumar et al. (2018), Vul et al. (2009), and Vul et al. (2014) proposed that likelihoods or posteriors in the Bayesian inference are approximated using sampling methods like Markov chain Monte Carlo (Metropolis et al., 1953) or particle filtering (Khan et al., 2004). Rather than calculating the likelihood or posterior probabilities precisely, humans sample around the true values. By incorporating approximate inference, researchers have explained various cognitive phenomena, including anchoring bias (Lieder et al., 2018), and recency and primacy effects (Lange et al., 2021), across different domains, from multi-object tracking (Vul et al., 2009) to probability induction (S. Gershman & Goodman, 2014). Some studies also explore the potential neural mechanisms of approximate inference in the brain (Haefner et al., 2016; Rullán Buxó & Savin, 2021; Savin et al., 2014; Shi & Griffiths, 2009; Shivkumar et al., 2018).

One practical advantage of the frameworks of sampling-based approximate inference is their easy adaptability in explaining the temporal processes occurring within approximate inference. Sampling-based approaches, such as Markov Chain Monte Carlo and particle filtering, allow for a dynamic and iterative process where the agent continuously refines its estimates based on new

samples. This iterative process inherently incorporates time as a factor, making it well-suited to study temporal processes of decision-making. By focusing on how samples are generated and refined over time, researchers can gain a deeper understanding of the cognitive processes involved in making perceptual decisions. In contrast, other frameworks such as primarily attributing approximate inference to noise in mental inference processes often provide a static snapshot of the decision-making outcome (Drugowitsch et al., 2016), without capturing the evolution of these decisions over time. They may explain why certain decisions are made but not how the decision-making process unfolds temporally. Thus, adopting a sampling-based approach not only aligns with empirical observations of how decisions are made but also offers a robust framework for exploring the dynamic nature of perceptual inference. This is the reason we adopt the framework of approximate inference through active sampling in the Chapter 2.

Let's focus on the sampling-based method. If the underlying mechanism of inference indeed works based on sampling, what would be an effective way of generating these samples? The philosophy behind active sampling is that in a task where an agent needs to iteratively sample from different sources, optimal performance is achieved by sampling in a way that rapidly reduces uncertainty. This method has been applied in value-based decision-making, explaining the choices and eye movements of participants when selecting from multiple valuable objects (Callaway et al., 2021; Jang et al., 2021; Z.-W. Li & Ma, 2021; Song et al., 2019).

Active sampling is similar to active sensing (Yang, Lengyel, & Wolpert, 2016; Yang, Wolpert, & Lengyel, 2016), which is defined as the process by which agents sample from unexplored areas in the physical world to maximize perceptual information gain. Both active sampling and active sensing reduce reaction time and improve task performance. However, they differ conceptually. Active sensing acquires new information by interacting with the physical world (e.g., orientation information in a different area), while active sampling decreases uncertainty by refining the agent's internal belief through the allocation of abstract computational resources. Variability in active sensing arises from the uneven distribution of perceptual information in the search field,

whereas in active sampling, it stems from the imperfect representation of belief in the brain.

A natural question that arises is how to actively sample, i.e., determining where it is most worthwhile to sample from. After assuming a sampling cost and reward for the task, this problem can be addressed by the Bellman equation and dynamic programming, but it is generally computationally intractable (Sutton & Barto, 2018). To manage this intractability, researchers approximate optimal sampling decisions using different methods, such as feature engineering (Callaway et al., 2021) or setting a finite bound so that the sampling process must stop after a preset number of samples (Jang et al., 2021). Another group of researchers bypasses this question by assuming humans use heuristics to decide where to sample (Z.-W. Li & Ma, 2021; Song et al., 2019). Both approaches have achieved some success in accounting for choice and eye movement data in value-based decision-making.

Lastly, we discuss the neural mechanisms of approximate inference. Numerous neural network models have been proposed to explain this complex process. Shi and Griffiths (2009) introduced a simple network structure comprising a few feature detection neurons based on importance sampling. Savin et al. (2014) developed a Hopfield network (Hopfield, 1982) that efficiently uncovers hidden states, despite its primary focus on memory. Haefner et al. (2016) demonstrated how feedback can be incorporated into neural network models and predicted how different types of feedback would influence approximate inference. Additionally, Rullán Buxó and Savin (2021) proposed a spiking network model where the output of neurons can be readily used for decision-making. In summary, these neural network models of approximate inference explain behavioral results and provide insights into the possible neural underpinnings of approximate inference.

In the next chapter of this dissertation, we will develop a novel framework of approximate inference through iterative active sampling. We will specify the structure of this framework and apply it to three different datasets, accounting for phenomena including choice, confidence, and response time in perceptual decision-making. The chapter aims to deepen our understanding of the temporal processes underlying perceptual decision-making by casting approximate inference

as an active-sampling process.

2 Approximate Inference through Active Sampling of Likelihoods Accounts for Human Categorization Behavior

Abstract

Bayesian computations are intractable and expensive, but this is rarely accounted for in existing Bayesian observer models. In this work, we propose that a) the brain can only compute imprecise (noisy) estimates of likelihoods and posteriors, and, b) since computations are expensive, the brain actively chooses which computations to perform to refine such estimates. We call our framework approximate inference through active sampling (AIAS) and study its implications in N-alternative categorization. AIAS accounts for several empirical findings. First, we account for a puzzling recent finding that decision confidence follows the difference between the two highest posteriors, rather than the highest posterior itself. AIAS not only provides better fits, but also yields an accurate prediction of response times based on the number of iterations. Second, we show that AIAS is able to explain how categorization behavior changes when the visual contrast varies. Third, we find that the mean response times predicted by AIAS grows approximately

logarithmically with the number of categories N, as per Hick's law. Overall, AIAS provides a novel approach to explain human categorization by casting approximate inference as an active-sampling process with imprecise computations.

Keywords: categorization, approximate inference, active sampling, Hick's law, confidence, response time

2.1 Introduction

The state of the world is typically unknown to agents. Therefore, reasoning about hidden states of the world based on noisy or ambiguous observations is central to cognition (al-Haytham, 1989). The leading theory of how humans perform such inference is the theory of Bayesian inference, whereby the brain computes probability distributions over world state variables of interest via Bayes' rule (Knill & Richards, 1996). The hypothesis that the brain implements Bayesian inference accounts well for data across a wide range of cognitive abilities, ranging from perception (Ma et al., 2023), motor control (Körding & Wolpert, 2004), language learning (Xu & Tenenbaum, 2007), to everyday reasoning (Griffiths & Tenenbaum, 2006). In Bayesian accounts, a key concept is the *likelihood* of a world state, which is equal to the probability of the current observations given that world state. The theory commonly assumes that the brain calculates the likelihood *exactly* based on the learned statistical structure of the world. However, this seems an unreasonable assumption given that high-precision neural computations are costly (S. Laughlin, 1981; Lennie, 2003). Rather, a more realistic and *resource-rational* (Griffiths et al., 2015) scenario is that the brain would compute cheaper imprecise, noisy estimates of the likelihood, refining them only as needed. Indeed, some evidence exists for noisy (Acerbi et al., 2014; Drugowitsch et al., 2016; Findling et al., 2019; Herce Castañón et al., 2019; Lange et al., 2021) or sample-based approximations (S. J. Gershman et al., 2012; Lange et al., 2021; Lieder et al., 2018; Vul et al., 2014). In this work, we commit to this notion of imprecise likelihood computations and ask: how should the brain perform Bayesian inference, when each computation step only provides a noisy estimate of the likelihood? That is, we abstractly model inference as an iterative process that starts with a crude approximation and proceeds by refining the computation until the costs outweigh the benefits.

Fig. 2.1 shows the basic intuition behind our framework. In an idealized world, computing the likelihood is not costly and therefore an agent, given the noisy sensory evidence, instantly

computes the exact likelihood over all hypothesized stimuli (Fig. 2.1A). However, it is more realistic that the brain can only compute imprecise (noisy) estimates of likelihoods and posteriors (Fig. 2.1B). When the likelihood computation is imprecise, the true likelihood function becomes unknown and must be recovered. We model this process as the agent iteratively drawing likelihood samples (Fig. 2.1B, top row). Using these imprecise samples, the true and unknown likelihoods are estimated with uncertainty (Fig. 2.1B, bottom row). Differently from other frameworks, the brain would actively decide where the new likelihood sample should be drawn from. As the computation process progresses (Fig. 2.1B, from left to right), more likelihood samples are actively drawn and aggregated in each iteration, gradually refining the belief over the likelihood and converging to the true likelihood. As the active process of shaping the approximation to the likelihood is key to this framework, we will refer to it as Approximate Inference with Active Sampling (AIAS).

This paper describes the formalism of AIAS and tests it with published human data from different categorization tasks (Adler & Ma, 2018; H.-H. Li & Ma, 2020). AIAS-based models predict response time, choice, and confidence with few free parameters. AIAS successfully accounts for various empirical findings in perceptual categorization, covering a phenomenon where decision confidence is governed by the two most probable choices across categories (H.-H. Li & Ma, 2020), a perceptual task where visual reliability varies (Adler & Ma, 2018), and Hick's law, which describes how mean response time changes with the number of alternatives (Hick, 1952). Collectively, these results provide evidence that humans may conduct approximate inference as an active-sampling process with imprecise likelihood computations.

2.2 Results

In this section, we first describe the framework of Approximate Inference with Active Sampling (AIAS), within the context of perceptual categorization tasks. We then apply AIAS to three dis-

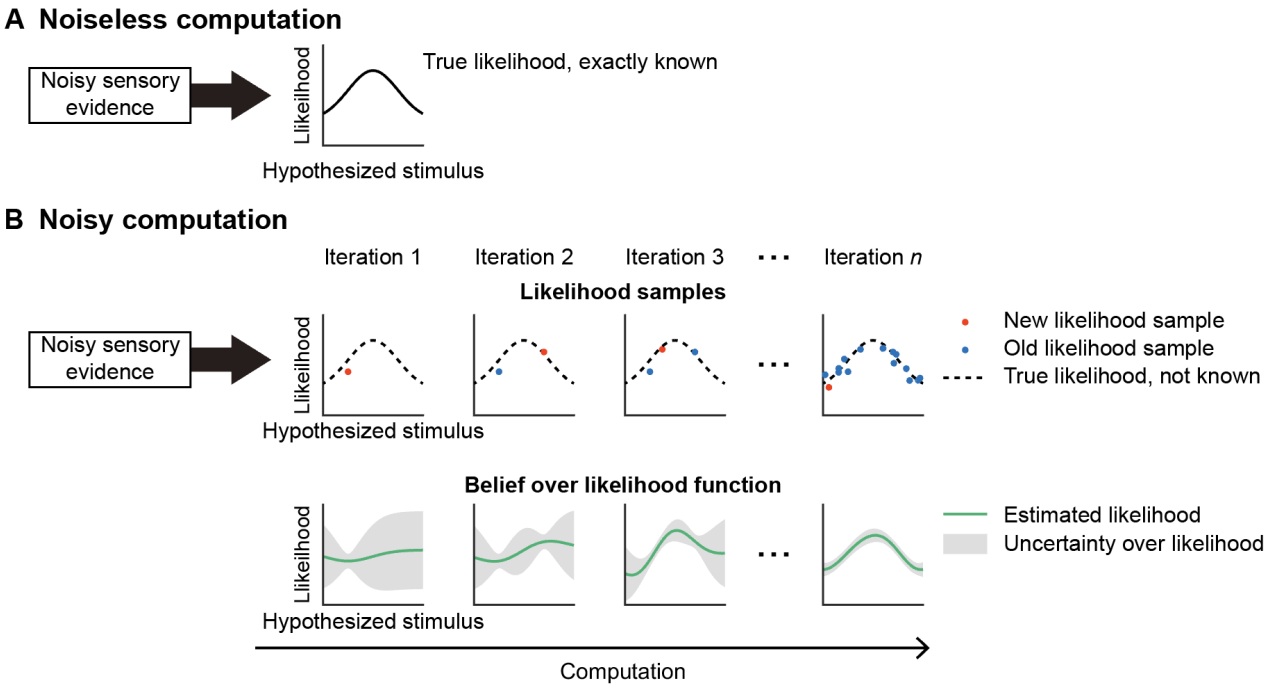

Figure 2.1: The illustrative plot of Approximate Inference with Active Sampling (AIAS) framework. A) When computing the likelihood is not costly and therefore precise, the true likelihood over all hypothesized stimuli are instantly calculated, based on the noisy sensory evidence. B) Likelihoods are approximated by actively drawing likelihood samples, when only imprecise estimates of likelihoods and posteriors can be computed. Top row: With noisy sensory evidence as input, at each iteration, a new noisy likelihood sample (red dot) is actively drawn, centered around the true and unknown likelihood (dashed line). All previous likelihood samples are shown as blue dots. Bottom row: The belief over the likelihood function is inferred from all likelihood samples, with the mean estimate as green line and uncertainty level in grey area. As the computation process progresses (from left to right), more likelihood samples are drawn, and the belief over the likelihood is refined and converges to the true likelihood.

tinct case studies, demonstrating its ability to explain a variety of perceptual and metacognitive findings.

2.2.1 AIAS-based Model for Categorization Tasks

Generative Model. We consider an N-alternative categorization task under uncertainty. In this section, we take $N = 3$ for simplicity of exposition (Fig. 2.2A). In a standard Bayesian observer model (Ma et al., 2023), the agent would receive a potentially noisy observation x and compute its likelihood $L_1 = p(x|C = 1)$, $L_2 = p(x|C = 2)$, $L_3 = p(x|C = 3)$. This likelihood would be then combined with the prior to compute the posterior over categories using Bayes' rule.

So far, this is textbook Bayesian inference. The key idea of this paper is that this standard approach ignores the fact that the human mind is unlikely to compute likelihoods L exactly, especially if the generative model is complex. That is to say, the true likelihoods L_1, L_2, L_3 and the true posterior vector $\mathbf{P} \triangleq (P_1, P_2, P_3)$ are unknown to the agent (Fig. 2.2B,C). Instead, here we consider *imprecise* computations, where the agent has only access to noisy samples l of the likelihood function. In other words, we abstract the process of computing the likelihood – and, thus, the posterior – as an iterative process whereby the agent refines their likelihood estimate by drawing noisy samples of the likelihoods (Fig. 2.2C), interpreting each sample as a small amount of computation. Crucially, the agent chooses which estimates need refining. At each time step, the agent can draw a noisy likelihood sample l_i^k for a chosen category i, where k is the position in the sequence of likelihood samples for that category. We remark that this process is not sampling information from the environment, but sampling is used as a representation of imprecise internal computations (see also Discussion).

Inference. Our generative model casts Bayesian inference as an inference process itself, where the agent aims to infer (compute) the true posterior given a set of noisy likelihood evaluations (Acerbi, 2020; Hennig et al., 2015). In particular, given a set of likelihood samples and a prior over the likelihoods, the agent is able to compute their belief about the true likelihoods, i.e. the

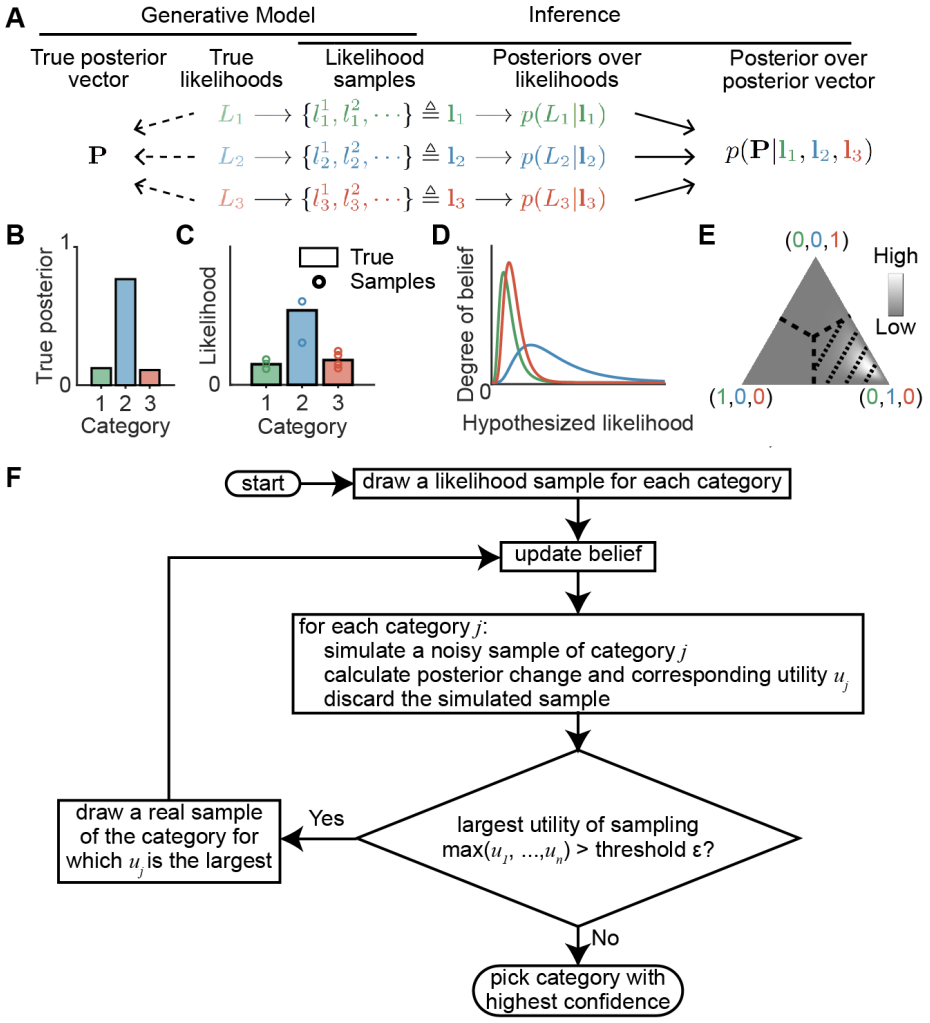

Figure 2.2: AIAS-based model for categorization tasks. We depict a three-alternative categorization problem for illustrative purposes. Colors correspond to the three categories. A) Flow chart of the generative model and inference process. The true likelihoods and posterior are unknown to the agent. B-C) Example true likelihoods, the corresponding true posterior vector, and the agent's noisy likelihood samples that have been drawn. D) Posterior distributions over the true likelihood of each category. E) The posterior distribution over the posterior vector. Each point corresponds to a hypothesized posterior vector. The intensity of the point represents the degree of belief. The shaded area weighted by the posterior over posterior vector is the confidence if the agent were to choose category blue for categorization. F) Flowchart of the active sampling, termination, and decision parts of AIAS.

posteriors over the likelihoods $p(L_1|\mathbf{l}_1), p(L_2|\mathbf{l}_2), p(L_3|\mathbf{l}_3)$ (Fig. 2.2D), using $p(L_i|\mathbf{l}_i) \propto p(L_i)p(\mathbf{l}_i|L_i)$, where $\mathbf{l}_i \triangleq (l_i^1, l_i^2, \cdots)$ is the vector of all likelihood samples of category i. They then again apply Bayes' rule to compute their belief about the true posterior vector, i.e. the posterior over the posterior vector $p(\mathbf{P}|\mathbf{l})$ (Fig. 2.2E), where $\mathbf{l} \triangleq \{\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3\}$ is the set of all likelihood samples of all categories. For computational tractability, in this paper we assume a factorized empirical log-normal distribution for the prior over each likelihood, $p(\log L_i) = \mathcal{N}(\log L_i; \mu_L, \sigma_L^2)$, log-normal noise on the likelihood samples $p(\log l_i^k|L_i) = \mathcal{N}(\log l_i^k; \log L_i, \sigma^2)$, and a uniform prior over categories (see Methods section). In the Discussion, we consider extensions of these baseline assumptions.

Given the posterior over the posterior vector, $p(\mathbf{P}|\mathbf{l})$, we further define the agent's confidence c_i as the subjective belief that the chosen category i has the highest true posterior probability when making a decision for categorization, which we name "comparative confidence". For example, in Fig. 2.2E, if the agent chooses the blue category, the shaded area weighted by $p(\mathbf{P}|\mathbf{l})$ will be their confidence. We will revisit the intuition of this concept in the model fitting section.

Active Sampling, Termination, and Decision. Considering that neural computations are costly in terms of time and metabolic resources (S. B. Laughlin et al., 1998; Lennie, 2003), drawing likelihood samples equally from all categories is a wasteful strategy. Instead, the agent should actively prioritize sampling from more likely categories. In a similar vein, the agent cannot sample indefinitely, but must decide when to terminate sampling and make a categorization decision to receive a reward. The choice of which category to sample from (active sampling), when to stop (termination), and which category to select at the end (decision) can be formalized as a sequential decision making problem (Garnett, 2023; Sutton & Barto, 2018). Optimally solving this problem requires considering all future steps and outcomes, which is computationally intractable in the general case (Sutton & Barto, 2018). In this paper, we consider a *myopic* policy that contemplates only one step ahead when deciding which category to sample from and when to stop sampling.

The active sampling, termination, and decision parts of AIAS are shown diagrammatically in Fig. 2.2F. We assume the agent starts with no or vague information about the categories, so

they first draw a likelihood sample for each category. Then, after updating the belief given the likelihood samples obtained so far, the agent considers each category as a candidate for drawing a likelihood sample for the next iteration. The goal is to estimate the expected utility for refining the likelihood of category j, where the utility is some metric related to the information gain about which category is most probable when it comes to making a decision, discussed next.

To perform this estimation, separately for each category j, the agent simulates a new, 'phantasized' likelihood sample based on current beliefs (Thompson, 1933). The phantasized likelihood sample leads to a change of the corresponding posterior over the likelihood, a change of the posterior over the posterior vector, and thus a change of the aforementioned confidence c_i for all N categories. We define the utility of such sampling as the absolute change of the largest confidence in all N categories, $u_j = |\Delta(\max(c_1, ..., c_N))|$. After getting the u_j, the agent discards the simulated likelihood sample. After one simulation per category, if the maximum of all utilities u_j $(j = 1, 2, ..., N)$ is **larger** than a predetermined termination threshold ϵ, the agent chooses the category j with the largest u_j to draw an actual likelihood sample. The agent then updates their posterior over the posterior vector and repeats the whole simulation process. Conversely, if the largest utility is **smaller** than the threshold ϵ, the agent stops sampling as further sampling would likely yield little information gain, and chooses the category i with the highest confidence c_i.

The intuition behind this strategy is that categories with potentially high true posteriors, which are important for the categorization, should be sampled more. This is realized in our framework in that categories with high true posteriors will cause a larger absolute change of maximum confidence across all categories, and thus are indeed sampled more frequently. The intuition behind the termination rule is that the agent should stop sampling if they find that sampling more is unlikely to substantially affect their largest confidence value, which is how the final category decision is made.

2.2.2 Multi-alternative Categorization and Decision Confidence

In psychophysical tasks, measures of decision confidence have emerged as crucial behavioral metrics to quantitatively test hypotheses about the underlying computational and metacognitive processes (Fleming, 2024; Kiani & Shadlen, 2009; Yeung & Summerfield, 2012). For this reason, we aim to experimentally validate the AIAS framework on rich datasets that combine both decision and confidence reports. We start by considering data from a recent study (H.-H. Li & Ma, 2020) that thoroughly tested different leading hypotheses about how perceptual confidence is reported. In this experiment (Fig. 2.3A), participants viewed spatial configurations of three colored dot clouds, with a black target dot randomly generated from one of the three clouds. Participants were asked to identify which color cloud the target dot came from and then reported their confidence level using a four-point scale from 1 (very unconfident) to 4 (very confident), while their response time was recorded. The experiment featured four different color cloud configurations (Fig. 2.3E). The two most prominent models in the paper were the Max model (Fig. 2.3B&D, left panel), the leading hypothesis that confidence represents the highest posterior of all categories (Kepecs & Mainen, 2012), and the Difference Model (Fig. 2.3B&D, middle panel), an alternative hypothesis proposed by the authors that confidence represents the difference between the two highest posteriors of all categories. Surprisingly, the authors found that the Difference model provided a better account of confidence by outperforming the Max model in a quantitative model comparison (H.-H. Li & Ma, 2020).

Despite giving a better fit to human confidence, the Difference model has two drawbacks. First, it only gives a quantitative prediction of confidence, without providing any rationale for why it should be the difference of the highest two posterior probability. Second, it is only a descriptive model for choices and confidence that does not specify any underlying process; as a result, the Difference model does not yield any prediction of response time in inference tasks. These issues are addressed in the AIAS-based model.

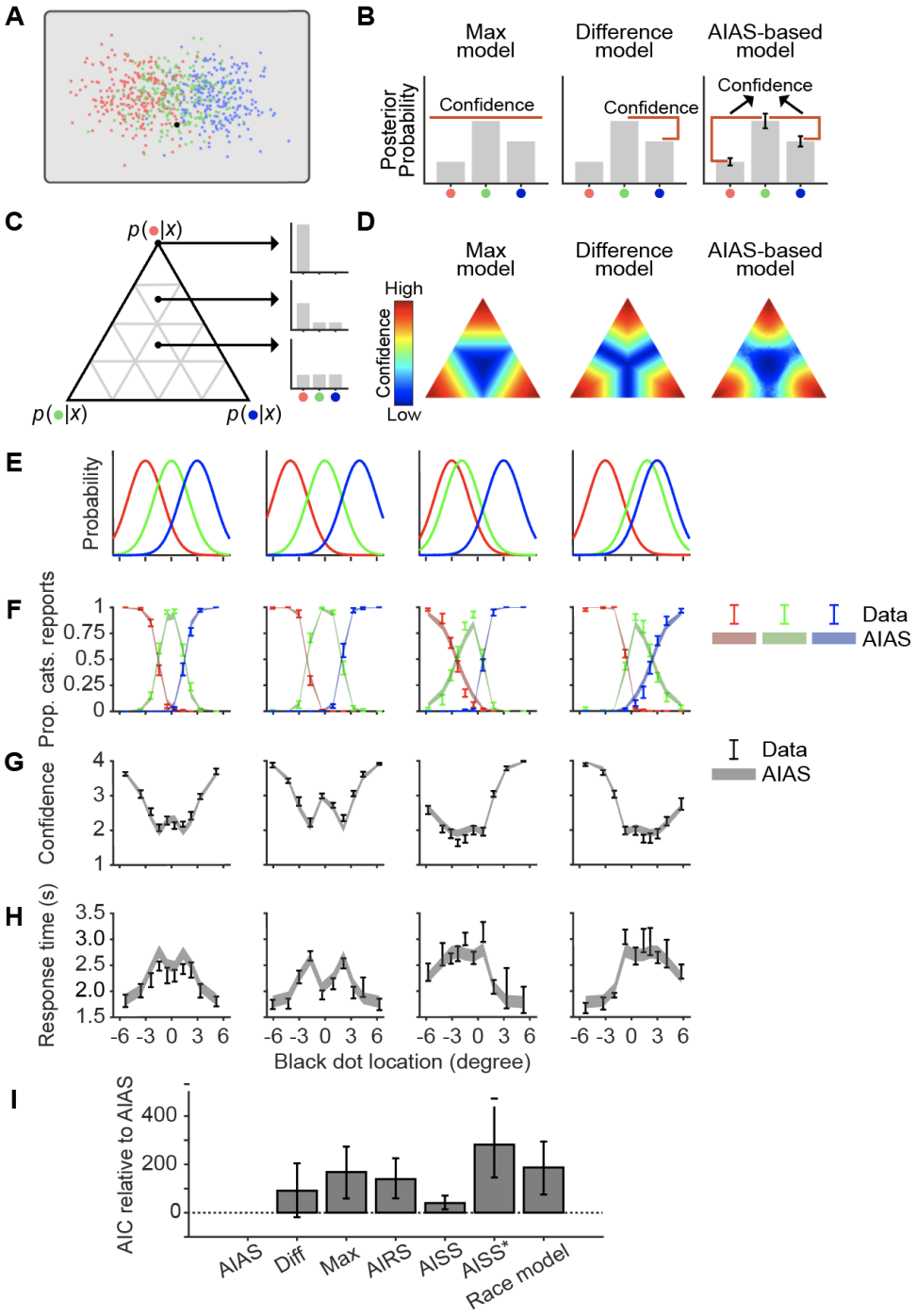

Figure 2.3: The model fitting results and predictions of AIAS-based model in H.-H. Li and Ma (2020)'s color cloud experiment. (Continued on the following page.)

Figure 2.3: A) Experimental paradigm. Participants viewed spatial configurations of three colored dot clouds, with a black target dot randomly generated from one of the three color clouds. Participants were asked to identify which color cloud the target dot came from and then report their confidence level using a four-point scale from 1-"very unconfident" to 4-"very confident", while their response time was recorded. B) Hypotheses of confidence mechanisms in the Max model (left panel, the highest posterior), the Difference model (middle panel, the difference between the two highest posteriors), and AIAS-based model (right panel, the subjective belief that the choice selected has the highest true posterior, a.k.a. comparative confidence). C) We use ternary plots to represent posterior distributions. Each point in the plot correspond to one posterior vector, whose individual elements are proportional to the distance from the opposite edges. For example, a point at the center represents the uniform posterior distribution; at the corners, the posterior probability of one category is one while the other two are zero. D) Predictions of confidence in the Max model (left panel), the Difference model (middle panel), and AIAS-based model (right panel). E) Four different spatial configurations of the three color clouds in the experiment. F) Human choice data (lighter color) and AIAS model fits (darker color). In F-H, error bars and shaded areas represent mean ±1 s.e.m across participants. G) Human confidence report (solid error bar) and AIAS model fits (shaded error bar). H) Human response time (solid error bar) and AIAS model predictions from fitted choice and confidence data (shaded error bar). I) Model comparison (AIC difference) of AIAS vs. Difference, Max, AIRS, AISS, AISS*, and Race models; lower is better. Error bars represent bootstrapped 95% confidence interval across participants.

In AIAS, we have defined the confidence for category j in an intrinsically comparative fashion, by defining c_j as the subjective probability that j is the best possible choice for the current beliefs (i.e., the category with largest posterior probability, with the available information). Note that this definition exploits 'second-order' uncertainty over the posterior, which is a distinctive property of our framework. A standard ideal Bayesian observer model would have no uncertainty about which choice is best with the current information: it is always the category with maximum posterior probability, which is known exactly. To compare the AIAS-based confidence predictions to data, we then simply assume that the reported confidence is the confidence $c_{\hat{j}}$ for the selected choice $\hat{j}$ (Fig. 2.3B&D, right panel).

We fitted AIAS to the color choice and confidence data on an individual basis for all eight participants. The model fitting results and predictions are shown in Fig. 2.3F-H. AIAS captures well both color choice (Fig. 2.3F) and confidence reports (Fig. 2.3G). We then used the number of AIAS iterations required to reach termination to predict response time in each trial. Since AIAS

iterations are expressed in abstract units, and participants' response time data are also affected by non-decision time, we conducted a linear regression mapping the number of iterations to human response times on an individual basis. Remarkably, the AIAS model fitted *on choice and confidence data only* (no response times) provides an accurate prediction for the shape of human response times (Fig. 2.3H).

We further conducted a quantitative model comparison to test AIAS against six main alternative models: 1) the Difference model; 2) the Max model; 3) approximate inference with random sampling (AIRS): A variant of AIAS which, instead of doing active sampling, randomly selects one category to draw a likelihood sample at each iteration; 4) approximate inference with same sampling for all categories (AISS): A variant of AIAS which instead of sampling only one category at each iteration, samples all categories at the same time; 5) approximate inference with same sampling for all categories, but only one time (AISS*): AISS* samples all categories once and makes the decision, with no iterative process involved; 6) the Race model (Ratcliff, 1978; Ratcliff et al., 2016): The Race model is widely used to explain the process of multi-alternative decision-making. The model assumes that evidence is accumulated for each alternative until a decision threshold is reached. Here, we used the true posteriors as the drift rate (the accumulation speed) and allowed for mutual inhibition between different categories. We defined as the confidence the reciprocal of the number of iterations for the Race model as a common practice. The model comparison results are shown in Fig. 2.3I. AIAS outperforms all six alternative models (see Methods for the details of the alternative models).

Finally, we explored the hypothesis that AIAS may be the underlying process at the heart of the empirical finding that the Difference model outperforms the Max model, as observed in H.-H. Li and Ma (2020). To test this hypothesis, we fitted the Difference model and the Max model to synthetic data generated by AIAS. If our hypothesis is correct, we should also observe that the Difference model fits the AIAS-generated data better than the Max model. The model comparison confirms our hypothesis in that the Difference model yields a better fit with a 95% confidence

interval of ΔAIC = [27.2, 129.4]. Overall, these findings show that AIAS is both qualitatively and quantitatively capturing key features of the human decision process.

2.2.3 Inference with Varying Perceptual Reliability

For a further investigation of our framework, we considered another rich dataset comprising of choice and confidence data as well as multiple perceptual reliability conditions (varying visual contrast levels) (Adler & Ma, 2018). In the experiment, participants were asked to perform a binary categorization task based on the orientation of a presented Gabor patch and then report their confidence level on a four-point scale from 1-"low confidence" to 4-"high confidence" (Fig. 2.4A, left panel). Their response time was recorded. The visual contrast of the patch varied across trials (Fig. 2.4A, right panel). Participants performed one of two different tasks, which differed in the distribution of orientation associated with each category (Fig. 2.4B). We fitted AIAS to the category choice and confidence data separately on an individual basis for all eleven participants. AIAS obtained a better quantitative fit than an extended Bayesian observer model including decision noise (see details in the Methods section), ΔAIC = [278.5, 1138.7], yielding a good qualitative fit on choice (Fig. 2.4C) and confidence (Fig. 2.4D) for the different levels of visual contrast. We then used the number of AIAS iterations before termination to predict response time in each trial, via a linear regression mapping the number of iterations to human response times on an individual basis. The AIAS model fitted on choice and confidence data (no response times) provides a reasonable prediction for the shape of human response times, which is somewhat lacking a strong structure in this task (Fig. 2.4E).

2.2.4 Hick's Law

We additionally investigated whether AIAS reproduces a classical psychophysical results for response times. Hick's law states that in a multiple-alternative choice task, the mean response

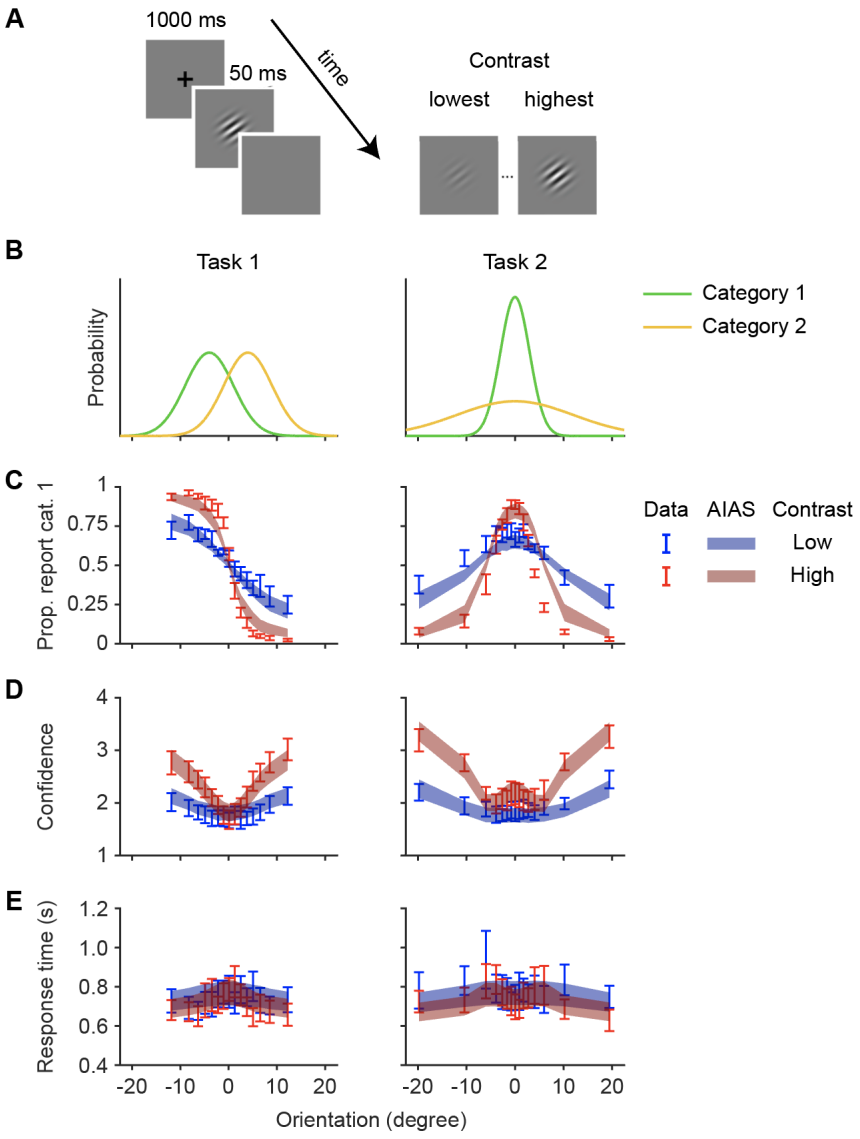

Figure 2.4: The model fitting results and predictions of AIAS-based model in Adler and Ma (2018)'s experiment. A) Experiment paradigm. Left: After fixation, participants were briefly shown a Gabor stimulus. They were required to report the category based on stimulus orientation as well as their confidence level using a four-point scale from 1-"low confidence" to 4-"high confidence". Right: The visual contrast of the Gabor patch varied across trials. B) Probability distribution of the orientation of the Gabor stimulus for the two categories in the two distinct tasks (left: task 1; right: task 2). C) Human choice (error bars) and AIAS model fits (shaded error bars) in the two different tasks (left: task 1; right: task 2) for different visual contrasts (blue: low; red: high). D) Human confidence (error bars) and AIAS model predictions (shaded error bars) in the two different tasks (left: task 1; right: task 2) for different visual contrasts (blue: low; red: high). E) Human response time (error bars) and AIAS model predictions from fitted choice and confidence data (shaded error bars) in the two different tasks (left: task 1; right: task 2) for different visual contrasts (blue: low; red: high). Error bars and shaded areas represent mean ±1 s.e.m across participants.

time increases logarithmically with the number of alternatives (Hick, 1952). In the previous two sections, we have acquired empirical estimates of human participants' parameters and associated linear mappings from the number of iterations to response time. Thus, we can simulate how response time changes with the number of categories N in AIAS for a population of synthetic participants whose model parameters match the ones estimated from the previous two experiments (nineteen synthetic participants in total). For this simulation, we assume that the true likelihoods of different categories are independent in the generative model. That is, our simulation does not commit to a specific task structure, in which the true likelihoods might be correlated.

Simulations of the response times of two example synthetic participants (i.e., parameter sets) are shown in Fig. 2.5A, along with the logarithmic fit and the linear fit of these simulated response times. For all nineteen parameter sets, estimated from all participants of the previous experiment, see Fig. 2.S1. The logarithmic fit qualitatively better captures the trend than the linear fit. To quantitatively validate this result, we computed the R^2 goodness-of-fit score of the logarithmic fit and the linear fit for all nineteen participants (Fig. 2.5B). The logarithmic fit gives a larger R^2 for all participants, meaning that the logarithmic fit better accounts for how the mean response time changes with the number of categories.

2.3 Discussion

In this paper, we developed AIAS, an approximate inference framework with active sampling of likelihoods, to understand how the human brain reasons about hidden states of the world when exact likelihoods are not known. Although it is well-documented that human reasoning appears Bayesian across multiple settings (Griffiths & Tenenbaum, 2006; Körding & Wolpert, 2004; Ma et al., 2023; Xu & Tenenbaum, 2007), many studies have not fully explored the consequences of the intractable and expensive nature of Bayesian inference as a computational process. To address this gap, in AIAS we postulated an agent actively drawing samples of the likelihoods to refine

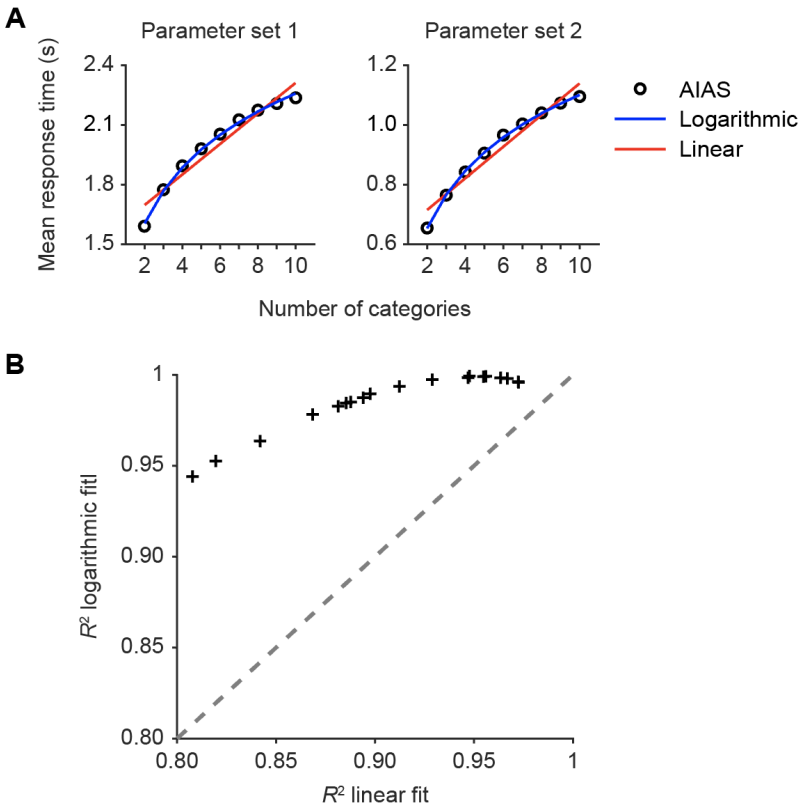

Figure 2.5: AIAS replicates Hick's law. A) The mean response time increases approximately logarithmically with the number of categories. The simulations generated from two example participants' parameter sets are shown on the left (participant 5 from H.-H. Li and Ma (2020)'s experiment) and right (participant 7 from Adler and Ma (2018)'s experiment). The black dot shows the mean response time simulated by AIAS. The blue line shows the logarithmic fit of AIAS's response time to the number of categories, and the red line shows the linear fit. B) The R^2 of logarithmic and linear fits were plotted against each other. Each plus sign represents the simulation generated from a participant's estimated parameter set (nineteen in total). The grey dashed line represents $y = x$. The area in the upper left triangle means that the logarithmic model better accounts for the simulated-data.

beliefs about the posterior, where each noisy sample represents a 'piece' of computation. We studied the implications of our model in the context of perceptual categorization and found that AIAS accounts for several empirical findings.

First, AIAS explained the somewhat puzzling recent finding that decision confidence follows the difference between the two highest posteriors, rather than the highest posterior itself (H.-H. Li & Ma, 2020). Second, we showed that AIAS could explain how categorization behavior changes with variations in visual contrast (Adler & Ma, 2018). In these studies, AIAS also *predicted* the shape of response times from the fitted parameter set based on choice and confidence responses alone. Third, we found that the mean response times predicted by AIAS increase approximately logarithmically with the number of categories N, consistent with Hick's law (Hick, 1952). Our results provide evidence for our framework as a useful process model for how how inference may be conducted in humans, particularly in the context of intractable and expensive Bayesian computations.

The observation that textbook Bayesian inference is intractable and expensive has been explored at length in the cognitive science and neuroscience literature, postulating that the brain must implement some form of approximate inference (Beck et al., 2012; Lange et al., 2021). Popular approximate inference techniques used in statistics and machine learning and co-opted to model brain computations include Markov Chain Monte Carlo (Metropolis et al., 1953) and variational inference (Jordan et al., 1999). However, almost all of these accounts do not question the computability of the likelihoods, which are assumed to be readily available. AIAS differs from common approaches to approximate inference in the brain in two key aspects. First, we assume that the likelihood itself is not available and needs to be computed (inferred, as a matter of fact). Second, unlike traditional approximate inference methods, and closer to more recent approaches in probabilistic and sample-efficient machine learning such as Bayesian optimization (Garnett, 2023; Hennig et al., 2015), we model the inference process as a sequential decision making process. That is, the brain actively decides at each step which costly computations to perform

(modelled as noisy samples), and when to stop.

Thus, at the heart of AIAS lies the problem of active sampling and termination. To better understand how AIAS functions, we compare it to other active sampling models in the field. These approaches generally fall into two categories: those based on the principle of resource rationality and those employing a heuristic utility function. In the resource-rational approach (Griffiths et al., 2015), researchers assume that the agent optimizes performance in a reinforcement learning manner – they get rewarded for being correct, and the sampling cost is deducted. By assuming this structure, the problem of active sampling and termination can be solved by using Bellman equation and dynamic programming (Sutton & Barto, 2018), but are intractable in the general cases. Jang et al. (2021) approximated it with dynamic programming method by assuming the active sampling process must stop after a preset horizon, while Callaway et al. (2021) used feature engineering to approximate the utility of sampling at each step. For heuristics like Z.-W. Li and Ma (2021) and Song et al. (2019), they both assumed the probability that an agent will sample an item is proportional to a weighted average of its current value and uncertainty level. In some aspects, AIAS aligns more closely with the heuristics approach, since it does not assume that the agent explicitly tackles the problem of reward-maximization. However, AIAS also shares similarities to the resource-rational approach, since it does not explicit encourage exploration like using uncertainty-boosted utility; instead, this happens naturally within the confidence-based active sampling process. Additionally, unlike these models, AIAS explicitly defines a new concept of 'comparative confidence' and significantly advances the understanding of confidence in decision-making. Lastly, AIAS emphasizes categorization. While the value-based decision-making models mentioned above could be adapted for categorization, this would require substantial effort.

The active sampling component within AIAS helps agents make accurate categorization decisions as quickly as possible. Similarly, in the field of perception, active sensing allows humans to interact efficiently with their environment by guiding exploration and gathering information about the physical world (Yang, Lengyel, & Wolpert, 2016; Yang, Wolpert, & Lengyel, 2016). Both

active sampling and sensing processes function similarly by reducing reaction time and improving task performance. However, they differ conceptually. Active sensing continuously gathers new perceptual evidence from the environment, such as texture in a different area, while AIAS refines the inference process within the agent itself by allocating computational resources to sample from the likelihoods. The variability in active sensing arises from the uneven distribution of perceptual information in the search field, whereas in AIAS, it stems from the imperfect representation of the likelihoods in the brain.

Within AIAS, we assumed a very specific structure for how the framework functions. It is worthwhile to list some of the assumptions that could potentially be loosened. First, we assume a log-normal structure for both the prior of the likelihoods and the likelihood samples. This assumption is made purely for computational tractability. Still, the crucial aspect here is that the brain maintains a distribution rather than a point estimate of the likelihoods. Second, we assume that the agent samples from each category once at the beginning. This assumption was adopted to save time when fitting AIAS by reducing iterative processes. It is likely that the brain initiates the AIAS process with zero samples, or only after gathering more than one sample for each category. Third, we assume that to decide which category to sample from and when to stop, the agent simulates a phantasized likelihood sample and bases its actions on the simulated outcomes. This process is known as *Thompson sampling*, a heuristic decision rule that approximates the intractable marginalization over beliefs by choosing the action that maximizes expected utility given a single sample from the current belief distribution (Thompson, 1933). Alternative sampling methods, such as the upper confidence bounds method, a commonly used acquisition algorithm in reinforcement learning and Bayesian optimization, could also be employed (Auer, 2002; Garnett, 2023).

A somewhat paradoxical aspect of AIAS is that we modelled the Bayesian inference process…with Bayesian inference, in that we assumed that the agent builds a posterior over the likelihoods (see Fig. 2.1). This second-order posterior is crucial and separates AIAS from the

majority of previous modelling attempts in that it provides uncertainty estimates used to guide the inference process itself (i.e., active sampling). This aspect of the model should not be taken literally (as any other part of AIAS): Bayesian inference here is a principled, abstract modelling choice to represent that the brain keeps track of the uncertainty of computations and of current estimates. The framework is agnostic with respect to how exactly this representation is kept in the brain, and is compatible with heuristic representations of second-order uncertainty. Importantly, second-order uncertainty (uncertainty about the posterior) provides an interesting new interpretation for confidence which we showed to explain empirical findings of human confidence reports (H.-H. Li & Ma, 2020).

Additionally, we would like to connect our study to another prominent area in Bayesian inference: approximate Bayesian computation (ABC), or more broadly, simulation-based inference (Cranmer et al., 2020; Marjoram et al., 2003; Peters et al., 2012; Sisson et al., 2007, 2018). A well-known example in this field is rejection sampling. Researchers in ABC face the challenge of deriving the posterior distribution from existing data when the likelihood function is computationally intractable. This parallels our hypothesis in AIAS, in which the true likelihood cannot be exactly known. In ABC, to estimate the posterior distribution without relying on a tractable likelihood function, algorithms employ an iterative process. This process involves simulating data from an assumed underlying process and gradually refining the posterior estimate. This iterative refinement is similar to the process in AIAS, in which noisy likelihood samples are progressively drawn to approximate the posterior with increasing accuracy. Moreover, in ABC, active learning techniques can be employed to expedite the simulation process, analogous to the active sampling mechanism in AIAS, which accelerates the refinement of the posterior belief. However, there are essential distinctions between ABC and AIAS. First, AIAS assumes a specific structure for the likelihood sample (a log-normal distribution), which, although an arbitrary way we choose to represent imprecise likelihood estimates, contrasts with the fundamental assumption of ABC that the likelihood function is computationally intractable. Second, in AIAS, all likelihood samples

contribute to improving the representation of the true likelihood and true posterior. In contrast, in ABC, simulated data that poorly match the existing data are typically discarded.

AIAS provides a framework for understanding how approximate inference occurs in the human brain. Though we found supporting evidence through model fitting and model predictions, the data we used – choice, confidence and response times – all reflect the outcomes of a hidden inference process. Is it possible to more directly examine the active sampling process at work? Eye movements might offer insights into the agent's thought processes, potentially indicating which category the agent is actively sampling from. For example, fixation times have been reported to correlate with the final choice in multiple-alternative value-based decision-making (Krajbich & Rangel, 2011), and researchers in the field of value-based decision-making used eye movement data to test their models (Callaway et al., 2021; Jang et al., 2021). A similar dataset in categorization can thus help determine whether eye movements correlates with AIAS's sampling process. However, researchers should exercise caution in experimental design, as eye movements can also serve the purpose of active sensing, as mentioned above.

An interesting future direction of research would be to investigate the neural basis of AIAS in the brain. Sequential sampling mechanisms have been well-established in neural systems for processing visual inputs in the visual cortex (Gold & Shadlen, 2001) and for memory retrieval in the hippocampus (Shadlen & Shohamy, 2016). Our theoretical framework proposes a possible third source of sequential sampling, which may have distinct neural correlates. One potential location for this mechanism is the medial frontal cortex, as has been identified for information sampling in economic decision-making tasks (Kaanders et al., 2021). Future imaging studies are needed to address this question.

2.4 Methods

In this section, we first cover in detail the mathematical framework of AIAS. We then provide task and model details for the two tasks presented in the main paper, the multi-alternative categorization and decision confidence task and the variable-reliability task. Finally, we outline the model fitting procedures, including model comparison and data visualization techniques.

2.4.1 Mathematical Description of AIAS

Generative model. We consider an N-alternative categorization task under uncertainty. We assume the agent receives a potentially noisy observation x. The true likelihoods $L_i = p(x|C = i)$ and the true posteriors $P_i = p(C = i|x), i = 1, 2, ..., N$ are unknown to the agent. We denote the true posterior vector as $\mathbf{P} \triangleq (P_1, P_2, \cdots, P_N)$. In AIAS, we assume that the agent estimates the true likelihood values (and hence the true posterior values) by iteratively drawing noisy samples of the likelihoods. At each time step, the agent can draw a noisy likelihood sample l_i^k for a chosen category i, where k is the position in the sequence of likelihood samples for that category. Specifically, it is from a log-normal distribution centered on the true likelihood value, $p(\log l_i^k|L_i) = \mathcal{N}(\log l_i^k; \log L_i, \sigma^2)$. We denote the vector of all likelihood samples of category i as $\mathbf{l}_i \triangleq (l_i^1, l_i^2, \cdots)$.

Inference. Given a set of likelihood samples $\{\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_N\}$ and a prior over likelihoods $p(L_i)$, the agent is able to compute their belief about the true likelihoods, i.e. the posteriors over the likelihoods using $p(L_i|\mathbf{l}_i) \propto p(L_i)p(\mathbf{l}_i|L_i)$. Specifically, we assume an empirical log-normal distribution for the prior over each likelihood $p(\log L_i) = \mathcal{N}(\log L_i; \mu_L, \sigma_L^2)$ and the agent's belief (likelihood in their perspective) of getting each individual likelihood sample is the same as the generative model, i.e. $p(\log l_i^k|L_i) = \mathcal{N}(\log l_i^k; \log L_i, \sigma^2)$. Combining the prior over the likelihood

and the likelihood of the likelihood, this gives us

$$p(\log L_i|\mathbf{l}_i) \propto \mathcal{N}\left(\log L_i; \frac{\frac{\mu_L}{\sigma_L^2} + \sum_k^{n_i} \frac{\log l_i^k}{\sigma^2}}{\frac{1}{\sigma_L^2} + \frac{n_i}{\sigma^2}}, \frac{1}{\frac{1}{\sigma_L^2} + \frac{n_i}{\sigma^2}}\right),$$

where n_i is the number of the drawn likelihood samples for category i. Now that the agent has their belief about the true likelihoods, they then apply Bayes' rule again to calculate their belief about the true posteriors, i.e. the posteriors over the posteriors $P_i = \frac{p(C=i)L_i}{\Sigma_{j=1}^N p(C=j)L_j}$. We assume an equal prior over all categories $p(C=i) = \frac{1}{N}$, and this gives us $P_i = \frac{L_i}{\Sigma_{j=1}^N L_j}$. We further denote the observer's belief about the true posterior vector as $p(\mathbf{P}|\mathbf{l})$, where $\mathbf{l} \triangleq \{\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_N\}$ is the set of all likelihood samples of all categories.

Given the posterior over the posterior vector, $p(\mathbf{P}|\mathbf{l})$, we further define the agent's confidence c_i as the subjective belief that the chosen category i has the highest true posterior probability when making a decision for categorization, i.e. $c_i = p(P_i = \max(P_1, P_2, \cdots, P_N)|\mathbf{l})$ which we name "comparative confidence". In the $N = 2$ case, the c_i can be analytically calculated. In the more general cases, to derive c_i, an $N-1$ dimensional integration is needed. In the model fitting, we conducted numerical integration (Riemann summation) to get this value.

Active Sampling, Termination, and Decision. Previously, we assumed that the agent can draw a noisy likelihood sample l_i^k for a chosen category i at each time step. Now we address the problems of the choice of which category to sample from (active sampling), when to stop (termination), and which category to select at the end (decision).

We assume the agent starts with no or vague information about the categories, so they first draw a likelihood sample for each category i. Then, after updating the belief given the likelihood samples obtained so far, the agent considers each category as a candidate for drawing a likelihood sample for the next iteration. The goal is to estimate the expected utility u_j for refining the likelihood of category j ($j = 1, 2, ..., N$), where the utility is some metrics related to the information gain about which category is most probable when it comes to making a decision. To perform this

estimation, separately for each category j, the agent simulates a new, 'phantasized' likelihood sample $\tilde{l}_j$ based on current beliefs,

$$
\begin{aligned}
& p(\log \tilde{l}_j | \mathbf{l_j}) \\
=& \int p(\log \tilde{l}_j | \log L_j) p(\log L_j | \mathbf{l_j}) d \log L_j \\
=& \mathcal{N}\left(\log \tilde{l}_j ; \frac{\frac{\mu_L}{\sigma_L^2} + \sum_k^{n_j} \frac{\log l_j^k}{\sigma^2}}{\frac{1}{\sigma_L^2} + \frac{n_j}{\sigma^2}}, \frac{1}{\frac{1}{\sigma_L^2} + \frac{n_j}{\sigma^2}} + \sigma^2 \right).
\end{aligned}
$$

The phantasized likelihood sample $\tilde{l}_j$ leads to a change of the corresponding posterior over the likelihood L_j, a change of the posterior over the posterior vector $\mathbf{P}$, and thus a change of the aforementioned confidence c_i for all N categories. We define the utility of such sampling as the absolute change of the largest confidence in all N categories, $u_j = |\Delta (\max(c_1, ..., c_N))|$. After getting the u_j, the agent discards the simulated likelihood sample $\tilde{l}_j$. After one simulation per category, if the maximum of all utilities u_j ($j = 1, 2, ..., N$) is larger than a predetermined termination threshold ϵ, the agent chooses the category j with the largest u_j to draw an actual likelihood sample. The agent then updates their posterior over the posterior vector and repeats the whole simulation process. Conversely, if the largest utility is below the threshold ϵ, the agent stops sampling as further sampling would likely yield little information gain, and chooses the category i with the highest confidence c_i.

2.4.2 Details of the Multi-alternative Categorization and Decision Confidence Task and Models

We describe here details of AIAS and other models used for the multi-alternative categorization and decision confidence ('color clouds') task (H.-H. Li & Ma, 2020). Model details are summarized in Table 2.1.

True likelihoods and posteriors. In the experiment, there were three color clouds in different colors. Each cloud consisted of $M = 375$ dots whose horizontal location was drawn independently from a normal distribution $\mathcal{N}(\mu_i, \sigma_s^2)$, where the mean μ_i for $i = 1, 2, 3$ depended on the cloud configuration, and the variance of the clouds σ_s^2 was the same for all clouds. For all dots, their vertical location was drawn independently from a normal distribution $\mathcal{N}(0, \sigma_s^2)$. The black (target) dot was generated first by randomly choosing one of the three clouds with equal probability and then by randomly choosing a dot inside. Since the vertical location distributions of the black dot of three clouds are the same, we denote the horizontal location of the black dot as s. Therefore, the true likelihoods are calculated as $L_i = \mathcal{N}(s; \mu_i, \sigma_s^2)$, $i = 1, 2, 3$. As per the generative model, we assume an equal prior across all categories. Then we apply Bayes' rule, and the true posteriors are calculated as $P_i = \frac{L_i}{\sum_{i=1}^{3} L_i}$. We ignored perceptual noise in our analyses, since the original authors reported that including perceptual noise in the observer models improved model fitting performance only marginally in the Max and Difference models (H.-H. Li & Ma, 2020).

AIAS. Model details are described in the section AIAS-based Model for Categorization Tasks. In order to fit the model to discrete confidence data in the 1-4 confidence scale used in the experiment, we discretize the AIAS-generated continuous confidence c into discrete c_d from 1 to 4 by using three boundary parameters b_1, b_2, b_3, such that $c_d = i$ if $b_{i-1} < c \leq b_i (b_0 = -\infty, b_4 = \infty)$. For the purpose of capturing outliers or participants' outright mistakes (e.g., a misclick), we also introduced a lapse rate $0 \leqslant \lambda \leqslant 1$, i.e. a probability λ that in each trial the category response and the confidence response are both randomly chosen from a discrete uniform distribution (Prins, 2012). Altogether, there are six free parameters in total: the noise level of likelihoods samples σ, the termination threshold ϵ, three boundaries to discretize the continuous confidence, b_1, b_2, b_3, and the lapse rate λ.

Max and Difference models. We describe here the Max and Difference models of confidence as introduced in H.-H. Li and Ma (2020). Both models assume, close in spirit to some of our assumptions, that agents can only access noisy estimates of the posteriors over the three

Model	Description	Parameters (Meaning)	#
AIAS	Approximate Inference with Active Sampling	σ (Noise level of likelihood samples), ϵ (Termination threshold), b_1, b_2, b_3 (Boundaries to discretize confidence), λ (Lapse rate)	6
Max	Confidence based on highest posterior	α (Dirichlet noise level), b_1, b_2, b_3 (Boundaries to discretize confidence), λ (Lapse rate)	5
Difference	Confidence based on difference of two highest posteriors	Same as Max	5
AIRS	Approximate Inference with Random Sampling	Same as AIAS	6
AISS	Approximate Inference with Same Sampling	Same as AIAS	6
AISS*	Approximate Inference with Same Sampling (Single iteration)	Same as AIAS, *without* ϵ (Termination threshold)	5
Race	Evidence accumulation with mutual inhibition	m (Drifting scalar), n (Mutual inhibition scalar), σ_d (Drifting noise level), b_1, b_2, b_3 (Boundaries to discretize confidence), λ (Lapse rate)	7

Table 2.1: **Multi-alternative categorization and decision confidence task models.** Names, descriptions, parameters and total free parameter count (#) for the models used in the 'color cloud' task.

categories (dot clouds). Specifically, posteriors are corrupted by Dirichlet-noise:

$$p\left(\mathbf{Q}\,|\mathbf{P};\alpha\right) = \frac{1}{B(\alpha\mathbf{P})}\prod_{i=1}^{3} q_i^{\alpha P_i - 1},$$

$$B(\alpha\mathbf{P}) = \frac{\prod_{i=1}^{3}\Gamma(\alpha P_i)}{\Gamma\left(\alpha\sum_{i=1}^{3} P_i\right)},$$

where the concentration parameter α denotes the Dirichlet noise level, Γ represents the gamma function, $\mathbf{P}$ is a vector consisting of the three true posteriors (P_1, P_2, P_3), and $\mathbf{Q}$ is a vector consisting of three noise-perturbed posterior probabilities (Q_1, Q_2, Q_3). Having observed the corrupted posterior $\mathbf{Q}$, the agent picks the category i whose posterior probability Q_i is the largest. In the **Max model**, the agent uses that Q_i to yield a confidence value. In the **Difference model**, the confidence value is defined as ΔQ, the difference between the two highest Q_i. In both models, the confidence value is then discretized by using boundary parameters (see AIAS above). A lapse rate λ is also included. Altogether, both models include five free parameters in total: the noise level of the Dirichlet distribution α, three boundaries to discretize the continuous confidence b_1, b_2, b_3, and the lapse rate λ. More details can be found in the original paper (H.-H. Li & Ma, 2020).

AIAS variants. We describe here three variants of the base AIAS models which differ from AIAS in some key aspects.

- **AIRS.** *Approximate inference with random sampling* (AIRS) differs from AIAS in two aspects. First, a random category is selected to draw a noisy likelihood sample from at each iteration. Second, when deciding when to stop, the agent will compare the absolute change of the largest confidence in all three categories before and after another sample $|\Delta(\max(c_1, c_2, c_3)|$ to the threshold ϵ; if the absolute change is smaller, the agent stops sampling.

- **AISS.** In *approximate inference with same sampling* (AISS), first the agent draws a likelihood sample from all three categories at each iteration. Second, when deciding when to stop, the agent will simulate the consequences of drawing a likelihood sample from all three categories and compare the absolute change of the largest confidence in all three categories before and after the simulation $|\Delta(\max(c_1, c_2, c_3)|$ to the threshold ϵ; if the absolute change is smaller, the agent stops sampling.

- **AISS*.** In *approximate inference with same sampling, only once* (AISS*) there is no iterative

process. The agent draws a likelihood sample from all three categories and immediately stops and decides which category to select.

All model variants share the same parameter set of AIAS, with the exception of AISS* where the termination threshold ϵ is not present.

Race model. The true posteriors are used as the drift rate, and mutual inhibition between different categories is allowed. The accumulation at time t for category i is updated as

$$x_i^t = x_i^{t-1} + mP_i - n\sum_{j\neq i} x_j^{t-1} + \eta, \qquad \eta \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_d^2),$$

where m is the drifting scalar, P_i is the true posterior for category i, n is the mutual inhibition scalar, and η is a zero-centered normal noise with a variance of σ_d^2. The accumulation is set to 0 when it gets below 0 and the bound was set as a constant. For the Race model, we defined the confidence value as the reciprocal of the number of iterations. The confidence is discretized by using the boundary parameters. As per the other models, we also included a lapse rate. Altogether, there are seven free parameters in total: the drifting scalar m, the mutual inhibition scalar n, the drifting random noise level σ_d, three boundaries to discretize the continuous confidence b_1, b_2, b_3, and the lapse rate λ.

2.4.3 Details of the Variable-reliability Task and Models

We describe here the tasks from Adler and Ma (2018) and the models used. Model details are summarized in Table 2.2.

True likelihoods and posteriors. For each task (1 and 2), the experiment used two categories of orientations, $\mathcal{N}(\mu_1, \sigma_1^2)$ and $\mathcal{N}(\mu_2, \sigma_2^2)$. On each trial, an orientation s was randomly drawn from one of the two categories. Like Adler and Ma (2018), we assume that people measured a noisy version of the true orientation $x \sim \mathcal{N}(s, \sigma_p^2)$. For the variance of the sensory noise, we also assume $\sigma_p^2 = \gamma + \alpha c^{-\beta}$, where c is the visual contrast of the Gabor and α, β, and γ are free parameters

(Adler & Ma, 2018). Therefore, the true likelihoods are calculated as $L_i = \mathcal{N}(s; \mu_i, \sigma_i^2 + \sigma_p^2)$, $i = 1, 2$. Equal prior is assumed across the two categories. Then Bayes' rule is applied, and the true posteriors are calculated as $P_i = \frac{L_i}{L_1 + L_2}$.

Model	Description	Parameters (Meaning)	#
AIAS	Approximate Inference with Active Sampling	σ (Noise level of likelihood samples), ϵ (Termination threshold), γ, α, β (Visual contrast parameters), b_1, b_2, b_3 (Boundaries to discretize confidence), λ (Lapse rate)	9
Extended Bayesian observer	Bayesian observer model with decision noise	σ_f (Decision noise magnitude), γ, α, β (Visual contrast parameters), b_1, b_2, b_3 (Boundaries to discretize confidence), λ (Lapse rate)	8

Table 2.2: **Variable-reliability task models.** Names, descriptions, parameters and total free parameter count (#) for the models used in the variable-reliability task.

AIAS. AIAS here is the same model as the previous section, except there are three additional visual contrast parameters. For this dataset, the model has nine free parameters in total: the noise level of likelihoods samples σ, the termination threshold ϵ, three visual contrast parameters γ, α, and β, three boundaries to discretize the continuous confidence b_1, b_2, b_3, and the lapse rate λ.

Extended Bayesian observer model. Adler and Ma (2018) consider in their analysis a variety of Bayesian and hybrid observer models with Bayesian components. Here we consider from their paper an extended Bayesian observer model, chosen to be of comparable complexity to AIAS (Adler & Ma, 2018). Besides standard Bayesian observer model assumptions, this model includes *decision noise* in that the decision variable d, i.e. the true log posterior odds, is corrupted by normal noise,

$$d = \log \frac{P_1}{P_2} + \eta, \qquad \eta \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_f^2),$$

where σ_f is the decision noise magnitude. Note that for $\sigma_f \to 0$, this model recovers a standard Bayesian observer. If $d > 0$, the agent chooses category 1; otherwise, the agent chooses category

2. The confidence is predicted by the absolute value of d and discretized using the boundary parameters b_1, b_2, b_3 as in AIAS: $c_d = i$ if $b_{i-1} < |d| \leq b_i$ (with $b_0 = 0$ and $b_4 = \infty$). Altogether, the model has eight free parameters: the inference noise σ_f, three visual contrast parameters γ, α, and β, three boundaries to discretize the continuous confidence b_1, b_2, b_3, and the lapse rate λ.

2.4.4 Model Fitting Details and Model Comparison

Model Fitting. We fit each model to individual-participant data by maximizing the log likelihood of $\boldsymbol{\theta}$, $\log L(\boldsymbol{\theta}) = \log p(\text{data}|\boldsymbol{\theta})$. All the considered models assume that the trials are conditionally independent. We denote the experimental condition, category response, and confidence report on the i-th trial by v_i, $\hat{C}_i$, and c_{di}. Then, the log likelihood becomes

$$\log L(\boldsymbol{\theta}) = \log \prod_i p\left(\hat{C}_i, c_{di} | v_i, \boldsymbol{\theta}\right) = \sum_i \log p\left(\hat{C}_i, c_{di} | v_i, \boldsymbol{\theta}\right)$$

where $\log p\left(\hat{C}_i, c_{di} | v_i, \boldsymbol{\theta}\right)$ is the log likelihood of each individual trial, described below. We optimized the parameters for each individual using Bayesian Adaptive Direct Search (BADS; Acerbi and Ma (2017)).

Per-trial log likelihoods. To fit the models we need to compute the per-trial log-likelihood $\log p\left(\hat{C}_i, c_{di} | v_i, \boldsymbol{\theta}\right)$. For most models in this paper, and particularly AIAS and its variants, this quantity is unavailable in closed form, but we are able to generate model responses (choice and confidence) by simulating from the model. Therefore, we estimated the log likelihood for a given trial and parameter vector $\boldsymbol{\theta}$ using a stochastic estimation method called Inverse Binomial Sampling (IBS), which only needs the ability to simulate from the model (van Opheusden et al., 2020). Crucially, IBS yields *unbiased* estimates of the log-probability with known variance. We used IBS to obtain an unbiased estimate of the joint log probability of category and confidence, $\log p(\hat{C}, c_d|\boldsymbol{\theta})$ for each trial.

Model Comparison. We used AIC (Akaike, 1974) for model comparison, which is calculated

as AIC $= -\log L^\star + 2k$, where $L^\star$ is the maximum value of a model's likelihood function and k is the number of fitted parameters. To report AIC, we computed the AIC for each participant and summed them up. The confidence interval of the group-summed AIS was estimated by non-parametric bootstrapping.

2.4.5 Datasets and Data Visualisation

We fitted two datasets in our analysis (Adler & Ma, 2018; H.-H. Li & Ma, 2020) (available at: https://github.com/WeiJiMaLab/confidence-multiple-alternatives and https://github.com/WeiJiMaLab/confidence). Since fitting AIAS is computationally expensive, we only fitted one experiment from each dataset. We briefly describe each dataset below.

We analysed data from experiment 3 in (H.-H. Li & Ma, 2020) in the section Multi-alternative Categorization and Decision Confidence, which consisted of 11 participants. As black dot locations were drawn from a continuous distribution, to plot the data we grouped nearby locations using a 0.6-unit moving-window. Data and model fitting results are shown in Fig. 2.3F-H.

We analysed data from experiment 2 in (Adler & Ma, 2018) in the section Inference with Varying Perceptual Reliability, which consisted of 8 participants. As Gabor's orientation were drawn from a continuous distribution, to plot the data we grouped nearby locations using a 3-degree moving-window. Data and model fitting results are shown in Fig. 2.4C.

2.5 Supplementary Figures

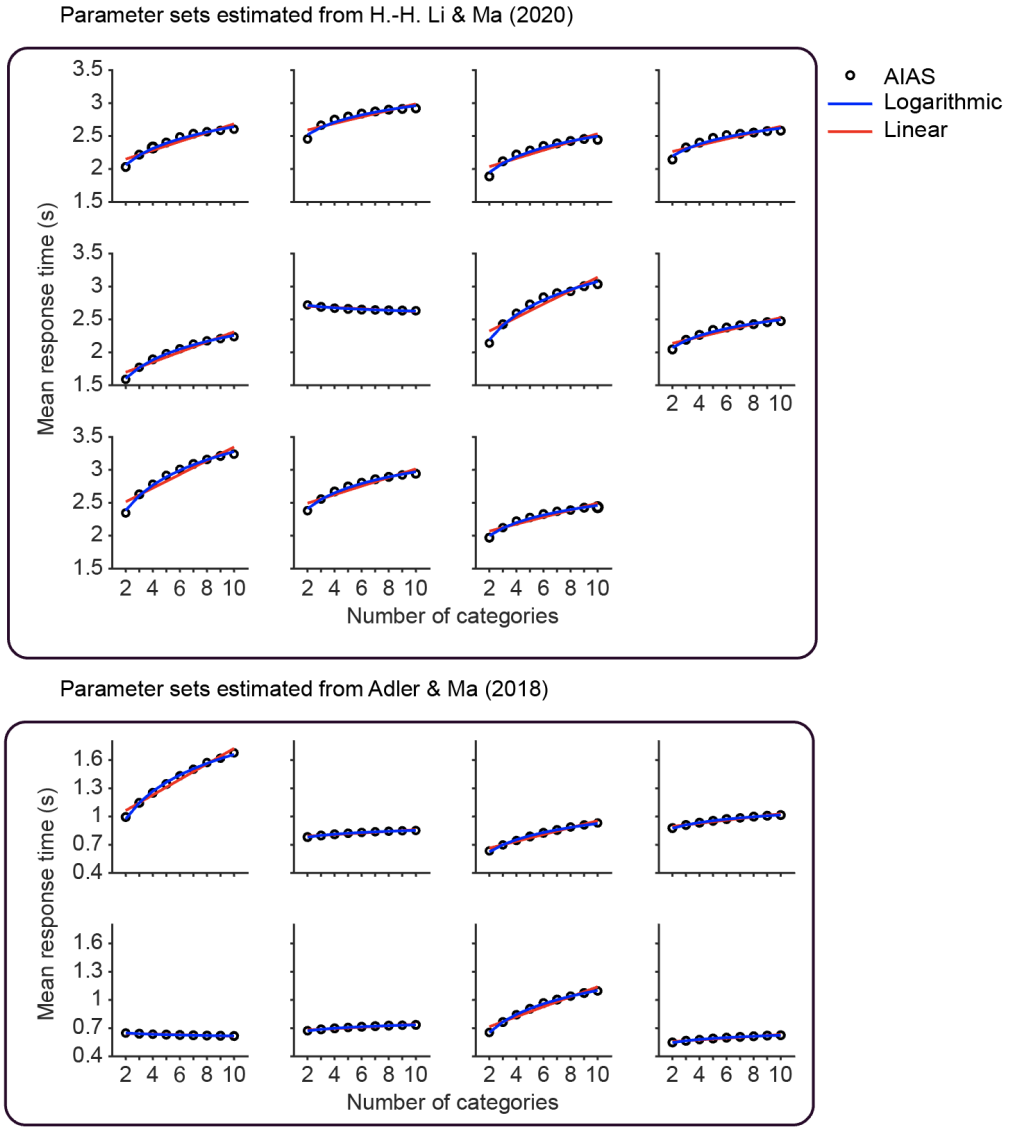

Figure 2.S1: AIAS replicates Hick's law. A) The mean response time increases approximately logarithmically with the number of categories. The simulations generated from participants' estimated parameters are shown on the top (eleven participants from H.-H. Li and Ma (2020)'s experiment) and bottom (eight participants Adler and Ma (2018)'s experiment). The black dot shows the mean response time simulated by AIAS. The blue line shows the logarithmic fit of AIAS's response time to the number of categories, and the red line shows the linear fit. While it can be hard to discern for some panels, the logarithmic fit generally provides a better match than the linear fit (see Fig. 2.5B in the main text).

3 | Cognitive Mechanisms of Collaboration: an Introduction

Collaboration is a cornerstone of human society, serving as the foundation for the advancement and well-being of communities at all levels. By working together, individuals and groups can pool their diverse skills, knowledge, and resources to tackle complex challenges that no single person or organization could address alone. In this dissertation, we define collaboration as behaviors in which two or more agents coordinate their actions, resulting in outcomes that benefit every individual involved, similar to previous studies in the field (e.g. Duguid & Melis, 2020). Daily examples include team sports like soccer, where eleven players on the field take different roles to get the ball into the opponent's goal, and working in industries, such as front-end and back-end software engineers collaborating to push an update to their app. Lifelong decisions like marriage also exemplify collaboration, as two people manage a household together.

Collaboration fosters innovation and better problem-solving by bringing together diverse perspectives and ideas, leading to creative and effective decision-making in fields such as healthcare (Morley & Cashell, 2017), education (Johnson & Johnson, 2002), technology (Hargadon & Sutton, 1997), and more. It also promotes unity and trust within communities, as working towards common goals requires mutual understanding and cooperation among different groups (Plötner et al., 2015; Tomasello & Vaish, 2013). Furthermore, collaboration drives economic growth and the advancement of human civilization by enabling more efficient resource allocation and facili-

tating the sharing of information (Balland et al., 2020; Hidalgo et al., 2007). In a globalized world facing intercontinental issues like climate change and pandemics, the ability to collaborate across borders is becoming even more crucial (Andersen et al., 2012). Ultimately, understanding collaboration and leveraging its best practices are vital for enhancing the quality of life for all members of human society.

Understanding collaboration in humans begins with studying its developmental process in early childhood, which reveals the building blocks of collaborative behavior. Infants as young as 12 months can point to provide information to others, demonstrating the ability to share information (Liszkowski et al., 2006). By 14 months, children start to grasp the concept of "we," indicating the very early development of shared intentionality (Moll et al., 2008). In their second year, children exhibit significant advancements in their collaborative abilities. They comprehend language conveying cooperative logic of requests (Grosse et al., 2010), recognize their partner as an intentional cooperative agent (Warneken et al., 2012), and successfully participate in cooperative problem-solving (Warneken et al., 2006). During the third year, children's collaboration skills become more sophisticated. They develop a better recognition of commitments of their own and their partners' in collaborative tasks (Gräfenhain et al., 2009) and show a more developed understanding of shared intentionality (Gräfenhain et al., 2013). They also show a preference for equal sharing after collaboration (Hamann et al., 2011; Warneken et al., 2011) and plan the division of labor, showcasing early organizational skills (Warneken et al., 2014). By the age of five, children can explicitly coordinate different roles in a shopping task, demonstrating an advanced level of role differentiation and cooperative planning (Baer & Odic, 2022; Gauvain & Rogoff, 1989). These developmental milestones highlight the natural progression of collaborative abilities and provide valuable insights into the foundational elements of collaboration (Tomasello & Hamann, 2012).

Another way to understand collaboration in humans is through comparative psychology, which involves studying collaboration in other species and comparing it to that in humans. Collaborative behaviors are prevalent in the animal kingdom and are observed both in the wild and

under experimental conditions. The most typical collaborative behavior in the wild is group foraging or group hunting, found in wild dogs (Estes & Goddard, 1967), baboons (Maples, 1969), chimpanzees (Boesch & Boesch, 1989), and many more. These studies qualitatively demonstrate that animals can collaborate, but without experimental controls, it is challenging to investigate the characteristics of their collaboration in depth.

Most laboratory studies of animal collaboration focus on chimpanzees, who exhibit high-level collaborative skills. They recruit collaborators only when necessary and can recognize and recruit the best collaborators, showcasing their ability to deliberate on the benefits of collaboration and mentalize their potential collaborators (Melis et al., 2006). They excel in complex cognitive collaborative games like the stag-hunt game (Bullinger, Wyman, et al., 2011), which typically requires advanced reasoning ability (Yoshida et al., 2008). Chimpanzees can also pick and transfer the right tool to their partner when different roles are required in a task, demonstrating more advanced skills than simple division of labor (Melis & Tomasello, 2013). Furthermore, they can communicate to coordinate with each other during collaboration (Melis & Tomasello, 2019). Overall, chimpanzees in laboratory settings demonstrate advanced collaborative abilities from multiple perspectives.

However, research still shows a significant gap between their collaborative behaviors and those of humans. For example, when a collaborator stops performing, 2-year-old human infants, unlike chimpanzees, attempt to communicate to reengage their collaborator (Warneken et al., 2006). When given choices to work alone or collaborate, chimpanzees prefer individual strategies (Bullinger, Melis, & Tomasello, 2011; Rekers et al., 2011). They also show a preference for less cognitively complex collaborative strategies (Bullinger, Wyman, et al., 2011). Unlike humans, collaboration does not lead chimpanzees to equally share rewards, and they do not adapt to their partner or converge on solutions as quickly as humans do (Duguid et al., 2020). These distinctions highlight the unique and advanced nature of human collaboration, setting it apart from even our close animal relatives.

We now turn to research on collaboration in human adults, as they are capable of more complex collaborative tasks with a greater potential to uncover the mechanisms of collaboration. A traditional line of research in this area is the cooperative games like public goods games (Camerer, 2011; Ostrom et al., 1994). A famous example is the prisoner's dilemma, where two agents decide whether to betray each other or keep a secret. Similar matrix payoff games include the Battle of the Sexes (Camerer, 2011), where participants decide whether to converge on movie A favored by player A or movie B favored by player B. If they cannot reach an agreement, neither receives a reward. The stag-hunt game involves both players deciding whether to hunt together for a stag, which requires more effort but offers a greater reward, or to catch a rabbit with less effort and reward, but independent of the partner's decision (Camerer, 2011). The weakest-link game determines the reward based on the minimum number chosen by all participants, with no extra return for efforts exceeding this minimum (Van Huyck et al., 1990).

These studies pioneered the quantitative investigation of humans' coordination processes in the context of possible competition, stemming from the field of game theory. While coordinating actions are a core component of collaboration and these games mimic aspects of real-life collaboration (i.e., each individual might have their own interests), they also pose obstacles for researchers aiming to investigate the mechanisms of collaboration. The intention to compete might overshadow the collaborative process. Additionally, these games typically involve turn-based actions, whereas real-life collaboration often occurs in real time. Consequently, although some research has examined how decisions evolve in the repeated version of these games (Browning & Colman, 2004; Helbing et al., 2005), these experiments do not fully capture the temporal dynamics of real-world collaboration.

Partially addressing these limitations, some studies employed the paradigm of real-world group discussions. For example, jury deliberations to determine the guilt of several crime suspects (Stewart & Stasser, 1995) or student council meetings to decide on the election of a chairperson (Stasser & Stewart, 1992). These scenarios arguably offer the most vivid representations of real-

world collaboration: each participant contributes to the discussion of a common goal, ultimately reaching a consensus. However, it is challenging to quantify communication patterns in free discussion settings. As a result, these studies can only qualitatively contribute to our understanding of collaboration, for example, how information sharing can be promoted, similar to developmental psychology studies in infants with their own limitations.

More recently, researchers have adopted more sophisticated research paradigms to investigate collaboration. For example, in collaborative search tasks, two players search together on a screen to determine if a stimulus is present (Andrade-Lotero & Goldstone, 2021; Brennan et al., 2008). This setup examines whether participants can divide the labor by coordinating specific areas to each player, thus addressing the problem more efficiently. However, in most of these cases, collaboration is beneficial but not strictly necessary, and participants often show varying degrees of labor division. Therefore, it would be ideal to create a game where collaboration is essential, such that the team cannot succeed without it.

In the field of artificial intelligence (AI), recent years have seen a surge of interest in collaboration, whether between humans and computers or AI, or between two computer agents. A notable benchmark in this area is the OpenAI paper on Dota 2 (Berner et al., 2019), where a computer-trained AI team of five outperformed the then best human teams in the world, OG. Similarly, in other collaborative games like Hanabi, researchers have developed AI agents that achieve above-human-level performance and are comfortable for human players to interact with (Bard et al., 2020). More examples can be found under terms such as AI alignment (Ji et al., 2023), multi-agent learning (Du et al., 2023), and human-robot interaction (Sheridan, 2016).

The emergence of collaboration studies in AI provides valuable insights into understanding collaboration in humans. Conversely, studying human collaboration also offers knowledge for AI researchers to develop better agents that can collaborate with each other or with humans (Carroll et al., 2019). For example, humans are adept at collaborating with unfamiliar partners, a challenge that AI has faced (Duguid et al., 2020; Mehta et al., 1994). By borrowing insights from

human studies, researchers have created AI agents capable of zero-shot learning for collaboration with novel agents (Hu et al., 2020; Strouse et al., 2021). These advancements in AI collaboration have also been crucial to the development of recent popular generative AI technologies (Ji et al., 2023).

Coordinating roles is an important part of collaboration, particularly when distinct functional roles are needed, the result of which is usually termed as role specialization or role differentiation. The benefits of role specialization are evident across various domains, from animals to AI-controlled agents. For example, a honey bee colony comprises several dozen female workers, several male productive drones, and only one egg-laying queen (Oster & Wilson, 1978; Page & Mitchell, 1990). Mammals like lions and bottlenose dolphins assign different roles during group hunting, such as 'flankers' circling the prey and 'drivers' herding the prey (Gazda et al., 2005; Stander, 1992). For AI agents, role coordination is crucial for effective interaction with human teams (M. Li et al., 2021). Multi-agent coordination algorithms incorporate either implicit (Wang et al., 2020) or explicit (Raileanu et al., 2018) role planning to improve their performance. In humans, the ability to coordinate roles in collaboration emerges as early as three years old, as has been shown in a study that children can choose complementary tools with respect to their partner (Warneken et al., 2014). However, more advanced role coordination, such as assigning more challenging tasks to the more skilled individuals, typically does not appear until later ages (Baer & Odic, 2022; Magid et al., 2018), indicating a gradual development of role coordination skills.

In the rest of this dissertation, we will explore human collaboration in a novel real-time continuous-space task. This task, called Zombie Escape, is a two-player game designed to require participants to coordinate different roles. A key focus of our exploration is understanding how humans coordinate roles and uncovering the temporal dynamics of this process. In Chapter 4, we will introduce the task and describe the dataset collected from an online experiment platform. We will then present statistical and linear regression results to the game. In Chapter 5, we develop a Bayesian linear extrapolation process model and fit the model to account for the

temporal process of role coordination in the game.

4 | Human Role Coordination in a Real-time Collaborative Game

Abstract

Collaboration enables humans to achieve feats together that no individual could achieve alone. To maximize utility through collaboration, agents must not only plan and react to the environment but also coordinate their roles dynamically in real-time. However, most previous studies have relied on grid-world, turn-based paradigms to study collaboration, which are less naturalistic than real-world collaboration and insufficient to probe role coordination in real time. In this study, we introduce Zombie Escape, a new two-player game paradigm where both time and space are continuous, designed to study human collaboration and the temporal process of role coordination. We found that players quickly adapted to the environment and efficiently coordinated their roles in response to changing conditions. The difficulty level was balanced—challenging enough to engage participants without discouraging effort or making the game monotonous. We observed a default pattern of role coordination among participants, and when the default role differentiation conflicted with the required role differentiation, dyads were less likely to adopt the correct roles. Moreover, we discovered that while deciding whether to maintain or switch their coordinated roles, players tracked past performance and adjusted future role coordination accordingly, showing a pattern of win-stay-lose-shift. These results demonstrate that individuals can successfully

collaborate in a real-time tasks and dynamically coordinate their roles in changing environments.

Keywords: collaboration; role coordination; role differentiation; real-time; continuous-space

4.1 Introduction

In the 2013-2014 Premier League match between Arsenal and Norwich, four attackers from Arsenal outmaneuvered Norwich's eight-man defense by passing six times in five seconds, including two back-heels, resulting in a goal by Jack Wilshere. What makes this difficult? Of course, the motor-perceptual coordination refined by years of practice is necessary. The presence of opponents also narrows the space the attackers can exploit. Additionally, the attackers need to plan in real time. Within a brief window like five seconds, there is no way to follow an alternating plan-move routine. Unsurprisingly, the core of this process is collaboration. Furthermore, role coordination occurs in real time as they need to decide who will assist and who will take the shot.

Real-time collaboration, including the ability to coordinate roles such as locating items and managing a shopping list, can develop as early as five years old, as demonstrated in a dyadic grocery shopping task (Gauvain & Rogoff, 1989). This skill is also evident in more perceptual tasks. In a dyadic visual search task, researchers found that through either gaze position visualization or voice communication, participants were able to assign different areas to search, thereby making the visual search more efficient (Brennan et al., 2008). However, in these tasks, collaboration is beneficial rather than necessary, thus only the efficiency, rather than the mechanism, of collaboration is examined.

Now let's look at the previous paradigms researchers have used to probe collaboration. Yoshida et al. (2008) used a stag-hunter or rabbit-catcher grid-world turn-based game, where two players either collaborate to catch a high-reward stag together (both get a high reward) or a relatively low-reward rabbit (the catcher gets a higher reward than the other). In this game, collaboration is necessary to hunt the stag. Researchers have also used grid-world turn-based navigation games to probe collaboration, where two players control their agents in a narrow grid world with collision, making collaboration necessary for both agents to navigate (Kleiman-Weiner et al., 2016). Adapted from the famous collaborative video game Overcooked, researchers have used a grid-

world turn-based cooking game to investigate collaboration when problems are comprised of sub-goals (Wu et al., 2021). All these games are constrained to grid-world and turn-based settings. However, in the real world, such as the example of soccer mentioned before, naturalistic collaboration involves continuous space and continuous time. Additionally, some of these games have competitive components, which might hinder the investigation of collaboration itself.

We therefore designed the experiment task such that: 1) it occurs in continuous space and in real time to better mimic collaboration happening in the real world, as opposed to grid-world and turn-based games, 2) it is non-competitive so that nothing hinders our investigation of collaboration, 3) both participants share the same reward at all times to maximize their potential for collaboration, and 4) collaboration is necessary rather than merely beneficial.

In the task, participants played as a “red survivor" and a “blue survivor," tasked with reaching the cure to save the world. Either survivor reaching the cure is a success for both of them. However, to get to the cure, they need to outmaneuver a faster zombie. If they are caught by the zombie, they are dead for the current trial and cannot move. If both of them are dead, the trial fails. One caveat is that to succeed, one of the players has to sacrifice themselves by being caught by the zombie first to let their partner reach the cure. Under this design, participants' role coordination can be characterized by whether they are performing as the baiter (the one attracting the zombie) or the dasher (the one reaching the cure). To proceed to the next environment, participants need to win three times before failing ten times in total. We evaluated the dynamics of collaboration, with an emphasis on role coordination, across multiple environments to systematically investigate collaboration in different settings. We predicted that participants would be able to adapt to different environments to successfully collaborate with each other. We also predicted that within one environment, they would learn from their previous history of role coordination to dynamically adjust to the best coordinated roles to conquer the current environment. We tested these predictions through statistical analysis.

Overall, we found that participants succeeded in collaborating with each other across different

environments. They quickly learnt to collaborate with each other and cruised through the task with some efforts. We found that they tended to have a default pattern of role coordination when completing the task, that is, to have the closer player to the zombie as the baiter. When the required role differentiation conflicted with this, they were less likely to coordinate their roles correctly. They also dynamically adapted their role coordination within a single environment as the trial progressed, showing a interesting pattern of win-stay-lose-sift.

4.2 Methods

4.2.1 Participant

Participants were recruited on Prolific (https://www.prolific.co) to take part in a psychology experiment. Participants were informed that the task would require them to play a collaborative video game with another online player. After passing several comprehension checks of the instructions, participants entered a waiting room to be paired with another anonymous player online. The task began immediately after two participants were paired; if a participant did not find a pair within five minutes, they exited the experiment and received a partial compensation for their time (about 5 minutes for the comprehension task and 5 minutes waiting) amounting to $3. The full experiment took about 50 minutes, and participants were paid $10 and a bonus of $0.08 for every trial they succeeded. The maximum bonus was $5. On average, participants were awarded a bonus of $4.3. If either participant in the dyad dropped out in the middle of the task, there data was discarded, and they received a partial payment of $3 plus $0.35 for every condition they completed. We recruited 111 participants. Eleven participants were unable to find another participant to pair with. Eighteen dyads (36 participants) were unable to finish the whole experiment due to internet connection issue or dropping out in the middle. After these exclusions, our dataset consisted of N = 32 dyads (64 participants; 20 female, 42 male, 2 other). The average age

of participants was 38.5 years old (SD = 10.4).

4.2.2 Paradigm

We name our paradigm “Zombie Escape”. The game is set in an apocalyptic scenario with zombies. A demo interface seen by a player in the experiment is shown in Fig. 4.1. The two agents, denoted by the red and blue dots, are the so-called “survivors”. The agent with the "self" text is controlled by the player, while the other one with the "partner" text is controlled by their partner. Their partner sees the same interface, but with the "self" and "partner" text swapped. The green triangle represents the computer-controlled zombie. The green line on the right represents the cure.

The mission for both players is to save the world by reaching the cure. If either player reaches the cure, both players succeed and the trial ends immediately. They receive a bonus of $0.08 for every trial they succeed. It does not matter who reaches the cure; the bonus depends only on the number of trials the dyad as a whole succeeds. That is to say, both players are always paid the same amount and there is no extra pay for individually reaching the cure. To reach the cure, they can press "F" to move left or "J" to move right. They cannot move in the vertical direction. Both players cannot communicate with each other during the experiment, except by controlling their own agents.

However, there is a zombie blocking the players from reaching the cure. The zombie chases players to stop them from reaching the cure. The zombie can move in any direction and always chases the player closer to itself, and this is explicitly instructed and tested in the questionnaire. If a player is caught by the zombie, they are "dead", disappear from the screen, and cannot further control their agent until the next trial. The dyad fails if both players are caught by the zombie before reaching the cure. The zombie is faster than the players, and to encourage collaboration between players, the game is designed so that if both players head toward the cure starting from the beginning of the trial, they will fail.

The game server is refreshed at a frequency of 15 Hz. The positions and movements of players and the zombie are recorded at this pace. At the beginning of each trial. there is three-second countdown. On average, each trial lasts for about 12 seconds.

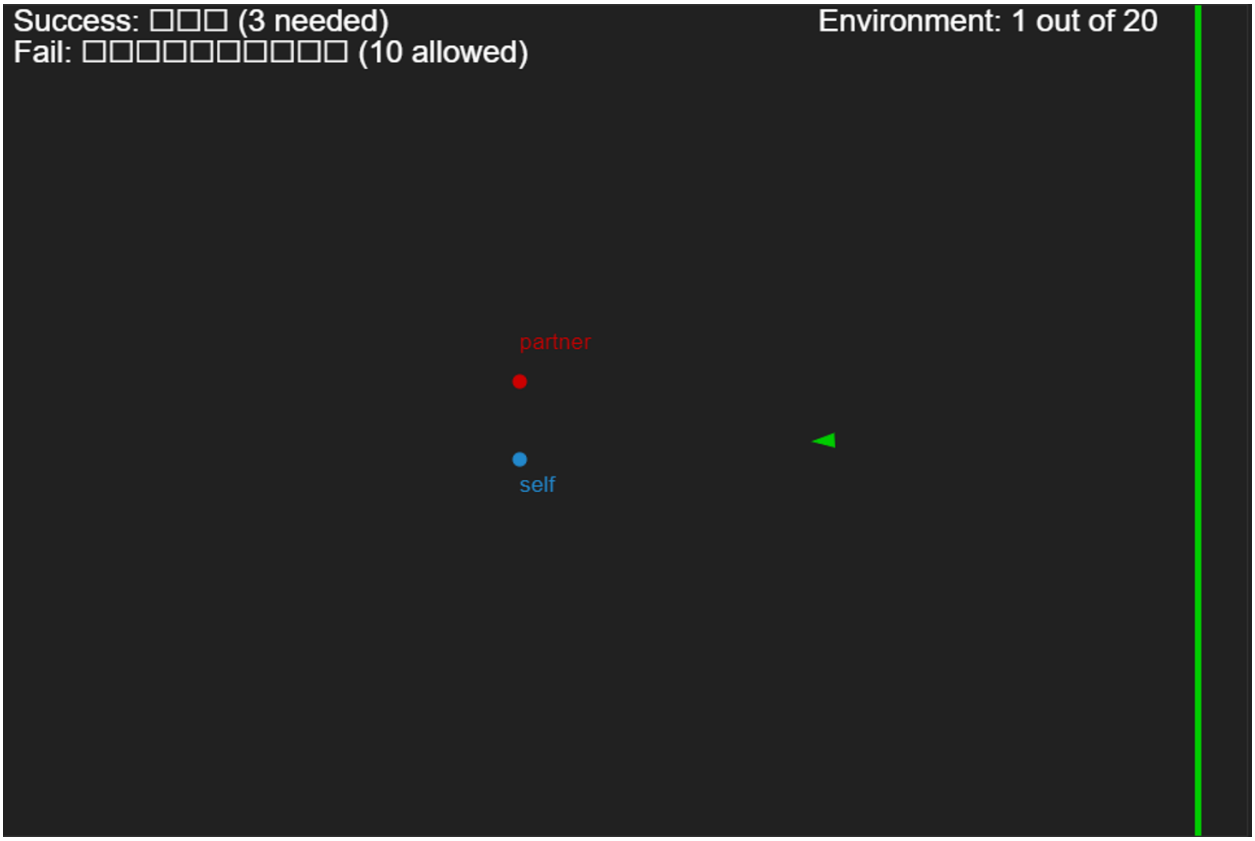

Figure 4.1: The interface of the Zombie Escape paradigm. Two players are represented as red and blue dots, while the computer-controlled zombie is shown as a larger green triangle. The green line on the right represents the cure. The goal for both players move left and right to reach the cure, which is blocked by the zombie. The zombie chases the closer player and catches a player when the green triangle touches them, causing the caught player to disappear until the next trial. If either player reaches the cure, both players succeed. If neither player reaches the cure, both players fail. The top left displays the progress in the current condition (environment), and the top right shows the overall progress of the game.

4.2.3 Experimental Design

There are 20 different conditions (environments) in the game, consisting of 2 practice conditions and 18 formal conditions. The two practice conditions always start at the beginning, and the order of the remaining 18 formal conditions is shuffled across different dyads. We will provide a detailed breakdown of these 20 conditions later. An alert will appear at the beginning of a new condition, demonstrating the speeds of both players. After every five conditions, there will be a one-minute break. The dyad can choose to extend the break to a maximum of four minutes before being considered idle and dropped from the experiment."

In each condition, the dyad needs to succeed three times to proceed to the next condition, with the assumption that they have fully mastered it. If the dyad fails ten times in total, they will also proceed to the next condition, assuming the current condition is too difficult for them. The number of trials for a condition can vary from three trials (three successes in a row) to twelve trials (three successes plus nine failures or two successes plus ten failures). This information, along with the overall progress of the game (i.e. how many conditions they have finished) is displayed at the top of the screen (Fig. 4.1).

Now we provide a detailed description of the 20 conditions. The game is displayed on a 960×640 pixel canvas (Fig. 4.2), with the top left corner of the screen at (0, 0). The average horizontal starting position of both players is always at $x = 400$, and the average vertical starting position of both players is always at the middle of the screen, $y = 320$. One player is always at the bottom throughout the game, while the other is always at the top (randomly decided when the game is created). The average speed of both players is always $v = 46.67$ pixels/second. The horizontal offset of both players Δx, their vertical deviation Δy, vertical distance from the middle of the screen $y = 320$, and their speed difference from the average speed Δv vary across different conditions. The zombie always starts in the vertical middle of the screen, $y = 320$, and has a speed of $1.65v = 77$ pixels/second. Its horizontal position x_z varies across different conditions. The cure

is always positioned at $x = 920$.

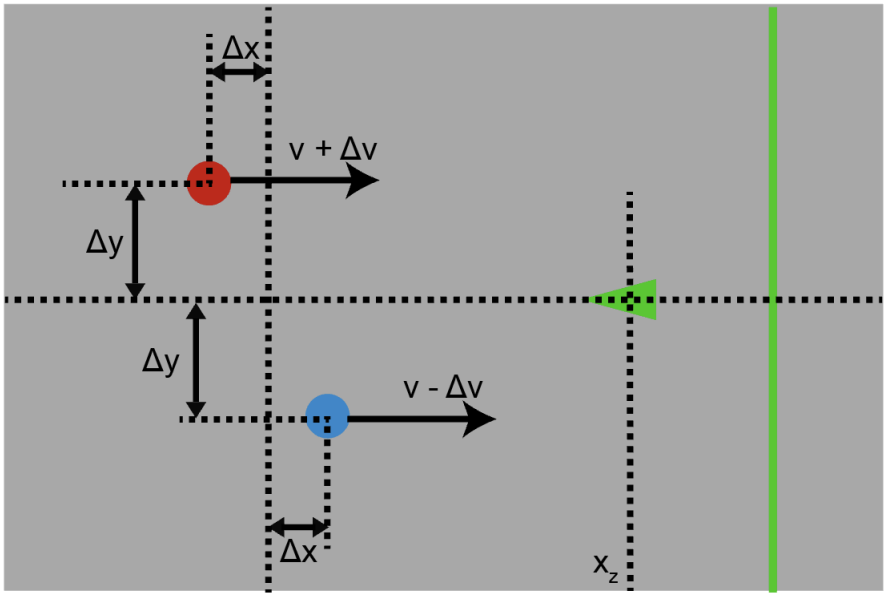

Figure 4.2: The experimental design of the Zombie Escape paradigm. The experimental design includes four varying variables across different conditions: 1) The horizontal offset of both players' starting positions (Δx). 2) The vertical deviation of both players' starting positions from the middle of the screen (Δy). 3) The speed difference from the average speed (Δv). 4) The zombie's horizontal starting position (x_z).

In the two practice conditions (the left panel in Fig. 4.3), both players do not have horizontal offsets $\Delta x = 0$, their vertical distance from the middle of the screen is $\Delta y = 60$, and the speed difference is $\Delta v = 0.3v$. In the first practice condition, the speed of the bottom player is $1.3v$, while the speed of the bottom player is $0.7v$. In the other practice condition, the speeds are reversed. This helps players familiarize themselves with the speed differences in the formal conditions. The zombie's horizontal starting position is at $x_z = 800$.

The eighteen formal conditions can be categorized into two mega-conditions: the slower-as-baiter mega-condition (the middle panel in Fig. 4.3) and the closer-as-baiter mega-condition (the right panel in Fig. 4.3). Each mega-condition consists of nine conditions and follows a three-by-three design (three levels of horizontal offset Δx and three levels of speed difference Δv). The vertical deviation Δy and the zombie's horizontal starting position x_z are fixed within each mega-

condition. Specifically, in the slower-as-baiter mega-condition, $\Delta y = 30$ and $x_z = 880$. The three levels of Δx are [-25, 0, 25] and the three levels of Δv are [-0.3, 0, 0.3] $\times v$. In the closer-as-baiter mega-condition, $\Delta y = 60$ and $x_z = 660$. The three levels of Δx are [-35, 0, 35] and the three levels of Δv are [-0.3, 0, 0.3] $\times v$.

In the slower-as-baiter mega-condition, when there is a speed difference (conditions 1, 2, 3, 7, 8, 9), the player with the slower speed must be the baiter (i.e., the one caught by the zombie) for the dyad to succeed. Similarly, in the closer-as-baiter mega-condition, when there is a horizontal offset (conditions 10, 12, 13, 15, 16, 18), the player whose starting position is closer to the zombie either must be the baiter (conditions 12, 13, 15, 16) or, if they perform as the baiter, they can succeed (conditions 10, 18). By designing these two mega-conditions and systematically varying variables within, we created a rich set of environments that demand the dyad adapt to different conditions by taking on different roles.

4.2.4 Illustrative Game Play

To aid understanding of our paradigm and the necessity of the role of baiter, we provide some example game plays in Fig. 4.4, with different rows representing different trials and different columns representing different phases of a trial. The zombie is designed to chase the closer player. Although the zombie only indirectly cares about which player reaches the cure, without collaboration the dyad will fail the task because the zombie has a significant speed advantage.

In all three trials, both players head directly to the cure at the beginning (Fig.4.4 rows 1-3, column A). In the first trial (Fig.4.4 row 1), the blue player continues to move towards the cure even after confirming that the zombie is chasing them (Fig.4.4 row 1, column B). When the blue player is caught by the zombie, the red player is still far from the cure (Fig.4.4 row 1, column C), and they are eventually caught by the zombie as well (Fig.4.4 row 1, column D).

In the second trial (Fig.4.4 row 2), the game is solved as follows: when the blue player (or the red player) realizes they are being chased (Fig.4.4 row 2, column B), they head away from the cure

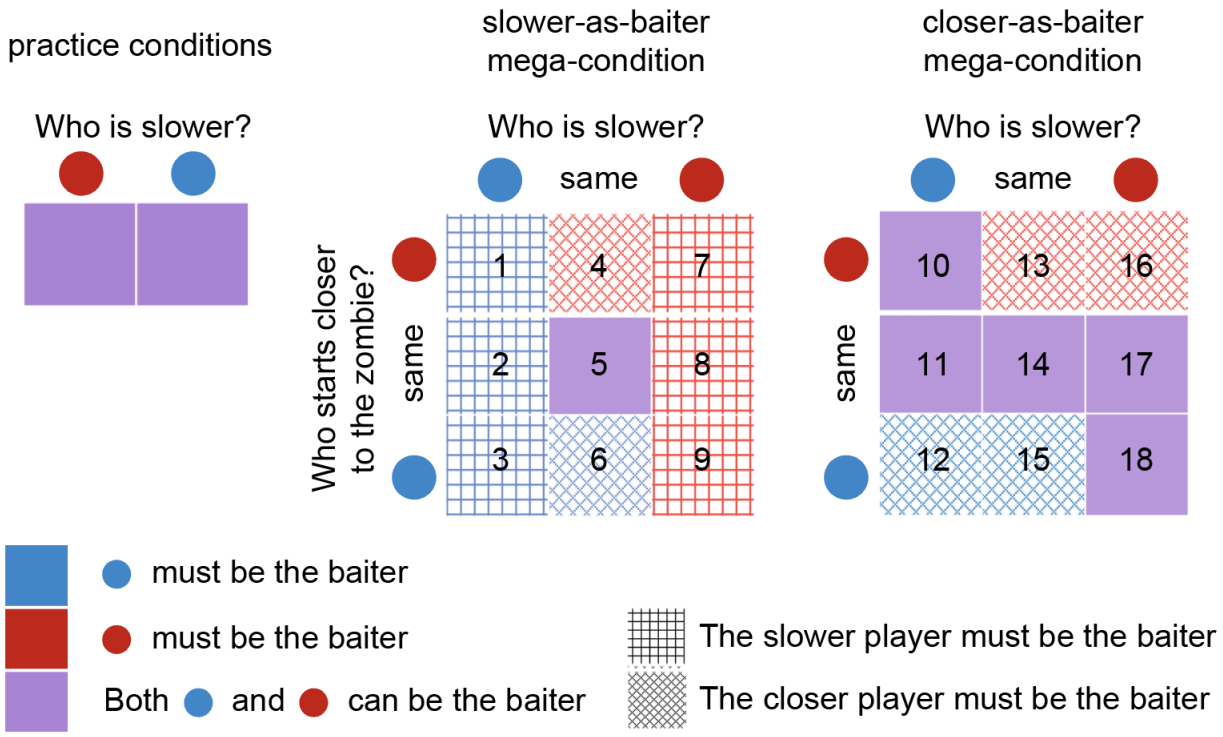

Figure 4.3: Experiment conditions and their corresponding strategies. Twenty conditions of the experiment can be broken down into three categories: two practice conditions (left), one slower-as-baiter mega-condition (middle), and one closer-as-baiter mega-condition (right). The two practice conditions differ in the speed difference (Δv). Each mega-condition follows a three-by-three design (three levels of horizontal offset Δx and three levels of speed difference Δv). In order to succeed in a condition, one player must be the baiter. Possible solutions for different conditions are plotted. The colors red and blue correspond to the player who must be the baiter in that condition, while the color purple means that either player can be the baiter to succeed. The rectangular grids represent conditions where only the slower player can be the baiter, while the diagonal grids represent conditions where only the player starting closer to the zombie can be the baiter.

to delay being caught by the zombie (Fig.4.4 row 2, column C), allowing their partner to rush to the cure without being caught (Fig. 4.4 row 2, column D).

However, if the player being chased only turns around without considering the position of their partner (Fig.4.4 row 3, column B), it might result in the zombie shifting its target in the midst of chasing (Fig.4.4 row 3, column C), causing the dyad to fail the task (Fig. 4.4 row 3, column D).

In summary, to succeed in the Zombie Escape, role coordination is necessary: one player needs to be the baiter while the other is the dasher. Moreover, even after coordinating proper roles, fine tuning their movements is crucial to solving the game.

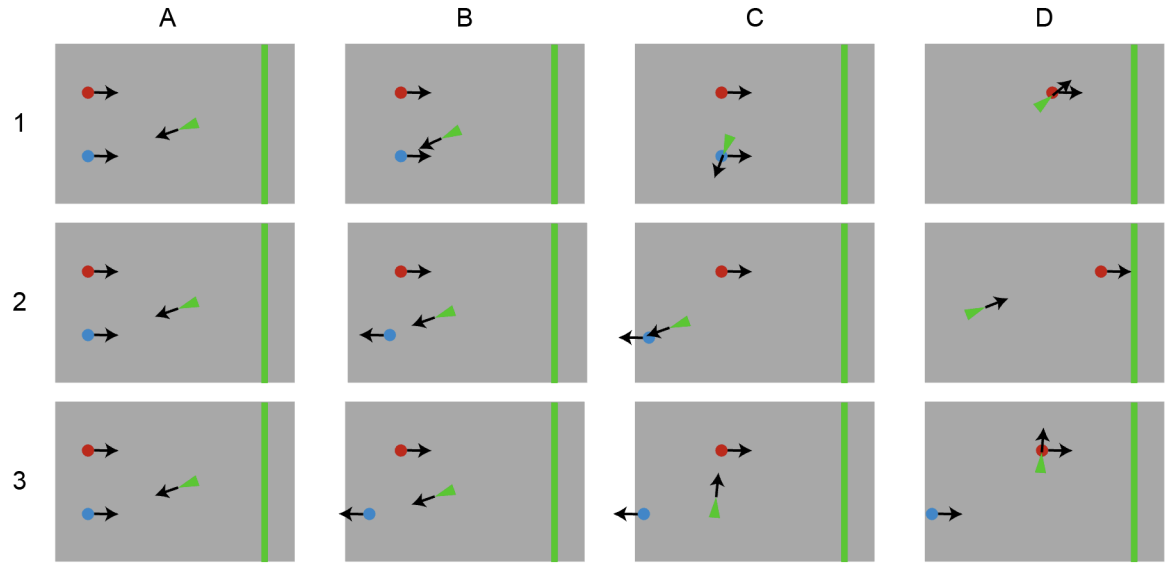

Figure 4.4: Example game plays of the Zombie Escape paradigm. Different rows (1-3) represent different trials, and different columns (A-D) represent different phases of a trial. The movement direction is denoted by the black arrow. The dyad only succeeds in row 2. In row 1, both players head to the cure from the start until the end (columns A to D), resulting in both being caught by the zombie one by one (columns C and D). In row 2, both players initially head to the cure (column A). When the blue player realizes they are the target of the zombie, they turn around at a specific position by taking into account their partner's position (column B), to bait the zombie farther away from their partner (column C), allowing the partner to reach the cure in time (column D). In row 3, the blue player, who is being chased, dodges away too early (column B), causing the zombie to shift its target mid-chase (column C), resulting in failure when their partner is caught before reaching the cure (column D).

4.3 Results

There are three levels of performance that can be examined here. The highest level is the across-condition level. The experiment consists of 2 practice plus 18 formal conditions (environments), so we can ask how people adapt to a new condition. The middle level is the within-condition level. At this level, we can ask how participants adapt their role coordination based on their history of collaborating in the condition. The lowest level is the within-trial level. There, we can ask how participants react to their partner's and the zombie's movements in real time. In this chapter, we focus on the highest and middle levels. We qualitatively described participants' behavior at both the across-condition and the within-condition level.

Before going into the detailed analysis of participants' behavior, we need some definitions. In the Introduction and Method sections, we roughly defined the roles of baiter and dasher. Here, we are providing stricter definitions. When the dyad succeeds, we define the player first caught by the zombie as the baiter and the player reaching the cure as the dasher. When the dyad fails, a slightly different definition is needed. At first glance, we could keep the definition that the baiter is the player first caught by the zombie. However, it is possible that one player tries to be the baiter but because of execution errors (such as in Fig. 4.4 row 3 column B), the other player is caught first. Therefore, on failure trials, we instead calculate the amount of time the zombie is chasing each player before the first player is caught. Whoever is being chased for a longer time is designated the baiter.

We would like to quantify a dyad's performance beyond mere success or failure in a trial. We propose using the distance of the last-surviving player to the cure when the trial ends, which we call the "margin". When the trial fails, we adopt this definition (Fig. 4.5A). However, when the trial succeeds, it stops as soon as the dasher reaches the cure, leading to a zero margin. To better quantify the success of a successful trial, we simulate the trial after the dasher reaches the cure: we let the dasher continue moving to the right and the zombie continue chasing until the dasher

is caught. Then we calculate the distance between the cure and the position where the dasher is finally caught and use the negative of this distance as the margin (Fig. 4.5B). The margin we define here provides a better description of how successful a trial is and enhances the analysis below about collaboration and role coordination.

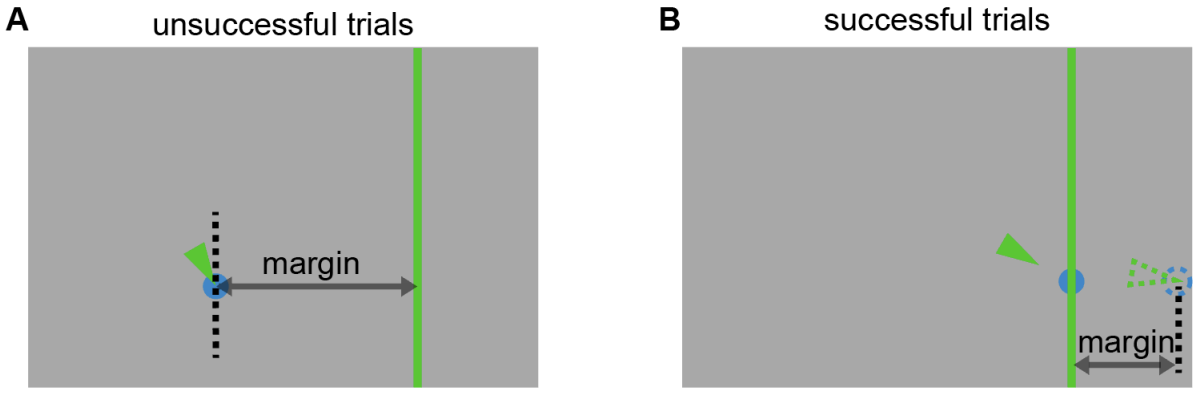

Figure 4.5: We define "margin" to quantify a dyad's performance. A) In unsuccessful trials, we use the distance of the last-surviving player to the cure when the trial ends. B) In successful trials, we simulate the trial after the dasher reaches the cure. We let the dasher continue moving to the right and the zombie continue chasing until the dasher is caught. Then we calculate the distance between the cure and the position where the dasher is finally caught and use the negative of this distance as the margin. The positions of the dashed-lined blue circle and the green triangle indicate where the dasher and the zombie end up in the simulation.

4.3.1 Across-condition Result

In the main analysis, we include only the data collected from the formal conditions. However, before excluding the practice data, we can explore several questions about participants' initial interactions with the game. For example, how many trials did it take before they achieved their first success, i.e., for the first time, i.e. successfully finding the collaborative baiter-dasher role differentiation? The average number is 2.0 ± 0.4 trials (Fig. 4.6A), demonstrating fast acquisition in the beginning of the game. Ten dyads were able to succeed on their first trial, learning the collaboration in a zero-shot way. Another interesting statistic is that nine dyads successfully tried out both role differentiations in their first practice condition (see Fig. 4.3). We neither provided

any incentive nor instructed them to do so. This demonstrated their exploratory nature in the collaborative game.

Then we investigated participants overall performance in the formal conditions. The average number of trials per condition was 7.2 ± 2.9 trials in the formal experiments out of a possible range of 3 to 12. The average number of successes per condition was 2.2 ± 0.1 trials in the formal experiments out of a possible range of 0 to 3. These two results suggest the Zombie Escape paradigm has an intermediate level of difficulty. This ensures that participants were not bored because it is too easy or too frustrated because it is too challenging.

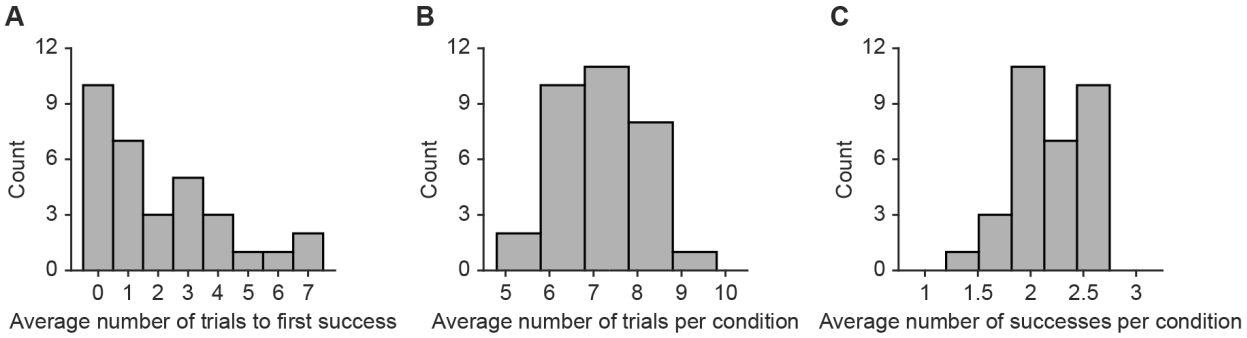

Figure 4.6: Participants' performance in the first practice and the formal conditions. A) Histogram of the number of trials before the first success in participants' first practice condition. The average number was 2.0 ± 0.4 trials, demonstrating a fast acquisition of the collaborative strategy at the beginning of the game. B) Histogram of the number of trials per condition in the formal conditions. The average number was 7.2 ± 2.9 trials out of a possible range of 3 to 12. C) Histogram of the number of successes per condition in the formal conditions. The average number was 2.2 ± 0.1 trials out of a possible range of 0 to 3.

Now we look into the proportion of success in all 18 formal conditions in Fig. 4.7. Despite there are experimental setting difference in different mega-conditions and possible strategies in different condition might be more than one. The general impression here is that the strategy to let slower player to be the baiter might be harder to find. We verified this by only taking into account the conditions that have a unique role coordination strategy (i.e. only one player can be the baiter). We found that in conditions with the unique strategy to allow the closer player as the baiter, participants were more likely to succeed $t(31) = 5.65, p = 3.36 \times 10^{-6}$, than those in

which the unique strategy is to allow the slower player as the baiter (Fig. 4.8A). This suggests that the default role differentiation in the Zombie Escape for the participants is to have the closer player to be the baiter. We further tested out this hypothesis focusing on the conditions only allowing the slower player to be the baiter. In these conditions, when the closer player is the slower player, which we regarded as "congruent", the participants were more likely to use the correct role coordination strategy, then when the closer player is the faster player, which we regarded as "incongruent" (Fig. 4.8B), $t(31) = 7.49, p = 1.93 \times 10^{-8}$.

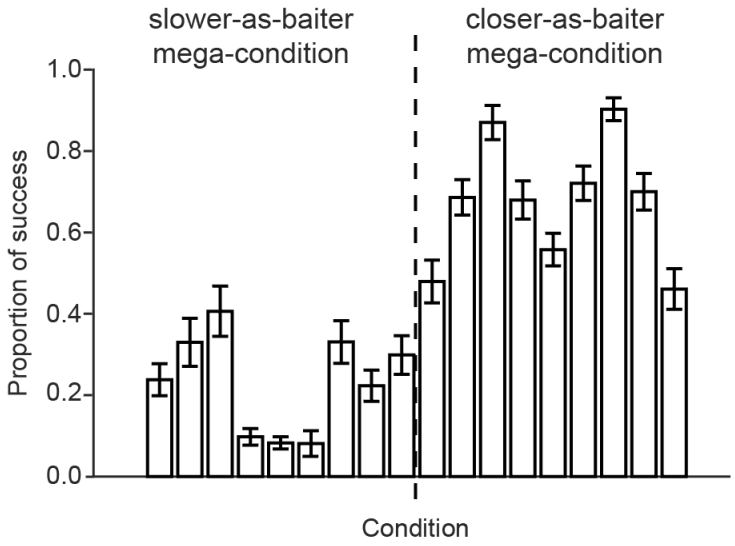

Figure 4.7: Participants' proportion of success in all 18 formal conditions. Error bars represent mean ±1 s.e.m across participants.

4.3.2 Within-condition Result

So far we have found that participants indeed were able to succeed in the game, and tried out the game with the set of mind to coordinate different roles when facing different conditions. How did this dynamic look like at the within-condition level? To answer this, we derived all instances of role switch, i.e. inside a condition, the role differentiation in the current trial different from the role differentiation in the last (previous) trial. We ran a generalized linear mixture model with the

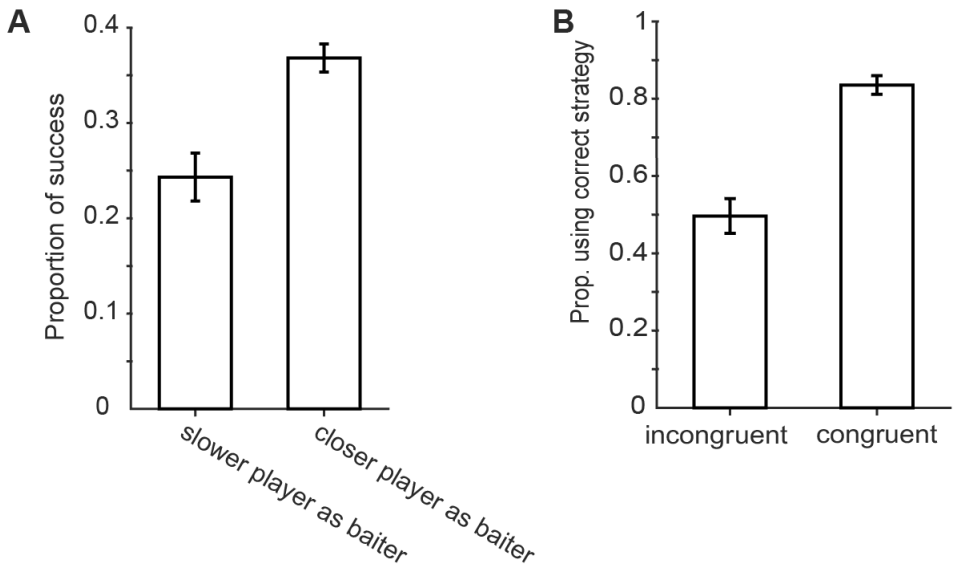

Figure 4.8: Participants' behaviors with varying required role coordination. A) Participants were more likely to succeed in conditions only allowing the closer player to be the baiter then conditions only allowing the slower player to be the baiter. B) In conditions only allowing slower player to be the baiter, participants were more likely to try the correct strategy in conditions that are congruent, i.e. the slower player is the closer player.

condition index and the dyad ID as random variables using the logit link function to statistically test the significance of the related variables. The list of independent variables include: whether the last trial succeeds, the margin of the last trial, the trial count in the current condition, the trial count in total, the number of fails of the last-trial role, the number of successes of the last-trial role, the condition count, and the mega condition. There was a clear trend of win-stay-lose-shift (Fig. 4.9A), i.e. the dyads were more likely to switch after a fail trial. However, the effect is insignificant, $z = 0.54, p = 0.59$, probably due to the high collinearity with another significant variable margin, $r = 0.73$. Since the number of successes of a specific role differentiation before the condition ends vary from 0 to 2, we also tested out if the number of successes of the role coordination in the last trial decreases the switch probability. The answer is yes (Fig. 4.9B), $z = -14.23, p < 2 \times 10^{-16}$. We found a non-linear inverse-U shape relationship between the switch probability and the number of fails of the last-trial role, (Fig. 4.9C), though the linear effect is also

significant $z = -13.85, p < 2 \times 10^{-16}$, meaning that more fails lead to less switches. To test out if the feeling of making progress in the last-trial role could decrease the switch probability, we plotted how the slope of the margin of the last-trial role (which meant that at least two trials of the same role coordination were finished) influenced the switch probability (Fig. 4.9D). However, there was not a visible trend. In the linear model we did not include the win rate of the last-trial role coordination to avoid collinearity issue, but we found a dipping pattern (Fig. 4.9E). The temporal progress of the current condition did have a significant influence on the switch probability (Fig. 4.9F), $z = 12.21, p < 2 \times 10^{-16}$, though the trend is not clear on the plot. The margin of the last trial also had an significant influence on the switch probability, meaning that if the last trial failed harder the dyads were more likely to switch roles (Fig. 4.9G), $z = 6.60, p = 4.2 \times 10^{-11}$. Effects of the rest independent variables are not significant, $ps > .05$.

4.4 Discussion

In this chapter, we developed a new collaborative paradigm Zombie Escape with continuous space and time. Participants must collaborate with each other to coordinate their roles in real time to succeed. We analyzed the performance and the dynamics of collaboration, with a keen interest in the role coordination. Participants quickly learnt to collaborate with each other and were able to complete the task without either finding it too easy or failing too much. Their performance varied across different conditions. The results suggested that there is a default role differentiation strategy in the game, that is to have the closer player to be the baiter. When this is against the required role coordination of a condition, participants failed more, spent more time in finding the right role coordination, and adopted the correct role differentiation less. Within a condition, we have found several variables dramatically increased the chance that the dyad will switch roles. Participants showed a win-stay-lose-shift pattern, switching roles more after a fail and switching roles more if a role differentiation fails harder in the last trial. If the dyad succeeded in a role

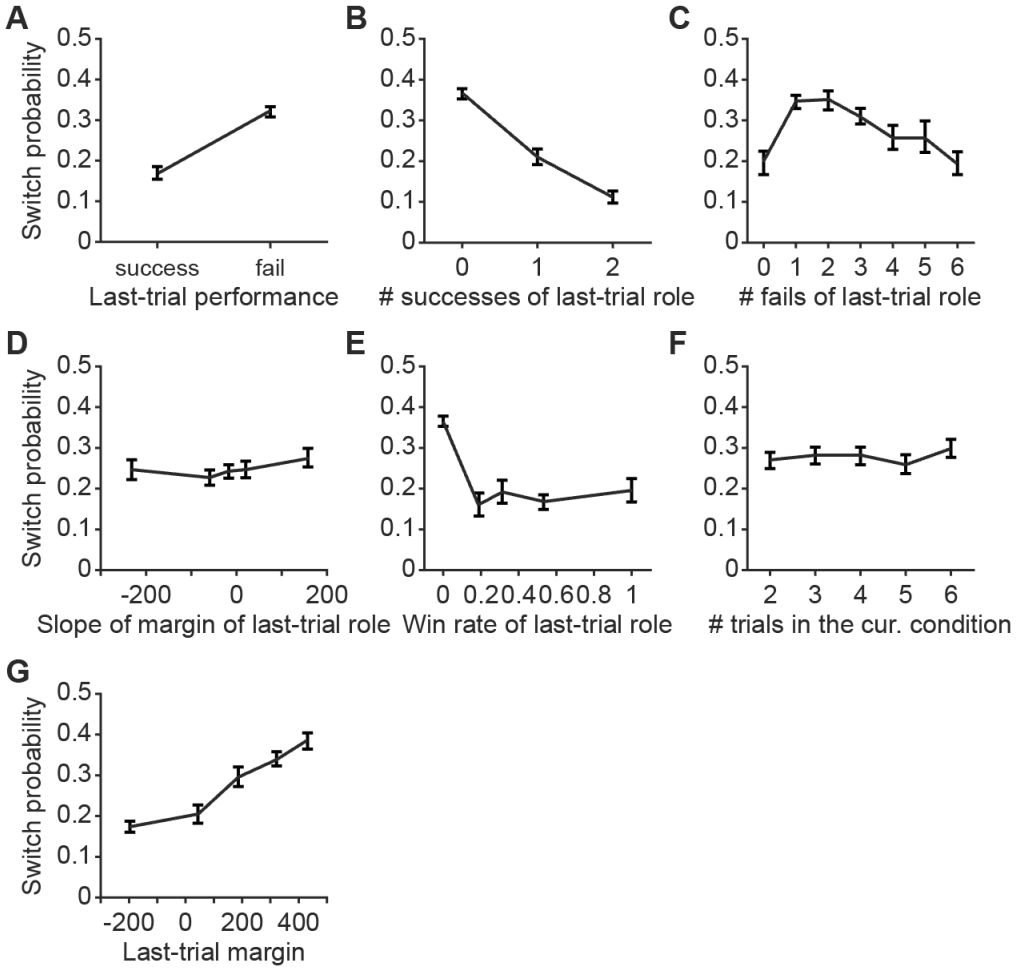

Figure 4.9: Probability of switching roles as a function of different variables. We plotted seven variables against the probability of switch roles, consisting of A) The performance (success/fail) of the last trial. B) The number of successes of the role differentiation of the last trial. C) The number of fails of the role differentiation of the last trial. D) The slope of the margin history of the role differentiation of the last trial. E) The win rate of the role differentiation of the last trial. F) The number of trials in the current condition. G) The margin of the last trial. Error bars represent mean ±1 s.e.m across participants.

differentiation before, it significantly decreased their chance of switching roles. As the number of fails of the role differentiation of the last trial increased, the switch probability showed an inverse U-shape.

The analysis in this chapter mostly focused on collaboration at the across-condition and within-condition levels. However, the Zombie Escape paradigm provides a richer dataset beyond these two levels, as we also have data at the within-trial level. This makes the paradigm very promising for investigating people's collaboration. Future directions include analyzing, for example, who initiates the role switch within a trial and how a player signals their intention to switch roles (e.g., by quickly moving back and forth, which is generally not useful to the task but serves as a form of "sign language").

Following the point of "signaling their intention," one interesting question is the role and influence of verbal communication in collaboration. In the current experiment, we did not allow any verbal communication between participants. However, in real life, people collaborate with highly intensive communication beyond mere behavioral interactions. We want to ask, for example, how participants would take advantage of verbal communication in diverse ways. They could discuss the overall strategy or cheer each other up during the break between conditions, discuss different roles and possible switches within a condition between trials, or coordinate movements inside a trial. Will they reach consensus faster, or will they be too overwhelmed by the game happening in real time to communicate effectively? These questions about verbal communication will be open to investigation if we allow participants to chat during the experiment.

Although we are very specific about the game design of the Zombie Escape, such as the differences in players' starting positions, we hope these results could generalize beyond these specific settings. Our aim is to analyze collaboration happening in real time, and with so many environment variables freely adjustable, we wanted to create a subset of conditions to encourage different collaborative behaviors in the paradigm. One obvious omission here is how different zombie strategies can influence player collaboration and how people adapt to these strategies.

In this paradigm, we only investigate situations where the zombie chases the closer player, as stated explicitly in the instructions. Alternative zombie strategies include, for example, chasing the player closer to the cure or the player who is expected to reach the cure given the current moving direction and speed. Future research might be able to answer these questions by setting up different zombie agents with drastically different chasing strategies.

In the Zombie Escape, we focused on how people collaborate in pairs. However, in real life, collaboration often involves more than two people. While existing research has already explored multi-person collaboration (Tang et al., 2022), it is crucial to dig deeper about role coordination in larger groups. For example, in a recent study of online multiplayer games (Vélez, 2024), researchers investigated how different players took on different roles in a online multiplayer survival game. It will be interesting to explore how we can either build a richer set of roles in an experimental control setting while still maintaining the naturalness of collaboration or applying our modeling method to their datasets.

5 | Computational Models of Role Coordination

Abstract

Role coordination is crucial in collaboration. To better understand the role coordination dynamics within conditions, we developed computational process models to test several hypotheses derived from the experiment, treating each dyad as a unit. The main model in the chapter, Bayesian linear extrapolation (BLE) model, assumed each dyad deliberates between possible role differentiations, summarizing past performance through linear regression and extrapolating future success probabilities. This model was compared to the classic win-stay-lose-shift (WSLS) model. The BLE model successfully captured how switch probabilities varied with trial performance, demonstrating a superior fit to the data compared to the the WSLS model. Key findings included the BLE model's ability to explain inverse-U patterns in switch probabilities with respect the number of fails of the last-trial role and its response to the quantitative performance of the last trial. Further analysis revealed the BLE model's strength in reasoning about the value of different coordinated roles when with no success. Our findings contribute to a better computational explanation of role coordination in human collaboration and have implications for building more adaptive and human-like AI systems. Future research should consider modeling individual agents' opinions and decision-making processes.

Keywords: collaboration; role coordination; role differentiation; Bayesian linear regression; extrapolation; hierarchical learning; win-stay-lose-shift

5.1 Introduction

In the last chapter, we developed a new real-time cooperative game called Zombie Escape to investigate how people collaborate and allocate roles in a challenging task. We found that participants were able to quickly learn to collaborate and showed interesting patterns about how they collaborated. Among these patterns, the most thought-provoking one was how the dyad decided the roles based on their history of performance. However, so far we have only used statistical tests or linear regressions to check the significance of related variables, without digging into the process of deciding roles. In this chapter, our goal is to build a computational model to explain the within-condition role coordination dynamics as the trial progresses.

Recently, researchers in the field of artificial intelligence have made significant breakthroughs in large-scale collaborative games. Computer agents are now able to achieve above human-level performance in highly collaborative board games like Hanabi (Bard et al., 2020) and video games like Dota 2 (Berner et al., 2019). One immediate thought would be to borrow frameworks from these papers to study their counterparts in humans. However, all these models are based on deep reinforcement learning with a heavy load of parameters. Therefore, while they might serve as good benchmark models, they provide little insight into the process of human collaboration, especially regarding role coordination.

To gain insight into potential models for human collaboration, we focus on studies that explicitly consider other agents with their own minds, especially models incorporating theory of mind (TOM). Researchers have found Bayesian TOM, where the observer actively infers the preferences of the agent from their movements and predicts their future actions accordingly, to be highly successful in predicting human beliefs (e.g., Baker et al., 2017). In this model, hierarchical planning is adopted; in the generative model direction, a high-order variable (e.g., preferences) guides a low-level variable (e.g., movements). In the context of collaboration, since multiple agents have their own beliefs, the concept of level K reasoning is involved. Level 0 reasoning treats all teammates

as passive parts of the environment and acts accordingly. Level 1 reasoning treats all teammates as level 0 reasoning agents and acts accordingly, and so on. A study showed that participants adapted their level of reasoning when playing with computer agents of different and fixed levels of reasoning (Yoshida et al., 2008). These promising results suggest that models with hierarchical structure built have the potential to capture the process of human collaboration.

What do we mean by hierarchical learning here? Botvinick (2012) said "Whereas a standard reinforcement learning agent selects among primitive actions, the hierarchical reinforcement learning agent can also select among subroutines". In the Zombie Escape, the exact movements are the primitive actions while the role coordination is the subroutine. Thus, the intuition of hierarchical learning in our paradigm is as such: the dyad tries out a potential role differentiation and improve the performance; if they find no improvement (i.e. no hope of success with the current coordinated roles), they will switch to another role settings.

In this chapter, we aimed to develop a model with the specified hierarchical structure equipped to account for the role coordination dynamics at the within-condition level. In the model we built, Bayesian linear extrapolation (BLE) model, the agent tracks their history of performance within a condition, extrapolates their probability of succeeding at least once based on the history, and decides between the two role differentiation based on the extrapolation result. The BLE model was able to account for how the probability of switching roles changes with all the variables we came up with, and outperformed a classic strategy: win-stay-lose-shift both qualitatively and quantitatively. We further investigated why the BLE model could explain the data and found out that the ability of the BLE model to reason about the value of different role differentiation played an important role when there were only fails .

5.2 Methods

We built computational process model to test several conjectures we derived in the Zombie Escape experiment chapter, In this section, we treat each dyad as a unit without specifying process happening on the individual participant level and model their role coordination sequence happening in a condition. Before moving forward, we specify the terms needed in the Table. 5.1.

Notation	Meaning
A_t	role coordination on the t-th trial
$A_t = a_1$	top player is the baiter
$A_t = a_2$	bottom player is the baiter
$A_{0:t} \triangleq (A_0, A_1, ..., A_t)$	role coordination history from the 0-th to the t-th trial
n	the horizon of a condition
S_t	$\in \{0, 1\}$, whether the t-th trial is successful
$S_{0:t} \triangleq (S_0, S_1, ..., S_t)$	success/fail history from the 0-th to the t-th trial
M_t	margin on the t-th trial
$M_{0:t}^{(j)}$	margin history from the 0-th to the t-th trial of role coordination a_j

Table 5.1: Notations for modeling the within-condition role switch.

5.2.1 Bayesian Linear Extrapolation Model

To further investigate how dyads adapt their roles within a condition, we built a Bayesian linear extrapolation (BLE) model to account for this. The intuition here is that participants in the dyad deliberate between two possible role differentiations: either they themselves will be the baiter or their partner will be the baiter. They summarize both role differentiations' previous history of results in the current condition using linear regression and extrapolate into the future to estimate the probability of succeeding at least once in the remaining trials. They then coordinate their roles in the next trial on the difference of the probability of success between the two possible role differentiations.

Why do we need Bayesian linear regression instead of a simple linear regression? It is because, at the beginning of the trial, the model starts with either no history or very limited history of

margins for both role differentiations. The vanilla linear regression can only work when there are at least two data-points to fit and three data-points to estimate confidence interval. By adopting Bayesian linear regression, the intended probability of success can be calculated with limited data or even no data at all.

At each trial, the dyad evaluates the probability of achieving at least one success before encountering too many failures (i.e. fails ten times in total), assuming that the margin in the future continues along the current trend. After estimating this probability for both role differentiations, they will base on it to decide if to continue the current role coordination or switch. Therefore, the utmost important part of the modeling is to calculate the probability of succeeding at least once in the game-allowed horizon of the role coordination a_j. Mathematically, it is equivalent to one minus the probability of all fail in the rest of the trials, i.e.

$$
\begin{aligned}
& p_{\text{win}}^{(j)} \\
\triangleq & p^{(j)}(\text{at least one success in the rest of the trials}) \\
= & 1 - p^{(j)}(\text{no success in the rest of the trials})
\end{aligned}
$$

If we know the prior of linear regression parameters, it would be possible to calculate $p^{(j)}(\text{no success in the rest of the trials})$ by

$$
\begin{aligned}
& p^{(j)}(\text{no success in the rest of the trials}|M_{0:t-1}^{(j)}) \\
= & p(\min\{M_{t:n}^{(j)}\} > 0|M_{0:t-1}^{(j)}) \quad i = t, t+1, \cdots, n \\
= & \iiint p(\min\{M_{t:n}^{(j)}\} > 0|a, b, \sigma^2) p(a, b, \sigma^2|M_{0:t-1}^{(j)}) da db d\sigma^2 \\
= & \iiint p(\min\{M_{t:n}^{(j)}\} > 0|a, b, \sigma^2) \frac{p(a, b, \sigma^2) p(M_{0:t-1}^{(j)}|a, b, \sigma^2)}{p(M_{0:t-1}^{(j)})} da db d\sigma^2 \\
= & \iiint \left[\prod_{i=t}^{n} p(M_i^{(j)} > 0|a, b, \sigma^2) \right] \frac{p(a)p(b)p(\sigma^2) \left[\prod_{i=0}^{t-1} p(M_i^{(j)}|a, b, \sigma^2) \right]}{p(M_{0:t-1}^{(j)})} da db d\sigma^2
\end{aligned}
$$

where a is the slope, b is the intercept, σ^2 is the variance term of the normal-distributed noise in the Bayesian linear regression, and $p(M_i^{(j)}|a,b,\sigma^2) = \mathcal{N}(M_i^{(j)}; ai + b, \sigma^2)$. We assume the independence of the prior of slope, intercept, and the variance. For slope and intercept, we assume the prior follows the normal distribution, while for the variance, we assume the prior follow the inverse gamma distribution. Each individual participant has their own prior and this prior is shared across all conditions and all role differentiations. This adds up to 6 parameters (two for each prior).

The above triple integral is hard to calculate analytically. In order to calculate this, we sample from the priors of the slope, the intercept, and the variance. After getting the parameter set, we first calculate the likelihood of each parameter set given the margin history $p(a,b,\sigma^2|M_{0:t-1}^{(j)})$. Then for each sampled parameter set, we calculate $p_{\text{win}}^{(j)}$ and aggregate it weighted by the likelihoods to derive an samples-based estimate of $p_{\text{win}}^{(j)}$ (Fig. 5.1B).

After getting $p_{\text{win}}^{(1)}$ and $p_{\text{win}}^{(2)}$, i.e. the probabilities of succeeding at least once in each role differentiation, we derive the probability that the dyad will let top player be the baiter as a logistic function:

$$\log \frac{p(A_t = a_1|M_{0:t-1}^{(1)}, M_{0:t-1}^{(2)}, A_{t-1})}{1 - p(A_t = a_1|M_{0:t-1}^{(1)}, M_{0:t-1}^{(2)}, A_{t-1})} = \beta \left(p_{\text{win}}^{(1)} - p_{\text{win}}^{(2)} + B_1 + \text{sgn}(A_{t-1} = a_1) B_{\text{stay}} \right)$$

where β is the inverse temperature term, B_1 is the bias term for the top player to be the baiter, and the last term is the bias resulting from the dyad's unwillingness to switch roles, which is 0 if it is the first trial in a condition $t = 0$, positive if the top player is the baiter in the trial $t-1$, $A_{t-1} = a_1$, and negative if the bottom player is the baiter in the trial $t-1$, $A_{t-1} = a_2$. The logistic function has 3 additional free parameters, bringing the total number of the free parameters of the BLE model to 9.

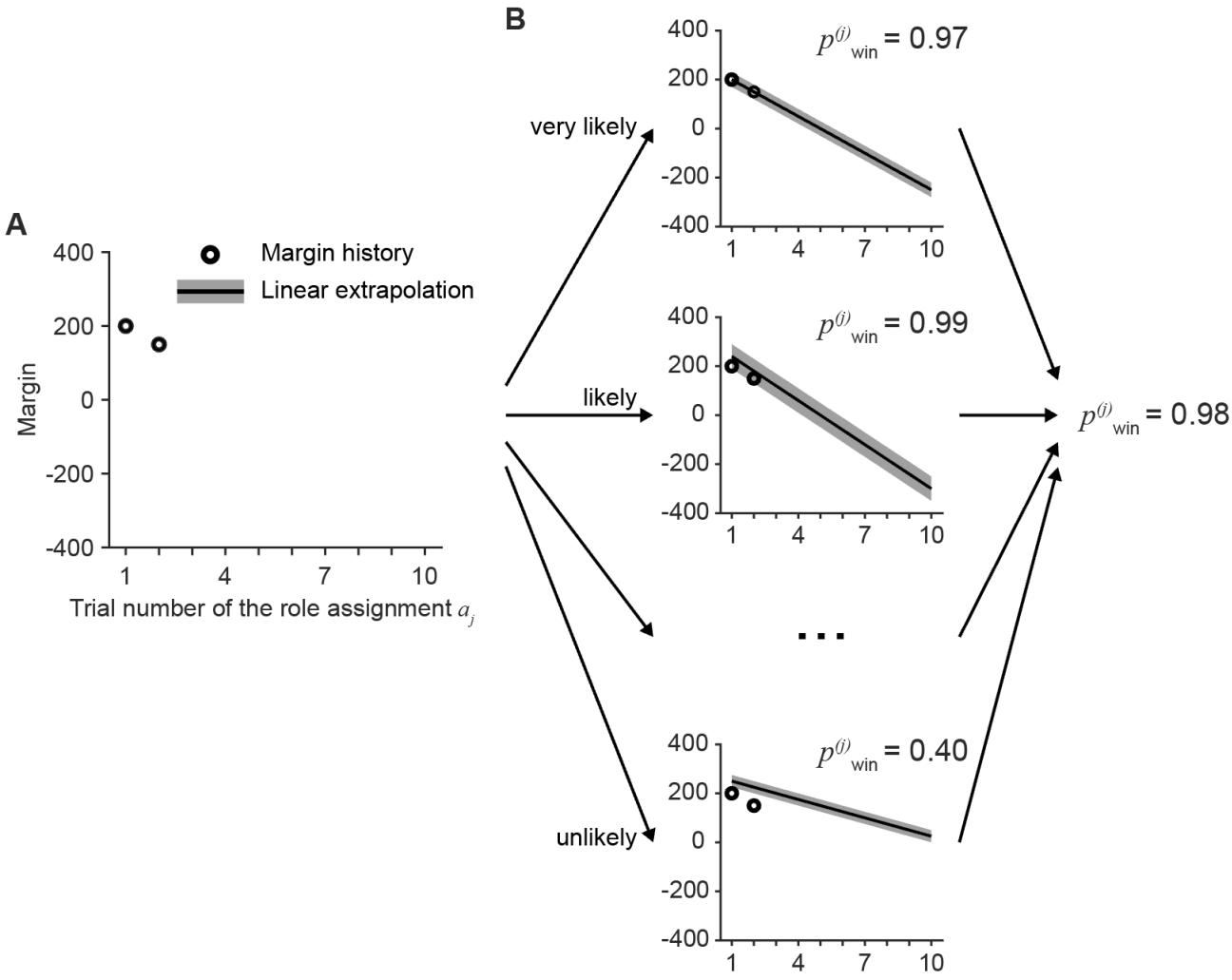

Figure 5.1: Illustration of Bayesian linear extrapolation model. A) For the role differentiation a_j, the dyad tried two times but all failed, with the margin history shown as black circle. B) In order to calculate the probability of at least succeeding once in the rest of the condition if the dyad holds on to the current role coordination $p^{(j)}_{\text{win}}$, we sample linear regression parameters from the prior and calculate the likelihood of each parameter set, illustrated by "very likely", "likely", or "unlikely". Then for each sampled parameter set, we calculate $p^{(j)}_{\text{win}}$ and aggregate it weighted by the likelihoods.

5.2.2 Win-stay-lose-shift Model

The alternative model we consider here is the classic win-stay-lose-shift (WSLS) model that is very intuitive (Nowak & Sigmund, 1993). When the dyad succeeds in the last trial, they are more likely to keep the coordianted roles in the last trial. When the dyad fails, they are more likely to switch roles in the next trial. The WSLS model only pays attention to the success/fail outcome of trials without taking into account the exact performance (e.g. margin). Therefore, we have

$$
\begin{aligned}
p(A_0 = a_1) &= q, \\
p(A_0 = a_2) &= 1 - q \\
p(A_t | A_{0:t-1}, S_{0:t-1}) &= p(A_t | A_{t-1}, S_{t-1}) \\
&= ((1 - w_s) S_{t-1} + (1 - w_f)(1 - S_{t-1}))^{\delta(A_i = A_{i-1})} (w_s S_{t-1} + w_f (1 - S_{t-1}))^{\delta(A_i \neq A_{i-1})}
\end{aligned}
$$

where q is the bias term, w_s is the switch probability parameter when the last trial succeeds, and w_f is the switch probability parameter when the last trial fails. In sum, there are three free parameters in total. Therefore, the log likelihood for a condition is calculated as

$$
\begin{aligned}
&\log L \\
&= \log p[A_{0:n} | S_{0:n}] \\
&= \log [p(A_0) \prod_{t=1}^{n} p(A_t | A_{t-1}, S_{t-1})] \\
&= \log p(A_0) + \sum_{t=1}^{n} \log p(A_t | A_{t-1}, S_{t-1})
\end{aligned}
$$

5.2.3 Model Fitting.

We fit each model to the data of each dyad by maximizing the log likelihood of $\boldsymbol{\theta}$, $\log L(\boldsymbol{\theta}) = \log p(\text{data} | \boldsymbol{\theta})$. We assume that the conditions are conditionally independent. Then, the log

likelihood becomes:

$$
\begin{aligned}
&\log L(\boldsymbol{\theta}) \\
&= \log p(\text{data} \mid \boldsymbol{\theta}) \\
&= \log \prod_k p(\text{data in condition } k \mid \boldsymbol{\theta}) \\
&= \sum_k \log p(\text{data in condition } k \mid \boldsymbol{\theta})
\end{aligned}
$$

We optimized the parameters for each individual using a new method called Bayesian Adaptive Direct Search (Acerbi & Ma, 2017).

5.2.4 Model Comparison

We used AIC for model comparison, which is calculated by the equation $\text{AIC} = -\log L^* + 2k$, where L^* is the maximum value of a model's likelihood function and k is the number of fitted parameters. To report the AIC, we computed the AIC for each dyad and summed them up. The confidence interval of the group-summed AIC was estimated by non-parametric bootstrapping.

5.3 Results

The modeling results are shown in Fig. 5.2 for the BLE model and Fig. 5.3 for the WSLS model. We plotted seven matrices against the probability of switching roles as in the last result section. Both models are also able to predict the win-stay-lose-shift trend in the data (Fig. 5.3A). The BLE model successfully matches the data as how the switch probability decreases as the number of successes of the role differentiation in the last trial increases (Fig. 5.2B). The WSLS shows a similar trend,

but fails to match quantitatively (Fig. 5.3B). The most striking result here is that how the BLE model follows the inverse-U pattern about how the switch probability changes with the number of fails of the role differentiation in the last trial (Fig. 5.2C), which is completely missing in the WSLS model (Fig. 5.3C). We will later dig deeper about why the BLE model can fit this trend. The step-function-like pattern of how the switch probability changes with the slope of margin history of the role differentiation of the last trial can be explained by the BLE model (Fig. 5.2D); however, this can not be captured by the WSLS model (Fig. 5.3D). The dipping pattern of how the switch probability changes with the win rate of the role differentiation of the last trial is predicted by the BLE model (Fig. 5.2E). The WSLS model shows a decreasing trend here but cannot explain the dip (Fig. 5.3E). Both BLE and WSLS models capture how the switch probability stays flat as the trial progresses in the current condition (Fig. 5.3F, (Fig. 5.3F). The BLE model successfully accounts for how the switch probability monotonically increases with the margin of the last-trial (Fig. 5.2G). However, for the WSLS model, although showing a increasing trend when the margin is negative, the trend disappears when the margin is very positive (Fig. 5.3G). In summary, the BLE model provides a much better fit. This is confirmed in the model comparison of AIC: $\Delta\text{AIC} = 54.5$, 95% confidence interval = $[22.2, 91.4]$.

One surprisingly good model fitting result is how the BLE model can explain the inverse-U pattern of how the switch probability changes as the number of fails of the role differentiation of the last trial increases. Inside the BLE model, we have a bias term about sticking to the same role differentiation in the last trial but we do not have an extra bonus if the role differentiation in the last trial happens a lot in the current condition. One way to investigate this is to split the switch probability in the Fig. 5.2C by if the role differentiation of the last trial has succeeded before in the current condition (Fig. 5.4A). When the dyad has succeeded using the last-trial role, the switch probably stays low and is not sensitive to the number of fails of the last-trial role. However, when the dyad has not succeeded using the last-trial role, the switch probability drops dramatically. Therefore, the inverse-U shape we observe previously is the outcome of combining

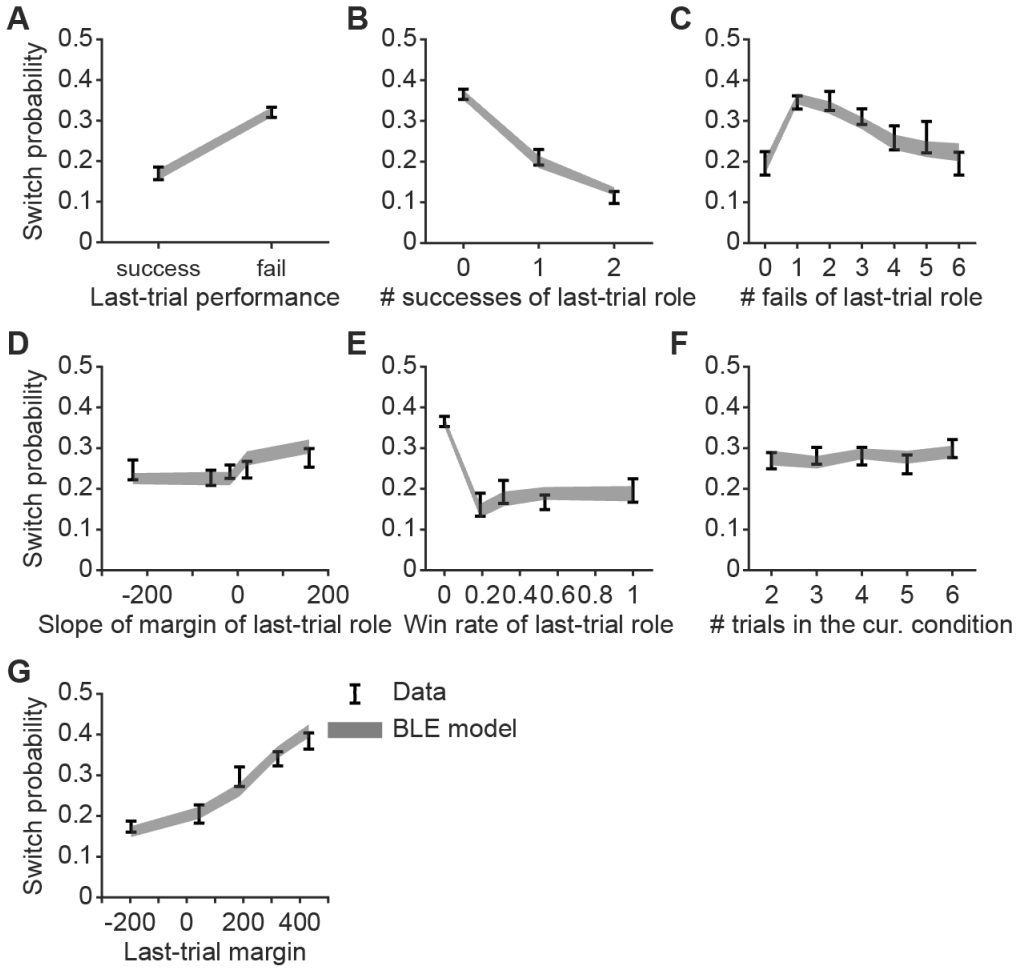

Figure 5.2: Model fitting results of the BLE model. We plotted seven variables against the probability of switch roles, consisting of A) The performance (success/fail) of the last trial. B) The number of successes of the role differentiation of the last trial. C) The number of fails of the role differentiation of the last trial. D) The slope of the margin history of the role differentiation of the last trial. E) The win rate of the role differentiation of the last trial. F) The number of trials in the current condition. G) The margin of the last trial. The black error bars are the data, and the grey shaded error bars are the model fitting results. Error bars and shaded areas represent mean ±1 s.e.m across participants.

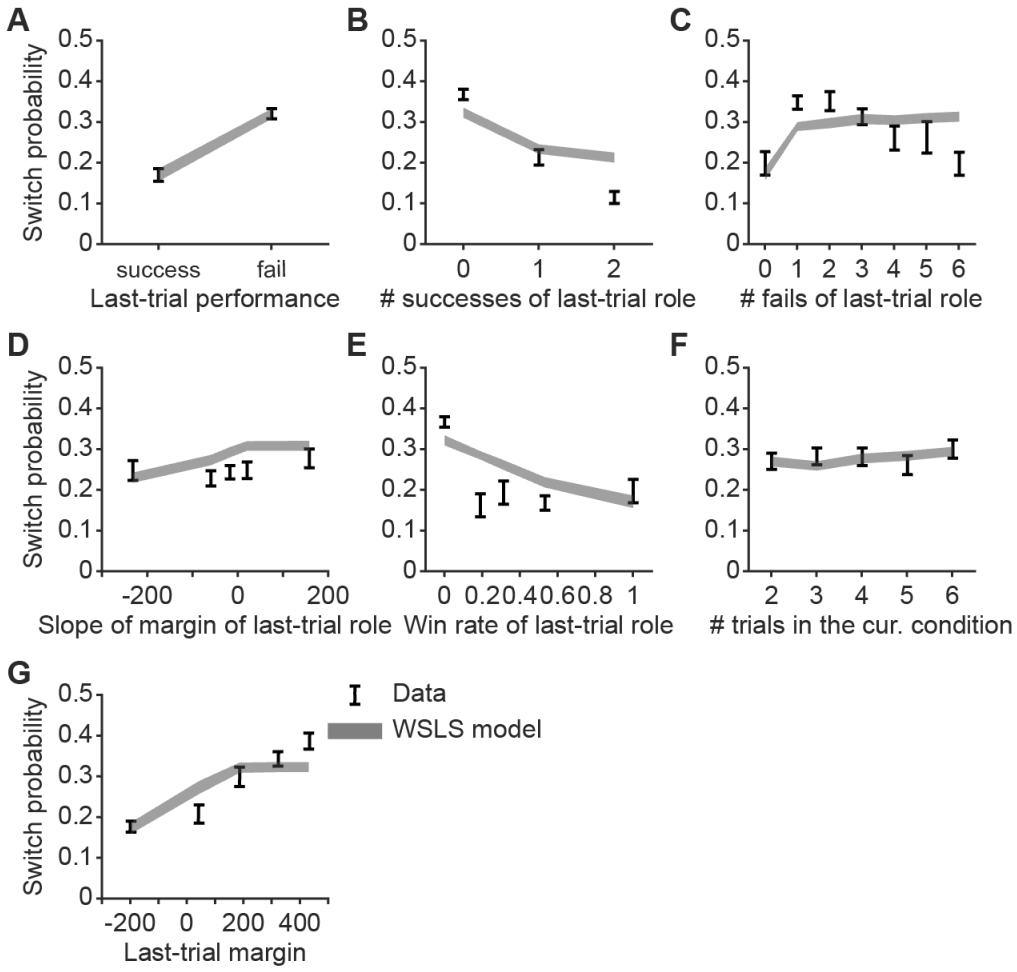

Figure 5.3: Model fitting results of the WSLS model.We plotted seven variables against the probability of switch roles, consisting of A) The performance (success/fail) of the last trial. B) The number of successes of the role differentiation of the last trial. C) The number of fails of the role differentiation of the last trial. D) The slope of the margin history of the role differentiation of the last trial. E) The win rate of the role differentiation of the last trial. F) The number of trials in the current condition. G) The margin of the last trial. The black error bars are the data, and the grey shaded error bars are the model fitting results. Error bars and shaded areas represent mean ±1 s.e.m across participants.

two rather distinct processes.

The flat trend is intuitive: when the dyad has succeeded before on a role differentiation, they might be confident on that role differentiation so failing more does not change their thoughts much. But the decreasing trend still remains as a mystery. To understand this, for these trials when the last-trial role has not succeeded before, we plotted the extrapolated probability of succeeding at least once in the rest of the condition if the dyad switch roles $p_{\text{win}}^{\text{switch}}$ or not $p_{\text{win}}^{\text{stay}}$ in Fig. 5.4B. The two extrapolated succeeding probabilities do not differ from each other when the number of fails of the last-trial role is low, but greatly diverges as $p_{\text{win}}^{\text{stay}}$ increases and $p_{\text{win}}^{\text{switch}}$ decreases as the number of fails of the last-trial role increases. The result suggests that even when the dyad had no success at all in a condition, they still actively deliberated on the history of the outcomes of the trials so far and collectively adopted the role differentiation with the higher possibility to succeed over the course of the condition.

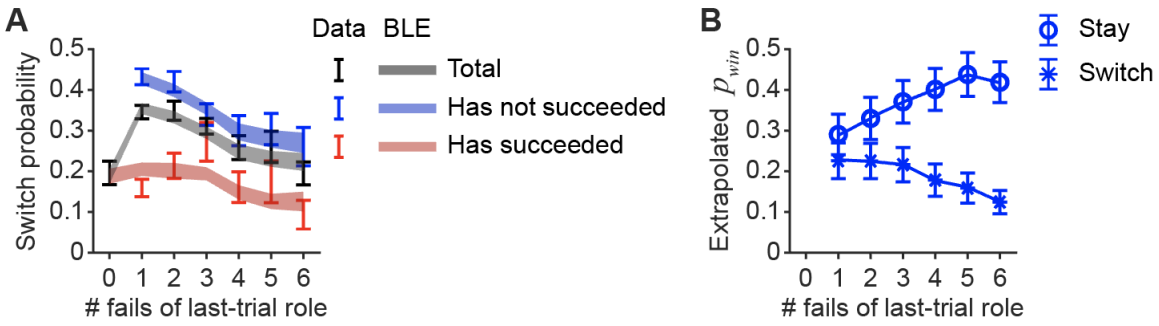

Figure 5.4: Model fitting results of the WSLS model.We plotted seven variables against the probability of switch roles, consisting of A) How the switch probably changes with the number of the fails of the role differentiation in the last trial, split by if the role differentiation has succeeded before or not. The blue color is when the dyad has not succeeded using the role differentiation of the last trial in the current condition, and the red color is when the dyad has. The black color is the aggregated result. The error bars represent the data, while the shaded error bars represent the model fitting result of the BLE model. B) We further investigated the model fitting result of the BLE model of the "has not succeeded" group in the panel A. We calculated the extrapolated probability of succeeding at least once during the rest of the horizon of the condition $p_{\text{win}}^{(j)}$ either if the dyad sticks to the role differentiation in the last trial or if the dyad switch roles. $p_{\text{win}}^{\text{stay}}$ increases while the $p_{\text{win}}^{\text{switch}}$ decreases as the number of the fails of the role differentiation in the last trial increases. Error bars and shaded areas represent mean ±1 s.e.m across participants.

5.4 Discussion

In this chapter, we developed a model called Bayesian linear extrapolation to account for the role coordination dynamics in the within-condition level. In the BLE model, the agent tracks their history of performance within a condition, extrapolates their probability of succeeding at least once based on the history, and decides between the two ways of role differentiation based on the extrapolation result. The BLE model was able to account for how the probability of switching roles changes with all the variables we came up with, and outperformed a classic strategy: win-stay-lose-shift both qualitatively and quantitatively. We further investigated why the BLE model could explain the data and found out that the ability of the BLE model to reason about the value of different ways of role differentiation when there were fails only played an important role.

One missing component in the current modeling is that we treated the dyad as a whole. Since the two agents in the game are controlled by two players, it is likely that each agent has a different opinion about which role they and their partner should take when coordinating roles. What we actually modeled in this chapter, in one way, was the outcome of the coordination of the dyad. In future work, we might be able to model the two agents separately, with each having their own evaluation of how good different role differentiations are based on the history of performance, and then deciding. It is also worth noting that players have different power in deciding role differentiations under varying conditions. For example, one intuition is that the faster player or the player starting closer to the zombie is more likely to decide the role differentiation. Thus, in the future, we might take these factors into consideration by further extending our BLE model.

Research on how humans collaborate, especially the role coordination component, is important in building AI that interacts with humans (Carroll et al., 2019). The investigation in this chapter, studying how humans decide to switch roles, provides a better computational explanation and helps with building machines that help humans.

6 | Conclusion

In this dissertation, we explored the temporal processes of decision-making in two distinct subfields of psychology: perceptual decision-making and collaborative decision-making.

In Chapter 2, we focused on perceptual decision-making. We developed a novel framework of approximate inference through active sampling of likelihoods (AIAS) and applied it to the tasks of perceptual categorization. AIAS accounts for several empirical findings. First, it explains the recent puzzling finding that decision confidence follows the difference between the two highest posteriors, rather than the highest posterior itself (H.-H. Li & Ma, 2020). AIAS not only provides better fits but also accurately predicts response times based on the number of iterations. Second, we demonstrated that AIAS can explain how categorization behavior changes with varying visual contrast (Adler & Ma, 2018). Third, we found that the mean response times predicted by AIAS grow approximately logarithmically with the number of categories N, consistent with Hick's law (Hick, 1952). By developing AIAS and successfully applying it to three datasets, we uncovered the temporal processes of perceptual decision-making by casting approximate inference as an active-sampling process with imprecise computations.

In Chapters 4 and 5, we turned our focus to collaborative decision-making. We developed a two-player real-time collaborative game to investigate how human players collaborate, with a particular interest in role coordination. We found that participants quickly adapted to the game and effectively coordinated roles across different environments. Additionally, participants tracked their performance history and used this information for future role coordination. We

then constructed a Bayesian linear extrapolation model to explain the temporal processes of role coordination observed in the experiment. This model successfully accounted for the data and outperformed a win-stay-lose-shift model, highlighting the complex temporal patterns of role coordination in collaboration.

Overall, the work presented in this dissertation examined the temporal processes of decision-making. Our research enhances the understanding of human decision-making by shedding light on the underlying temporal processes.

Bibliography

Acerbi, L. (2020). Variational bayesian monte carlo with noisy likelihoods. *Advances in neural information processing systems*, *33*, 8211–8222.

Acerbi, L., & Ma, W. J. (2017). Practical Bayesian optimization for model fitting with Bayesian adaptive direct search. *Advances in neural information processing systems*, *30*.

Acerbi, L., Vijayakumar, S., & Wolpert, D. M. (2014). On the origins of suboptimality in human probabilistic inference [Publisher: Public Library of Science San Francisco, USA]. *PLoS computational biology*, *10*(6), e1003661.

Adler, W. T., & Ma, W. J. (2018). Comparing Bayesian and non-Bayesian accounts of human confidence reports [Publisher: Public Library of Science San Francisco, CA USA]. *PLoS computational biology*, *14*(11), e1006572.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, *19*(6), 716–723.

al-Haytham, I. (1989). *The Optics of Ibn al-Haytham: Books I-III: On Direct Vision*. (A. I. Sabra, Trans.). Warburg Institute, University of London.

Andersen, S. O., Sarma, K. M., & Taddonio, K. N. (2012). *Technology transfer for the ozone layer: Lessons for climate change*. Routledge.

Andrade-Lotero, E., & Goldstone, R. L. (2021). Self-organized division of cognitive labor [Publisher: Public Library of Science San Francisco, CA USA]. *PloS one*, *16*(7), e0254532.

Auer, P. (2002). Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, *3*(Nov), 397–422.

Baer, C., & Odic, D. (2022). Mini managers: Children strategically divide cognitive labor among collaborators, but with a self-serving bias [Publisher: Wiley Online Library]. *Child Development*, *93*(2), 437–450.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4).

Balland, P.-A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P., Rigby, D. L., & Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities [Publisher: Nature Publishing Group UK London]. *Nature human behaviour*, *4*(3), 248–254.

Bard, N., Foerster, J. N., Chandar, S., Burch, N., Lanctot, M., Song, H. F., Parisotto, E., Dumoulin, V., Moitra, S., Hughes, E., et al. (2020). The hanabi challenge: A new frontier for ai research [Publisher: Elsevier]. *Artificial Intelligence*, *280*, 103216.

Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., & Pouget, A. (2012). Not noisy, just wrong: The role of suboptimal inference in behavioral variability [Publisher: Elsevier]. *Neuron*, *74*(1), 30–39.

Berner, C., Brockman, G., Chan, B., Cheung, V., Dębiak, P., Dennison, C., Farhi, D., Fischer, Q., Hashme, S., Hesse, C., et al. (2019). Dota 2 with large scale deep reinforcement learning. *arXiv preprint arXiv:1912.06680*.

Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Tai National Park [Publisher: Wiley Online Library]. *American journal of physical anthropology*, *78*(4), 547–573.

Botvinick, M. M. (2012). Hierarchical reinforcement learning and decision making [Publisher: Elsevier]. *Current opinion in neurobiology*, *22*(6), 956–962.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search [Publisher: Elsevier]. *Cognition*, *106*(3), 1465–1477.

Browning, L., & Colman, A. M. (2004). Evolution of coordinated alternating reciprocity in repeated dyadic games [Publisher: Elsevier]. *Journal of theoretical biology*, *229*(4), 549–557.

Bullinger, A. F., Melis, A. P., & Tomasello, M. (2011). Chimpanzees, Pan troglodytes, prefer individual over collaborative strategies towards goals [Publisher: Elsevier]. *Animal Behaviour*, *82*(5), 1135–1141.

Bullinger, A. F., Wyman, E., Melis, A. P., & Tomasello, M. (2011). Coordination of chimpanzees (Pan troglodytes) in a stag hunt game [Publisher: Springer]. *International Journal of Primatology*, *32*, 1296–1310.

Callaway, F., Rangel, A., & Griffiths, T. L. (2021). Fixation patterns in simple choice reflect optimal information sampling [Publisher: Public Library of Science San Francisco, CA USA]. *PLoS computational biology*, *17*(3), e1008863.

Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton university press.

Carroll, M., Shah, R., Ho, M. K., Griffiths, T., Seshia, S., Abbeel, P., & Dragan, A. (2019). On the utility of learning about humans for human-ai coordination. *Advances in neural information processing systems*, *32*.

Cranmer, K., Brehmer, J., & Louppe, G. (2020). The frontier of simulation-based inference [Publisher: National Acad Sciences]. *Proceedings of the National Academy of Sciences*, *117*(48), 30055–30062.

Drugowitsch, J., Wyart, V., Devauchelle, A.-D., & Koechlin, E. (2016). Computational precision of mental inference as critical source of human choice suboptimality [Publisher: Elsevier]. *Neuron*, *92*(6), 1398–1411.

Du, Y., Leibo, J. Z., Islam, U., Willis, R., & Sunehag, P. (2023). A review of cooperation in multi-agent learning. *arXiv preprint arXiv:2312.05162*.

Duguid, S., & Melis, A. P. (2020). How animals collaborate: Underlying proximate mechanisms [Publisher: Wiley Online Library]. *Wiley Interdisciplinary Reviews: Cognitive Science*, *11*(5), e1529.

Duguid, S., Wyman, E., Grueneisen, S., & Tomasello, M. (2020). The strategies used by chimpanzees (Pan troglodytes) and children (Homo sapiens) to solve a simple coordination problem. [Publisher: American Psychological Association]. *Journal of Comparative Psychology*, *134*(4), 401.

Estes, R. D., & Goddard, J. (1967). Prey selection and hunting behavior of the African wild dog [Publisher: JSTOR]. *The Journal of Wildlife Management*, 52–70.

Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., & Wyart, V. (2019). Computational noise in reward-guided learning drives behavioral variability in volatile environments [Publisher: Nature Publishing Group US New York]. *Nature neuroscience*, *22*(12), 2066–2077.

Fleming, S. M. (2024). Metacognition and confidence: A review and synthesis [Publisher: Annual Reviews]. *Annual Review of Psychology*, *75*(1), 241–268.

Garnett, R. (2023). *Bayesian optimization*. Cambridge University Press.

Gauvain, M., & Rogoff, B. (1989). Collaborative problem solving and children's planning skills. [Publisher: American Psychological Association]. *Developmental psychology*, *25*(1), 139.

Gazda, S. K., Connor, R. C., Edgar, R. K., & Cox, F. (2005). A division of labour with role specialization in group–hunting bottlenose dolphins (Tursiops truncatus) off Cedar Key, Florida [Publisher: The Royal Society]. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1559), 135–140.

Gershman, S., & Goodman, N. (2014). Amortized inference in probabilistic reasoning [Issue: 36]. *Proceedings of the annual meeting of the cognitive science society*, *36*.

Gershman, S. J., Vul, E., & Tenenbaum, J. B. (2012). Multistability and perceptual inference [Publisher: MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . .]. *Neural computation*, *24*(1), 1–24.

Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli [Publisher: Elsevier]. *Trends in cognitive sciences*, *5*(1), 10–16.

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. [Publisher: American Psychological Association]. *Developmental psychology*, *45*(5), 1430.

Gräfenhain, M., Carpenter, M., & Tomasello, M. (2013). Three-year-olds' understanding of the consequences of joint commitments [Publisher: Public Library of Science San Francisco, USA]. *PloS one*, *8*(9), e73039.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic [Publisher: Wiley Online Library]. *Topics in cognitive science*, *7*(2), 217–229.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition [Publisher: SAGE Publications Sage CA: Los Angeles, CA]. *Psychological science*, *17*(9), 767–773.

Grosse, G., Moll, H., & Tomasello, M. (2010). 21-month-olds understand the cooperative logic of requests [Publisher: Elsevier]. *Journal of Pragmatics*, *42*(12), 3377–3383.

Haefner, R. M., Berkes, P., & Fiser, J. (2016). Perceptual decision-making as probabilistic inference by neural sampling [Publisher: Elsevier]. *Neuron*, *90*(3), 649–660.

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees [Publisher: Nature Publishing Group UK London]. *Nature*, *476*(7360), 328–331.

Hargadon, A., & Sutton, R. I. (1997). Technology brokering and innovation in a product development firm [Publisher: JSTOR]. *Administrative science quarterly*, 716–749.

Helbing, D., Schönhof, M., Stark, H.-U., & Hołyst, J. A. (2005). How individuals learn to take turns: Emergence of alternating cooperation in a congestion game and the prisoner's dilemma [Publisher: World Scientific]. *Advances in Complex Systems*, *8*(01), 87–116.

Hennig, P., Osborne, M. A., & Girolami, M. (2015). Probabilistic numerics and uncertainty in computations [Publisher: The Royal Society Publishing]. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *471*(2179), 20150142.

Herce Castañón, S., Moran, R., Ding, J., Egner, T., Bang, D., & Summerfield, C. (2019). Human noise blindness drives suboptimal cognitive inference [Publisher: Nature Publishing Group UK London]. *Nature communications*, *10*(1), 1719.

Hick, W. E. (1952). On the rate of gain of information [Publisher: Taylor & Francis]. *Quarterly Journal of experimental psychology*, *4*(1), 11–26.

Hidalgo, C. A., Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations [Publisher: American Association for the Advancement of Science]. *Science*, *317*(5837), 482–487.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, *79*(8), 2554–2558.

Hu, H., Lerer, A., Peysakhovich, A., & Foerster, J. (2020). "other-play" for zero-shot coordination. *International Conference on Machine Learning*, 4399–4410.

Jang, A. I., Sharma, R., & Drugowitsch, J. (2021). Optimal policy for attention-modulated decisions explains human fixation behavior [Publisher: eLife Sciences Publications, Ltd]. *Elife*, *10*, e63436.

Ji, J., Qiu, T., Chen, B., Zhang, B., Lou, H., Wang, K., Duan, Y., He, Z., Zhou, J., Zhang, Z., et al. (2023). AI alignment: A comprehensive survey. *arXiv preprint arXiv:2310.19852*.

Johnson, D. W., & Johnson, R. T. (2002). Learning together and alone: Overview and meta-analysis [Publisher: Taylor & Francis]. *Asia Pacific Journal of Education*, *22*(1), 95–105.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., & Saul, L. K. (1999). An introduction to variational methods for graphical models [Publisher: Springer]. *Machine learning*, *37*, 183–233.

Kaanders, P., Nili, H., O'Reilly, J. X., & Hunt, L. (2021). Medial frontal cortex activity predicts information sampling in economic choice [Publisher: Soc Neuroscience]. *Journal of Neuroscience*, *41*(40), 8403–8413.

Kepecs, A., & Mainen, Z. F. (2012). A computational framework for the study of confidence in humans and animals [Publisher: The Royal Society]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1322–1337.

Khan, Z., Balch, T., & Dellaert, F. (2004). A rao-blackwellized particle filter for eigentracking. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, *2*, II–II.

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex [Publisher: American Association for the Advancement of Science]. *Science*, *324*(5928), 759–764.

Kleiman-Weiner, M., Ho, M. K., Austerweil, J. L., Littman, M. L., & Tenenbaum, J. B. (2016). Coordinate to cooperate or compete: Abstract goals and joint intentions in social interaction. *CogSci*.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press.

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning [Publisher: Nature Publishing Group UK London]. *Nature*, *427*(6971), 244–247.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions [Publisher: National Acad Sciences]. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857.

Lange, R. D., Chattoraj, A., Beck, J. M., Yates, J. L., & Haefner, R. M. (2021). A confirmation bias in perceptual decision-making due to hierarchical approximate inference [Publisher: Public Library of Science San Francisco, CA USA]. *PLoS Computational Biology*, *17*(11), e1009517.

Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Zeitschrift für Naturforschung c*, *36*(9-10), 910–912.

Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost of neural information [Publisher: Nature Publishing Group]. *Nature neuroscience*, *1*(1), 36–41.

Lennie, P. (2003). The cost of cortical computation [Publisher: Elsevier]. *Current biology*, *13*(6), 493–497.

Li, H.-H., & Ma, W. J. (2020). Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis [Publisher: Nature Publishing Group UK London]. *Nature communications*, *11*(1), 2004.

Li, M., Kwon, M., & Sadigh, D. (2021). Influencing leading and following in human–robot teams [Publisher: Springer]. *Autonomous Robots*, *45*, 959–978.

Li, Z.-W., & Ma, W. J. (2021). An uncertainty-based model of the effects of fixation on choice [Publisher: Public Library of Science San Francisco, CA USA]. *PLoS computational biology*, *17*(8), e1009190.

Lieder, F., Griffiths, T. L., M. Huys, Q. J., & Goodman, N. D. (2018). The anchoring bias reflects rational use of cognitive resources [Publisher: Springer]. *Psychonomic bulletin & review*, *25*, 322–349.

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12-and 18-month-olds point to provide information for others [Publisher: Taylor & Francis]. *Journal of cognition and development*, *7*(2), 173–187.

Ma, W. J., Kording, K. P., & Goldreich, D. (2023). *Bayesian models of perception and action: An introduction*. MIT press.

Magid, R. W., DePascale, M., & Schulz, L. E. (2018). Four-and 5-year-olds infer differences in relative ability and appropriately allocate roles to achieve cooperative, competitive, and

prosocial goals [Publisher: MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . .]. *Open Mind*, *2*(2), 72–85.

Maples, W. R. (1969). Adaptive behavior of baboons [Publisher: Wiley Online Library]. *American Journal of Physical Anthropology*, *31*(1), 107–109.

Marjoram, P., Molitor, J., Plagnol, V., & Tavaré, S. (2003). Markov chain Monte Carlo without likelihoods [Publisher: National Acad Sciences]. *Proceedings of the National Academy of Sciences*, *100*(26), 15324–15328.

Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games [Publisher: JSTOR]. *The American Economic Review*, *84*(3), 658–673.

Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees recruit the best collaborators [Publisher: American Association for the Advancement of Science]. *Science*, *311*(5765), 1297–1300.

Melis, A. P., & Tomasello, M. (2013). Chimpanzees'(Pan troglodytes) strategic helping in a collaborative task [Publisher: The Royal Society]. *Biology letters*, *9*(2), 20130009.

Melis, A. P., & Tomasello, M. (2019). Chimpanzees (Pan troglodytes) coordinate by communicating in a collaborative problem-solving task [Publisher: The Royal Society]. *Proceedings of the Royal Society B*, *286*(1901), 20190408.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines [Publisher: American Institute of Physics]. *The journal of chemical physics*, *21*(6), 1087–1092.

Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-month-olds know what "we" have shared in a special way [Publisher: Wiley Online Library]. *Infancy*, *13*(1), 90–101.

Morley, L., & Cashell, A. (2017). Collaboration in health care [Publisher: Elsevier]. *Journal of medical imaging and radiation sciences*, *48*(2), 207–216.

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game [Publisher: Nature Publishing Group UK London]. *Nature*, *364*(6432), 56–58.

Oster, G. F., & Wilson, E. O. (1978). *Caste and ecology in the social insects*. Princeton University Press.

Ostrom, E., Gardner, R., & Walker, J. (1994). *Rules, games, and common-pool resources*. University of Michigan press.

Page, R. E., & Mitchell, S. D. (1990). Self organization and adaptation in insect societies [Issue: 2]. *PSA: Proceedings of the biennial meeting of the philosophy of science association*, *1990*, 289–298.

Peters, G. W., Fan, Y., & Sisson, S. A. (2012). On sequential Monte Carlo, partial rejection control and approximate Bayesian computation [Publisher: Springer]. *Statistics and Computing*, *22*, 1209–1222.

Plötner, M., Over, H., Carpenter, M., & Tomasello, M. (2015). The effects of collaboration and minimal-group membership on children's prosocial behavior, liking, affiliation, and trust [Publisher: Elsevier]. *Journal of experimental child psychology*, *139*, 161–173.

Prins, N. (2012). The psychometric function: The lapse rate revisited [Publisher: The Association for Research in Vision and Ophthalmology]. *Journal of Vision*, *12*(6), 25–25.

Raileanu, R., Denton, E., Szlam, A., & Fergus, R. (2018). Modeling others using oneself in multi-agent reinforcement learning. *International conference on machine learning*, 4257–4266.

Ratcliff, R. (1978). A theory of memory retrieval. [Publisher: American Psychological Association]. *Psychological review*, *85*(2), 59.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history [Publisher: Elsevier]. *Trends in cognitive sciences*, *20*(4), 260–281.

Rekers, Y., Haun, D. B., & Tomasello, M. (2011). Children, but not chimpanzees, prefer to collaborate [Publisher: Elsevier]. *Current Biology*, *21*(20), 1756–1758.

Rullán Buxó, C., & Savin, C. (2021). A sampling-based circuit for optimal decision making. *Advances in Neural Information Processing Systems*, *34*, 14163–14175.

Sanborn, A., & Griffiths, T. (2007). Markov chain Monte Carlo with people. *Advances in neural information processing systems*, *20*.

Savin, C., Dayan, P., & Lengyel, M. (2014). Optimal recall from bounded metaplastic synapses: Predicting functional adaptations in hippocampal area CA3 [Publisher: Public Library of Science San Francisco, USA]. *PLoS computational biology*, *10*(2), e1003489.

Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory [Publisher: Elsevier]. *Neuron*, *90*(5), 927–939.

Sheridan, T. B. (2016). Human–robot interaction: Status and challenges [Publisher: SAGE Publications Sage CA: Los Angeles, CA]. *Human factors*, *58*(4), 525–532.

Shi, L., & Griffiths, T. (2009). Neural implementation of hierarchical Bayesian inference by importance sampling. *Advances in neural information processing systems*, *22*.

Shivkumar, S., Lange, R., Chattoraj, A., & Haefner, R. (2018). A probabilistic population code based on neural samples. *Advances in neural information processing systems*, *31*.

Sisson, S. A., Fan, Y., & Beaumont, M. (2018). *Handbook of approximate Bayesian computation*. CRC press.

Sisson, S. A., Fan, Y., & Tanaka, M. M. (2007). Sequential monte carlo without likelihoods [Publisher: National Acad Sciences]. *Proceedings of the National Academy of Sciences*, *104*(6), 1760–1765.

Song, M., Wang, X., Zhang, H., & Li, J. (2019). Proactive information sampling in value-based decision-making: Deciding when and where to saccade [Publisher: Frontiers Media SA]. *Frontiers in human neuroscience*, *13*, 35.

Stander, P. E. (1992). Cooperative hunting in lions: The role of the individual [Publisher: Springer]. *Behavioral ecology and sociobiology*, *29*, 445–454.

Stasser, G., & Stewart, D. (1992). Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgment. [Publisher: American Psychological Association]. *Journal of personality and social psychology*, *63*(3), 426.

Stewart, D. D., & Stasser, G. (1995). Expert role assignment and information sampling during collective recall and decision making. [Publisher: American Psychological Association]. *Journal of personality and social psychology*, *69*(4), 619.

Strouse, D., McKee, K., Botvinick, M., Hughes, E., & Everett, R. (2021). Collaborating with humans without human data. *Advances in Neural Information Processing Systems*, *34*, 14502–14515.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tang, N., Gong, S., Zhao, M., Gu, C., Zhou, J., Shen, M., & Gao, T. (2022). Exploring an imagined "we" in human collective hunting: Joint commitment within shared intentionality [Issue: 44]. *Proceedings of the annual meeting of the cognitive science society*, *44*.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples [Publisher: Oxford University Press]. *Biometrika*, *25*(3-4), 285–294.

Tomasello, M., & Hamann, K. (2012). The 37th sir frederick bartlett lecture: Collaboration in young children [Publisher: SAGE Publications Sage UK: London, England]. *Quarterly Journal of Experimental Psychology*, *65*(1), 1–12.

Tomasello, M., & Vaish, A. (2013). Origins of human cooperation and morality [Publisher: Annual Reviews]. *Annual review of psychology*, *64*, 231–255.

Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure [Publisher: JSTOR]. *The American Economic Review*, *80*(1), 234–248.

van Opheusden, B., Acerbi, L., & Ma, W. J. (2020). Unbiased and efficient log-likelihood estimation with inverse binomial sampling [Publisher: Public Library of Science San Francisco, CA USA]. *PLoS computational biology*, *16*(12), e1008483.

Vélez, N. (2024). Collaboration in One Hour One Life.

Vul, E., Alvarez, G., Tenenbaum, J., & Black, M. (2009). Explaining human multiple object tracking as resource-constrained approximate inference in a dynamic probabilistic model. *Advances in neural information processing systems*, *22*.

Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples [Publisher: Wiley Online Library]. *Cognitive science*, *38*(4), 599–637.

Wang, T., Gupta, T., Mahajan, A., Peng, B., Whiteson, S., & Zhang, C. (2020). Rode: Learning roles to decompose multi-agent tasks. *arXiv preprint arXiv:2010.01523*.

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees [Publisher: Wiley Online Library]. *Child development*, *77*(3), 640–663.

Warneken, F., Gräfenhain, M., & Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities [Publisher: Wiley Online Library]. *Developmental science*, *15*(1), 54–61.

Warneken, F., Lohse, K., Melis, A. P., & Tomasello, M. (2011). Young children share the spoils after collaboration [Publisher: Sage Publications Sage CA: Los Angeles, CA]. *Psychological science*, *22*(2), 267–273.

Warneken, F., Steinwender, J., Hamann, K., & Tomasello, M. (2014). Young children's planning in a collaborative problem-solving task [Publisher: Elsevier]. *Cognitive Development*, *31*, 48–58.

Wu, S. A., Wang, R. E., Evans, J. A., Tenenbaum, J. B., Parkes, D. C., & Kleiman-Weiner, M. (2021). Too many cooks: Bayesian inference for coordinating multi-agent collaboration [Publisher: Wiley Online Library]. *Topics in Cognitive Science*, *13*(2), 414–432.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. [Publisher: American Psychological Association]. *Psychological review*, *114*(2), 245.

Yang, S. C.-H., Lengyel, M., & Wolpert, D. M. (2016). Active sensing in the categorization of visual patterns [Publisher: eLife Sciences Publications, Ltd]. *Elife*, *5*, e12215.

Yang, S. C.-H., Wolpert, D. M., & Lengyel, M. (2016). Theoretical perspectives on active sensing [Publisher: Elsevier]. *Current opinion in behavioral sciences*, *11*, 100–108.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring [Publisher: The Royal Society]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–1321.

Yoshida, W., Dolan, R. J., & Friston, K. J. (2008). Game Theory of Mind (T. Behrens, Ed.). *PLoS Computational Biology*, *4*(12), e1000254.