Intuiting Informavores

An empirical investigation of the recognition and evaluation of epistemic actions in infancy

Bálint Varga

Supervisor: Ágnes Melinda Kovács

Secondary Supervisor: Gergely Csibra

Submitted to Central European University,

Department of Cognitive Science

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Cognitive Science

Vienna, Austria

2024

Declaration of Authorship

I, the undersigned, Bálint Varga, candidate for the PhD degree in Cognitive Sciene, declare herewith that the present dissertation is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the dissertation infringes on any person's or institution's copyright. I also declare that no part of the dissertation has been submitted in this form to any other institution of higher education for an academic degree.

The dissertation includes work that appears in the following papers/manuscripts:

- Varga, B., Csibra, G., & Kovács, Á. M. (2021). Infants' interpretation of information-seeking actions. In *Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 43, pp. 472–478)*. Cognitive Science Society.
- Varga, B., & Kovács, Á. M. (under review). The epistemic sense: Human infants readily appreciate that information bears value for other individuals.

Other parts of the dissertation contain research that was performed collaboratively and will be submitted for publication with the following co-authors:

- Chapter 5 with Ágnes Melinda Kovács, Ernő Téglás
- Chapter 6 with Ágnes Melinda Kovács, Barbara Pomiechowska

Vienna, 28 August 2024

Bálint Varga

Copyright Notice

Acknowledgments

Ági Kovács — for believing in my ideas, encouraging me to move forward, brainstorming with me, guiding me to avoid conceptual and methodological pitfalls, and investing endless time and effort in mentoring me through this venture. It has been an honor.

Gergő Csibra — for his critical insight and clarity of thought.

Petra Kármán, Dorka Kerschner, Dóri Mészégető, Iulia Savos, Timi Sugár, Bori Széplaki-Köllőd, Mari Tóth, and the hundreds of babies and toddlers who participated in our experiments — for making these studies possible.

Réka Finta, Edit Vitrai, and Ági Volein — for handling the administration, so I could stay focused on the science.

Ernő Téglás and Nicolò Cesana-Arlotti — for giving me the chance to contribute to their ingenious projects.

Julian Jara-Ettinger and Marlene Berke — for providing endless inspiration and hospitality while working together on our project, which I hope we will finish one day.

Rachel Dudley, Gyuri Gergely, Christophe Heintz, Velisar Manea, Hanna Marno, Mara Mavridaki, Basia Pomiechowska, Tibor Tauzin, Ernő Téglás, and Ricardo Vieira — for helpful discussions on various studies and ideas.

Martin Dockendorff, Barbu Revencu, and Laura Schlingloff-Nemecz — for being my friends and academic companions.

Balázs Demény, Beni Kállai, Zoli Opánszky, Dani Pucz, Kristóf Szőcs, Márton Szőcs, and Boga Varga — for being my friends.

My parents — for their love, and respect of my decision to pursue an academic career.

Tímea Éva Bogya — for her love, patience, and belief in me.

Bendő — for being the best boy and the most adamant explorer I have ever known. This dissertation is dedicated to his memory.

Abstract

Research on cognitive development often portrays infants as little scientists, constructing intuitive theories to explain observed phenomena and actively gathering information to test these theories. This dissertation investigates whether infants can intuit when and why *others* pursue information, exploring the foundations that allow humans to interpret epistemic actions in their social environment.

In Chapter 1, I motivate the research project and discuss the inductive challenges observers face in recognizing epistemic goals. I propose that preverbal infants may possess an intuitive theory of goal-directed action that includes expectations about how others seek information. This intuitive theory is proposed to contain three principles: instrumental relevance, drive toward novelty, and efficiency.

Chapter 2 focuses on the first principle, *instrumental relevance*, which posits that individuals seek information that helps them achieve their higher-order goals. This chapter details Experiments 1 and 2, designed to assess whether 14-month-old infants expect an agent to seek information when uncertain about a goal object's location. We found that infants looked longer when a knowledgeable agent sought information unnecessarily, as opposed to an uncertain agent, which indicates an understanding that epistemic actions may be driven by uncertainty about goal-relevant states.

Chapter 3 explores the second principle, *drive toward novelty*, which suggests that individuals are naturally inclined to seek information about new or unfamiliar objects. Experiment 3 aimed to determine if infants apply this principle when observing others' behavior. Results indicated that infants expect agents to explore novel objects as opposed to familiar ones, providing evidence for the idea that infants use novelty as a cue for interpreting actions as epistemic.

Chapter 4 investigates the third principle, *efficiency*, which proposes that individuals gather information in a way that maximizes benefits while minimizing costs. Experiment 4 tested whether infants expect agents to act efficiently when seeking information, choosing the

more informative from two actions with equal cost. As hypothesized, infants looked longer when an agent chose a less informative action over a more informative one, suggesting that they expect epistemic behavior to be efficient.

Chapter 5 presents two experiments examining whether infants and older children (4-year-olds) use a modal representation of others' uncertainty (attributing beliefs representing a set of possibilities) when interpreting their behavior. The mixed results showed that while 4-year-olds (Experiment 6) might use such a representational scheme, the evidence for infants (Experiment 5) is inconclusive.

Chapter 6 then explores infants' ability to evaluate expected information gain from their own perspective, with Experiments 7 and 8 focusing on whether and how 12-month-old infants can proactively assess the value of information before obtaining it. We designed a gaze-contingent paradigm to test infants' capacity to seek out advance information that could predict future events. The results suggested that while infants may evaluate information prospectively, previous information gain also influences their choices.

Finally, in Chapter 7, I provide a general discussion integrating the findings from all the chapters and discuss some avenues for future research on humans' interpretation of the epistemic dimensions of their conspecifics' behavior.

Table of Contents

"It is easier to judge the mind of a man
by his questions rather than his answers."

— Pierre Marc Gaston, Duke of Lévis

from *Maximes et réflexions sur différents sujets*
de morale et de politique (Paris, 1808)

Chapter 1. **Introduction**

In contemporary cognitive science, there is a widespread interest in cognitively driven information-gathering activities (recent representative reviews include Bonawitz, Bass, & Lapidow, 2019; Bromberg-Martin & Monosov, 2020; Coenen, Nelson, & Gureckis, 2019; Foushee, Srinivasan, & Xu, 2023; Gottlieb & Oudeyer, 2018; Hunt et al., 2021; Mehlhorn et al., 2015; Kidd & Hayden, 2015; Poli et al., 2024; Raz & Saxe, 2020; Ruggeri, 2022; Schulz & Gershman, 2019; Yang, Wolpert, & Lengyel, 2016). This interest partly stems from the appreciation that biological information processing is resource-limited (Lieder & Griffiths, 2019), while the environments in which it operates are complex and often transient. Organisms regularly face environments that are marked by significant variability in distal, directly inaccessible states, the properties of which are crucial for survival (e.g., depth of a pit, location of food, tomorrow's weather, intentions of a stranger), as well as by non-deterministic, noisy relationships between these distal states and the proximal cues that they can directly perceive (Godfrey-Smith, 1996; Sterelny, 2003). These conditions collectively give rise to subjective uncertainty.

Agents may mitigate uncertainty by learning internal models that represent the stochastic nature and recurring structures of the environment, enabling probabilistic predictions based on sparse and noisy data (Tenenbaum et al., 2011). However, since nature cannot provide all potentially relevant knowledge in advance, most of these models must be learned (Margolis & Laurence, 2011; Ullman & Tenenbaum, 2020). Additionally, the accuracy of these predictions is limited by the available data and the number of dimensions such models can efficiently integrate (Frankland, Webb, & Cohen, 2022). Relying on limited predictions is often too risky and could result in lost opportunities, especially in contexts that tend to give rise to novel problems.

Fortunately, through their movements, organisms can influence the flow of information they perceive, providing an opportunity to address these problems more promptly than waiting for incidental information to arrive. Through evolution, many species develop mechanisms with the proper function (Millikan, 1989) of generating behaviors with the outcome of perceiving environmental signals, which, in turn, help them tackle relevant uncertainties. To use George A. Miller's famous analogy, most animals are not just herbivores or carnivores; they are also *informavores*, whose mind "*survives by ingesting information*" (Miller, 1983, p. 111).[1]

Of all species, humans are the most proficient information-seekers. First, beyond engaging in sophisticated inquiry, we can also reason about this process, which, coupled with processes of cultural transmission and retention (Boyd, Richerson, & Henrich, 2011; Sperber & Hirschfeld, 2004), enables us to carve, perfect and share tools and practices for acquiring knowledge with exceptional efficiency. Second, as Tooby and Cosmides (2020) put it, we "*have taken the mining of information to be found in the minds of others to remarkable zoological extremes*" (p. 31), since language, in contrast to other communication systems found in nature, enables the exchange of an infinity of different contents, virtually allowing any question to be asked and answered, drastically reducing the cost of information.

From early on, human infants and toddlers, in addition to being receptive to pedagogy (Csibra & Gergely, 2011), spend most of their waking hours actively learning about the generic structure of their environment through sensory exploration (Fontenelle et al., 2007; Rochat, 1989), hypothesis testing (Lapidow & Walker, 2020ab; Schulz, 2012), social interrogation (Begus & Southgate, 2012; Begus, Gliga, & Southgate, 2014; Goupil, Romand-Monnier, & Kouider, 2016; Jones, Swaboda, & Ruggeri, 2020; Kovács et al., 2014), and focusing on events that promise learning opportunities (Kidd, Piantadosi, & Aslin, 2012, 2014; Stahl & Feigenson,

[1] In ecology, it is often found that animals with greater exploratory tendencies are more likely to encounter and adapt to diverse environments, thereby expanding their ecological niches. For instance, Sommer-Trembo et al. (2024) integrated behavioral data from 57 species of the cichlid fishes of Lake Tanganyika with ecological and genetic information. They discovered that exploratory behavior is linked to adaptations to different macrohabitats as well as to a genetic mutation affecting a region of habenula involved in movement and fear response.

It is not to say that exploration is without risks, as its utility also depends on ecological conditions that influence the diversity and stability of resources, as well as the costs associated with gathering information. Tebbich et al. (2009) demonstrated that exploration among Darwin's finches correlates with diet diversity and the amount of fruit in their diet, which is consistent with theories suggesting that generalist species benefit more from exploration than specialists. They also report that species relying on more stable and predictable resources, such as concealed food, exhibit lower levels of exploration probably due to reduced necessity for discovering new resources. Conditions such as food abundance, competition, and predation risk also influence the extent of exploration in different species (Greenberg & Mettke-Hofmann, 2001; Mettke-Hofmann, 2014; Winkler & Leisler, 1999).

2015). In fact, Gopnik (2020) recently argued that the uniquely long and expensive human childhood might have been selected asle it provided a protected period of life history in which humans could exercise their specialized exploration capacities, building a stable knowledge base that they could exploit later, in adulthood.[2] Therefore, it is no coincidence that decades after the seminal works of Piaget (1937/1952), Bruner (1973), McCall (1974), Gibson (1988), and others, understanding early information gathering remains a central aspect of developmental research, underscoring its role as a critical driver of cognitive development.

However, information seeking persists throughout the lifespan, arguably just as pervasively, albeit often in less salient forms. It remains crucial not only for efficiently building internal models of novel phenomena (Bramley et al., 2018; Braunlich & Love, 2022; Coenen, Rehder, & Gureckis, 2015; Steyvers et al., 2003), acquiring various skills (Lövdén, Gerzón, & Lindenberger, 2020), or making informed judgments and decisions (Good & Card, 1971; Quillien, 2023; Schulz et al., 2019), but also for performing even our most basic activities in the short term. Consider, for instance, the following actions during a bus ride: reading news on your phone, glancing at the doors to select the best exit as your stop approaches, and watching your steps on the stairs while getting off. While browsing the news is a culturally well-recognized form of information seeking, the subsequent actions are also information-directed in nature, even though they seem so automatic that we may rarely think of them as such. In fact, most instrumental tasks are synchronously accompanied by active perceptual monitoring of task-relevant contingencies (Jackendoff, 2007; Hayhoe & Rothkopf, 2011; Hoppe & Rothkopf, 2016, 2019).

This ongoing nature of information gathering has important implications for theories that aim to explain how humans interpret each other's behavior from the third-person perspective. Given their high frequency, *epistemic* actions—those intended to change the actor's own knowledge state through information gathering—should be considered just as

[2] The idea of the safe and protected human childhood is nuanced by Frankenhuis and Amir (2020), who argue that children historically faced higher levels of threat and deprivation than modern industrialized societies typically assume. They claim that human children have evolved to be highly adaptable to a wide range of adverse conditions, not just safe and nurturing environments. Thus, the childhood safeguarded and secure for discovery that Gopnik (2020) describes may be just one end of a spectrum, with varying degrees of exploration emerging across more or less risky environments (although this hypothesis is currently up to empirical testing).

crucial a component of human behavior as *instrumental* actions, which are aimed at changing the environment.[3] Both types of actions require interpretation from observers, yet research on social cognition tends to focus almost exclusively on how this interpretation is achieved in the latter case.

The present dissertation aims to contribute to understanding the core processes involved in interpreting epistemic actions by examining its rudimentary forms in early development. By studying infants—who have not yet been extensively shaped by language or culture—there is a potential to uncover some of the fundamental conceptual structures crucial for reasoning about information-directed intentions and behavior of our conspecifics.

1.1 Arguments for an early concept of epistemic action

To motivate the research presented in the following chapters, I should start by asking why the ability to reason about epistemic actions would develop so early, already in infancy. There are at least four independent lines of argument for this hypothesis. These are based on (1) infants' presumed predisposition to account for real patterns, (2) the logical necessity of epistemic action interpretation as a prerequisite for later skills, (3) the existence of direct learning benefits, and (4) the assumed benefits of the inferences that the causal interpretation of epistemic action could have provided in human evolutionary history.

First, frequent instances of epistemic actions constitute real patterns that infants can observe. Assuming that infants learn through constructing intuitive causal models to explain various experienced phenomena (Gopnik & Wellman, 2012), there is reason to believe that

[3] The term '*epistemic action*' was first put forward by Kirsh and Maglio (1994) to stand for a wide set of phenomena. According to their definition, any act counts as an epistemic act that has the function of reducing uncertainty or complexity in mental computations via revealing information that is either unavailable or hard to compute. In this dissertation, I will only focus on one prominent form of these behaviors, that is carried out to obtain samples from the environment.

In a general typology of behavior, apart from instrumental and epistemic actions, we can distinguish a third type, *communicative actions*, that are carried out to change the epistemic states of another individual through communication. There is now considerable work on the understanding of such communicative actions in infancy as well (e.g., Tauzin & Gergely, 2018; Vouloumanos, Onishi, & Pogue, 2012).

they will try to account for such recurring social stimuli as well — even though their initial models will likely be considerably more simplified and error-prone than those of adults. As will be discussed in more detail, there is extensive evidence that infants start to have principled expectations about others' instrumental goal-directed actions already in their first year of life (Gergely & Csibra, 2003; Liu et al., 2017). If such expectations are indeed driven by such rudimentary intuitive models, these models could soon develop towards encompassing epistemic actions as well. The exact mechanisms infants could use to generate and sample from the vast space of hypotheses compatible with their observations are currently unclear. Nevertheless, this framework has been successful in accounting for various experimental results and provides a fruitful general approach to conceptual development (for a recent review, see Ullman & Tenenbaum, 2020).

Second, one may argue that an intuitive model of epistemic actions must be acquired early, as other abilities that emerge later in childhood necessitate it. Particularly, during early language acquisition, the pragmatic capacity to understand interrogative speech acts as requests for information may be scaffolded by a prelinguistic understanding of epistemic behaviors. By 12 months of age, children can distinguish between declarative and interrogative sentences (Geffen & Mintz, 2015), and by the age of 2, they anticipate a shift in the speaker more often after interrogatives than declaratives (Casillas & Frank, 2017). However, beyond developing turn-taking expectations following interrogatives – that might also be driven by probabilistic association between these speech acts and speaker change – a concept of epistemic action is likely required to develop a full-fledged understanding of questions as tools for information request.[4]

[4] In a recent conference proceeding, Goodhue et al. (2023) reported an ongoing experiment exploring whether 18-month-olds know the canonical links between declarative and interrogative clauses and their corresponding speech acts. They conducted a preferential looking study, predicting that infants would look more at informed puppets during declaratives and at uninformed puppets during interrogatives, assuming that people look at speakers more than addressees during speech. The experiment featured two puppets in a cookie delivery scenario. During training, infants observed puppets collecting cookies delivered by a mechanical arm. In the pre-test phase, one puppet left the scene and missed the cookie delivery, creating an epistemic asymmetry between the puppets. In the test phase, the puppets conversed using either declarative or interrogative clauses while slightly moving. Preliminary results indicate that as predicted, participants tended to look more at the uninformed puppet in the interrogative condition and at the informed puppet in the declarative condition, indicating a preliminary

Third, grasping the role of information seeking in mitigating uncertainty toward achieving goals would benefit children through augmenting observational social learning. Uncertainty can stem not only from inexperience in specific domains but also from the inherently probabilistic aspects of the environment, giving rise to expected variability. Attending to such contingencies is critical for effectively navigating and responding to immediate situations. For example, a child might observe a parent checking the water's temperature before bath or watch an adult looking to the left and right before crossing a street. Recognizing these epistemic sub-goals as key to managing risks and uncertainties at hand could aid infants, who are novices concerning the background knowledge governing complex action sequences, in learning about the relevance of environmental cues in different contexts.

Finally, one could also speculate that the capacity to recognize others' efforts to obtain information offered significant benefits in navigating the social environment throughout human history and thus has been favored by evolution, possibly leading to the early emergence of this skill in ontogenesis. Many of these potential benefits arise from the causal relationship between these actions and the uncertainty of the agents performing them.

On the one hand, interpreting epistemic actions as stemming from uncertainty can help individuals identify the specific information needs of others and possibly tailor their own messages to address these needs.[5] This ability to provide relevant information can be advantageous in situations where being perceived as a cooperative partner is essential. According to theories of partner choice (Barclay, 2016; Barclay & Willer, 2007; Baumard, André, & Sperber, 2013), ancestral human environments were just like that: they provided

understanding of these linguistic mappings. However, a complete statistical evaluation of this difference awaits the completion of data collection from 21 more participants.

[5] In fact, the capacity of choosing relevant information to share is often seen as a precondition of linguistic communication. For instance, relevance theory (Sperber & Wilson, 2002) aims to provide a solution to the interpretational problem of communication by positing that hearers infer the intended meaning of an utterance by a relevance-based comprehension procedure: computing cognitive benefits via following the path of least effort and testing interpretative hypotheses in order of accessibility until the expectations of relevance are satisfied. The authors claim that this algorithmic procedure can be successful because speakers reliably communicate relevant information. Similarly in this aspect, the Rational Speech Act model of linguistic communication captures the speaker's need to be informative, by assuming that her utility scales with the social benefit of providing epistemic help to the listener (Goodman & Frank, 2016).

opportunities for mutualistic collaboration that could be endangered by free-riders, giving rise to supply and demand for true cooperators, constituting a 'biological market'. To navigate this market, humans evolved psychological capacities, allowing them to (1) scrutinize potential collaborators and abandon cheaters for new partners if necessary, but also to (2) manage their own reputation as high-value partners through displays of generosity and effectiveness. As part of the latter skill set, recognizing and responding to others' informational needs could have been crucial in environments where information sharing was essential for survival (Bickerton & Szathmáry, 2011; Tomasello et al., 2005).

Beyond demonstrating value as a cooperator through sharing information, recognizing both uncertainty and its resulting information gathering could directly support performing joint actions towards shared goals by minimizing aggregate costs and achieving co-efficiency (Török et al., 2019). For example, in cooperative foraging scenarios, recognizing when a partner is already checking a possible location could lead to better coordination and resource allocation, thereby enhancing the group's overall efficiency and success.

In addition, inferred uncertainty and subsequent information seeking may also serve as indicators of an individual's domain-specific knowledge and competence, which are highly beneficial to scrutinize when choosing partners to learn from or collaborate with (Pasquini et al., 2007; Sperber, 2010). Evaluating others' expertise based on their exploratory behavior can inform decisions about whom to trust and follow, enhancing both individual and collective outcomes.

Evidence shows that some of these capacities, potentially benefiting from recognizing epistemic actions, occur early in development. Studies have shown that 12-month-old infants adjust their gestures based on their partners' information needs (Liszkowski et al., 2006; Liszkowski, Carpenter, & Tomasello, 2008). Infants pointed more frequently to an object that an adult had not seen fall than to an object the adult had witnessed falling. To point in the former case, infants must have represented the adult in a state of epistemic uncertainty regarding the object's location. Tailoring information sharing based on listeners' goals and prior knowledge develops further in the preschool years. 4-year-olds share more details about a toy with someone who wants to understand its mechanics than with someone who simply wants to observe its effects (Gweon & Schulz, 2019), and they offer more general information to those with little knowledge and more detailed information to those who are more informed (Baer & Friedman, 2018). By the age of 5, children begin to alter the information they provide to optimize the benefits for the learner while reducing the effort needed for them to learn independently (Bridgers, Jara-Ettinger, & Gweon, 2020).

Similarly, experimental evidence suggests early sensitivity to competence cues, demonstrating that young children prefer to learn from and imitate individuals who exhibit certainty and expertise in a particular domain (Brosseau-Liard & Poulin-Dubois, 2014; Tenney et al., 2011). The capacity to critically assess information sources develops into the preschool years, distinguishing and reacting differently to teachers who provide too much or too little information (Gweon et al., 2014, 2018).

It is at least conceivable that evolutionary pressures that led to the early inclination to transmit relevant information and the early vigilance towards potential information sources also favored the early development of reasoning about others' information-directed behavior.[6]

1.2 Inductive challenges of recognizing epistemic goals

Even if infants would profit from understanding information seeking, it is still unclear how they could solve the inductive challenge of recognizing such actions. In fact, this question equally applies to adults.

Recent research reveals that when provided with a constrained set of possible knowledge-seeking goals behind an observed actor's behavior, adults can distinguish these

[6] While young children may not be seen as reliable informants or potential cooperators due to their inexperience, they could nevertheless benefit from the early practice of involvement in cooperative activities that require information sharing. In many non-WEIRD (Western, educated, industrialized, rich, and democratic) societies, children start to be integrated into adult activities from an early age, and their involvement becomes more structured and autonomous as they grow older. Lancy's (2018) ethnographic survey reveals that children in various cultures transition seamlessly from play to work. For instance, in the Chiga community of Uganda, children are encouraged to take the initiative from infancy, with their responsibilities gradually increasing as they age. By middle childhood, these tasks become integral to their daily lives, contributing to the family economy. In the Hadza society, children as young as five are observed fetching items, assisting with chores, and mimicking adult roles through play, which gradually evolves into more meaningful participation in household and community activities.

goals solely based on kinematic cues that characterize the performed actions (specifically, the differences in how one shakes a box depending on whether she wants to find out what kind of or how many objects are inside; Croom, Zhou, & Firestone, 2023). These results point to a sophisticated adult competence to reason about information-directed behaviors. However, they do not address how they recognize an action as information gathering in the first place. Although there may be many stereotypical forms of active information seeking, it is crucial to note that basically any action can serve epistemic purposes, hence, it is not trivial how people identify them.

In principle, all actions may result in perceptual access to states or events for the actors producing them. Yet, third-party observers need to correctly judge whether this access was the actual goal or just a by-product of an action (e.g., grasping a bottle to drink or to find out whether its content is cold). Moreover, there is further underdetermination in the relationship between an observed behavior and the specific underlying question driving it. For instance, a person checking a thermometer may be interested in the current temperature, or she may be looking for fingerprints. In other words, identical behaviors may serve the investigation of different questions. Therefore, an additional problem any observer must face is the specification of the question that the observed agent has in mind and aims to reduce its uncertainty about — that is, the goal of investigation.

Such inductive challenges, however, are not exclusive to information seeking. In fact, as it is often formulated, in principle, any finite set of goals is compatible with an infinite number of action means, and any stream of behavior is compatible with an infinite number of goals (e.g., Davidson, 1963). Therefore, a central aim of research on action interpretation is to understand how different individuals arrive at the one goal they eventually attribute to a given action.

As a solution, several accounts have introduced the idea that what ultimately constrains how we make sense of behavior is a set of expectations of rationality (e.g., Baker et al., 2017; Baldwin et al., 2008; Dennett, 1971; Gergely & Csibra, 2003). In infant cognition, the teleological stance theory of Gergely and Csibra (2003) became a particularly influential account. Following a line of looking-time studies (Csibra et al., 1999; Gergely et al., 1995; Csibra et al., 2003), these authors proposed that 9-month-old infants (but see evidence for 6-

month-olds in Csibra, 2008; Liu & Spelke, 2017) already possess an intuitive theory of rational action. This intuitive theory consists of two central tenets: (1) actions have a fundamental function of achieving certain goal states in the world, and (2) acting agents will consistently apply to the most efficient (least costly) means to reach these states, given the constraints of the particular situation. With these premises in mind, observers can deduce latent purposes behind actions by focusing only on those possible goals for which the observed behavior appears to be most efficiently suited. If environmental constraints justify the observed action means, then the action, the outcome, and the environmental constraints produce a well-formed teleological schema, and the observed outcome is attributed as the agent's goal. In addition, this theory is hypothesized to be productive to the degree that it allows the prospective inference (prediction) to goal states and retrospective inference to situational constraints as well (Csibra et al., 2003; Southgate & Csibra, 2008; Wagner & Carey, 2005). While the assumption of efficiency is a simplification, it arguably yields sufficiently valid inferences, as human behavior, generally speaking, tends to approximate the maximization of utility.

According to a related proposal, adults and infants are not only able to represent action costs but can also integrate these variables with agents' preferences as utilities and take them as the underlying causes of observed behavior (Jara-Ettinger et al., 2016; Liu et al., 2017). More specifically, the suggestion is that the intuitive theory that lies beneath human action understanding represents agents as utility maximizers, that is, rational planners who calculate the utilities of candidate actions (via subtracting costs from rewards) and select the one with the highest positive utility. With this naïve utility calculus at hand, observers can (1) predict agents' future actions based on the rewards and costs they assign to them, and also (2) infer backward the costs and rewards from observed behavior.

In a similar vein, having a simplified, intuitive theory that accounts for fundamental principles of human information seeking could offer observers comparable inductive constraints for inferring intended epistemic outcomes. Here, I propose three such regularities.

First, individuals tend to gather information about variables that may impact their ability to meet their instrumental goals (e.g., Hoppe & Rothkopf, 2016, 2019; Schulz et al., 2019). There are multiple aspects of the environment that could make a difference in individuals' expected payoff when aiming to achieve their goals. Individuals are often uncertain about one or more of these aspects. Learning about these instrumental, goal-relevant factors usually leads to more efficient goal achievement. Infants could learn and integrate this pattern of behavior into action-specific expectations about others and assume that partial knowledge about a goal-relevant state is a valid reason to inquire about it.

Second, humans also gather information about states that are novel to them. That is, the dynamics of human inquiry may extend beyond resolving task-specific uncertainties, towards information that is not directly instrumental. Environmental changes, especially those involving novel or unexpected elements, naturally elicit curiosity and exploratory actions (e.g., Kidd, Piantadosi, & Aslin, 2012; Loewenstein, 1994; Stahl & Feigenson, 2015). This established pattern of detecting something new and subsequently exploring it could also be recognized by infants when discerning actions aimed at information acquisition.

Finally, when interpreting their conspecifics' behavior, it is helpful for individuals to presuppose that others engage in information-seeking behavior efficiently by trading off the costs of their actions and the expected value of information acquired. Rational models of inquiry (Chater et al., 1998; Coenen et al., 2018; Nelson, 2008) aim to provide a normative solution to how this should be done. They propose that given a choice between a set of mutually exclusive lines of investigation, a rational agent should choose the option with the highest expected utility. In line with standard decision theory, the suggestion is that these expected utilities are derived as the probability-weighted average utility of the possible outcomes of the given investigation. However, in contrast to instrumental acts, where intended outcomes are material, information gathering, as the very notion implies, ends in having information. If the value of any inquiry is determined by the value of the possible pieces of information it might provide, the problem ultimately boils down to the issue of what determines the value of information, or, from a normative point of view, how information should be valued. In this respect, Chater et al. (1998) make an instructive distinction between *interested* and *disinterested* sampling and their corresponding utility functions.

Interested sampling is the kind of information gathering that is at work when an agent has a higher-order instrumental decision problem and it tries to optimize its outcome. In this case, the gathered information is valuable if it changes the agent's expectations regarding the best option within this decision problem. For example, if one is faced with deciding whether to pack an umbrella for a trip, watching the forecast is valuable to the extent that it helps to make this decision.

Disinterested sampling, on the other hand, is not attached to any instrumental goal or task: the agent seeks information even though it lacks any specific knowledge about its potential material or strategic benefits. Instead, its goal is to identify the correct hypothesis in a given hypothesis space by maximizing information gain.[7] Here, information has value to the extent that it brings the agent closer to this desired epistemic state of minimal uncertainty.

Normative solutions do not fully align with the reality of human exploratory behavior that is often constrained by cognitive limitations and situational contexts, leading individuals to rely on various heuristics that provide outcomes that are only satisfactory but not perfectly optimal.[8] However, from the perspective of action interpretation, the presumption of

---

[7] This epistemic value of information has been formalized in several ways apart from information gain (Nelson, 2008; Nelson, Meder, & Jones, 2018; Coenen, Nelson, & Gureckis, 2019), including probability gain (the extent of increase in the probability of a correct guess) and impact (the extent of absolute change from prior to posterior representations).

[8] The computations required to arrive at optimal solutions can become quite demanding. For simplicity, let us consider disinterested sampling, where changes in expected utilities should not even be taken into account, only information gain. Even in this case, an ideal agent should not only consider which Y variable each possible sampling action could disambiguate at time $t+1$, and how the observation of this variable would update her probabilistic representation about a target variable X, but also how this fictious posterior representation, $P(X|Y_{t+1})$ would differ, in terms of entropy, from their current representation of $P(X)$. This is due to the normative consideration that the amount of information that an action could provide depends on the average information gain over each of the possible posterior updates that it could lead to. Now, the information gained by only one such possible update can be quantified via subtracting the expected entropy (the statistical equivalent of uncertainty) of $P(X|Y_{t+1})$, from the entropy of $P(X)$, as represented in the present. At the end of the day, a rational agent should, in principle, average the expected reduction in the entropy of $P(X)$ over all possible updates, for all possible disambiguated variables, yielded by all the possible actions it can take.

Normative approaches often provide computationally intractable functions of this kind, therefore a frequent theoretical scheme in such cases is the formulation of different, tractable functions, generating only approximately optimal output (van Rooij et al., 2019). In our case, one way out of this conundrum is granted by the fact that expected information gain can be expressed in a different, but mathematically equivalent formula, which can be interpreted as the difference between the uncertainty about the variables an action could disambiguate and the average uncertainty we would have about these variables if we already knew what the true value of the target variable X was. A slightly different heuristic simply takes the current uncertainty about Y_{t+1} as a proxy for the expected information gain of an action, however, this approach, called uncertainty (or maximum entropy) efficiency serves as a functional simplification, facilitating prediction and interpretation by abstracting away from non-essential complexities of behavior. Notably and congruent with this theoretical framework, Aboody, Zhou, and Jara-Ettinger (2021) have shown that preschoolers are ready to infer others' epistemic states and corresponding valuation of information based on the cost they are willing to pay for acquiring it.

Importantly, to expect agents to gather information efficiently, infants themselves should already have the resources to evaluate information. However, the developmental origins of humans' ability to proactively evaluate and search for relevant information are still unclear. There is a growing body of research on how infants' observational behavior may be sensitive to the information they gained in the past (Ghilardi et al., 2024; Kidd, Piantadosi, & Aslin, 2012, 2014; Poli et al., 2020); however, in these studies, participants have little role in actively selecting the target of their investigation in the first place — instead, the active component consists in paying or not paying further attention to already encountered stimuli. Selection in infant exploration is either left unexplained or presumed to be either random (e.g., Altmann, Bazhydai, & Westermann, 2021) or habitual, that is, based on learned, backward-looking associations between internal states of uncertainty and action routines, reinforced by previously experienced information gain (Carruthers, 2018, 2019; Perner, 2012). Importantly, these assumptions are not based on actual evidence, as to my knowledge virtually no study attempted to target the evaluation of expected information gain in infancy directly.

In conclusion, to solve the inductive challenges of recognizing epistemic goals, human infants from early on may utilize an intuitive model that represents at least three reliable properties of epistemic actions as guiding principles for interpretation: instrumental relevance, interest in novelty, and efficiency. The first part of this dissertation (Chapters 2–4) focuses on testing this hypothesis. In addition, since reasoning about the efficiency of others' epistemic behavior necessitates cognitive resources to evaluate information from one's own perspective, Chapter 6 focuses on testing this capacity in infants.

sampling, only identifies optimal strategies in case of deterministic likelihoods, that is, when knowledge of the variable of interest *X* would make it possible to predict Y_{t+1} with complete certainty, which is rarely the case (Yang, Wolpert, & Lengyel, 2016; Nelson, Meder, & Jones, 2018). For other solutions, including the likelihood difference and the split-half heuristic, see Nelson, Meder, & Jones (2018).

1.3 Attribution of certain and uncertain mental states

Epistemic actions differ from instrumental actions in that not only their causes but also their intended outcomes are unobservable epistemic states. Therefore, as an additional complication, to construct and utilize the intuitive theory hypothesized in the previous section, infants should already be able to represent and attribute epistemic states to others.

There is evidence both for and against the existence of Theory of Mind (ToM; the ability to reason about mental states such as beliefs, intentions, desires, and knowledge) in infancy. Studies using the violation-of-expectation paradigm (e.g., Onishi & Baillargeon, 2005) demonstrated that 15-month-old infants expected an agent to act based on her false beliefs, suggesting an early understanding of mental states. Interactive paradigms, where infants assist an adult taking into consideration the adult's false belief, further support the presence of early ToM abilities (Buttelmann, Carpenter, and Tomasello. 2009), besides numerous other methods that have underscored infants' sensitivity to others' visual and mental perspectives (Baillargeon, Scott, & He, 2010; Buttelmann & Kovács, 2019; Kampis et al., 2015; Kampis & Kovács, 2021; Kovács, Téglás, & Csibra, 2021; Kovács, Téglás, & Endress, 2010; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007).

However, replication efforts of some of these findings have produced mixed results, raising questions about the reliability of the initial findings (Kulke et al., 2018). Besides direct and conceptual replications of some paradigms (Buttelmann, Suhrke, & Buttelmann, 2015; Király et al., 2018; Schneider et al., 2012; Thoermer et al., 2012), recent studies have failed to reproduce some of the original results (Barone, Corradi, & Gomila, 2019; Crivello & Poulin-Dubois, 2018; Kampis et al., 2021; Kaltefleiter et al., 2022; Kulke et al., 2018; Poulin-Dubois et al., 2018). This ongoing debate about the replicability of these findings has led to multi-lab collaborations conducting large-scale studies to better understand infants' performance on different ToM tasks (Schuwerk et al., 2021).

Importantly, to my knowledge, no existing study investigated infants' belief attributions in the context of epistemic actions, and, in particular, whether infants can correctly interpret others' actions as directed towards information. Instead, they typically examine scenarios involving agents who act or are expected to act to cause changes in their environment based on information they have previously acquired. Moreover, since the signature of the full-fledged ToM is the ability to attribute beliefs to others that conflict with one's own understanding of reality, these tasks tend to test whether infants can attribute false or inaccurate beliefs to others and whether they understand the consequences of being ignorant about a crucial aspect of a situation. False beliefs are beliefs with definite, well-defined contents conflicting with the actual state of affairs, while ignorance, in the bulk of these studies, is simply the lack of any beliefs about some condition. However, information-gathering agents are usually driven by a different kind of ignorance, namely, *uncertainty*.

Information, at least according to the most prominent account, is the reduction of previous uncertainty about a set of possibilities (Shannon, 1948).[9] Implied by this definition, the information that the very same state or event provides is not an intrinsic property of it but is always relative to the prior space of possibilities that the information-gathering agent entertains. Similarly, it is argued that the kind of mental states we attribute to agents in inquiry are better described as having *questions* as contents rather than being attitudes towards a specific proposition as beliefs are (Friedman, 2013; Carruthers, 2018). Questions do not have truth conditions, but classical semantic analyses (Hamblin, 1958; Karttunen, 1977) identify their content with the set of possible answers that could satisfy them. For example, the content of the question that we could express linguistically as "*Where is the ball?*", is the set {the ball is at x, the ball is at y, …}. Thus, one way to conceptualize the uncertainty of other agents would be to represent them as having attitudes toward a set of possibilities.[10]

Accordingly, such a modal understanding would involve quantification over mutually exclusive possibilities that the observed agent entertains. There is some evidence that preverbal infants can represent alternatives disjunctively and make inferences by exclusion. For instance, 12-month-olds looked more at events that were inconsistent with the conclusion of disjunctive inferences (Cesana-Arlotti et al., 2018), 14-month-olds' pupil size was shown to

[9] Semantic theories of information differ from this formulation in that they usually neglect the probabilistic aspect of the hypothesis space. In such frameworks, information is what reduces the range of logical possibilities (e.g., Bar-Hillel & Carnap, 1952; Ciardelli et al., 2019). More formally, the information content of a proposition P is what updates a set of possible worlds W to {w in W | w makes P true} (Adriaans & van Benthem, 2008).

[10] Such a representation of inquiring agents' mental states would come with the assumption that these agents are always attributed possible answers to their questions, or at least they entertain a single possibility and its negation.

increase with the number of alternatives they should have sustained in memory (Cesana-Arlotti, Varga, & Téglás, 2022). Moreover, infants have been found to be sensitive to the likelihood of events, possessing early capabilities to manage uncertainty across various possibilities (Denison & Xu, 2010ab; Lawson & Rakison, 2013; Téglás et al., 2007, 2011, 2015, 2016).

In contrast, even preschoolers seem to struggle to consider possibilities in certain tasks that involve planning actions. In a study by Beck et al. (2006), three- and four-year-olds were instructed to place mats to collect items released from an inverted Y-shaped slide. Instead of placing two mats, they placed only one mat half the time, resulting in unsafe landings. Similarly, Redshaw and Suddendorf (2016) used the same kind of tube and asked participants to catch a ball released into the tube. Most four-year-olds used two hands, one under each opening, ensuring a successful catch. In contrast, younger children frequently covered only one opening, overlooking the possibility that the ball could exit from either side. Various theoretical accounts were put forward to resolve these conflicting findings, but the debate on early modal cognition is far from being settled (Bohus et al., 2023; Feiman, Mody, & Carey, 2022; Leahy & Carey, 2020; Turan-Küçük & Kibbe, 2024).

In any case, even if infants themselves can indeed represent possibilities, it is unknown whether (and when) they can also ascribe content with disjunct constituents to other agents – although one might argue that any representation that a subject can entertain could also serve as a content of the representation it attributes to others (Sperber, 2000).

On the other hand, infants may use a different format to represent others' uncertainty, which does not require the cognitive resources to think about possibilities. Discussing the content of questions, Friedman (2013) suggests that representing a set of answers requires only that the subject can represent what the answers have in common: an *open proposition*, one or more arguments of which are empty (e.g., *"The ball is at SOMEWHERE."*). Similarly, infants might represent uncertain epistemic states as mental states with open-ended contents, that is, open propositions that have an argument position that is occupied by a placeholder, without specifying the range of its possible values (in this example, the possibilities that the agent entertains about the actual location of the ball). Such an attribution could indicate to the ToM system that the agent has only partial knowledge without the involvement of modal concepts.

However, the inferential productivity of this open-ended representational scheme would be restricted. This is because, in relation to an open proposition, where there are no specified possibilities to which probabilities could be assigned, the value of information becomes underdetermined and cannot be used for definite prediction. For example, suppose you observe an agent trying to find a ball hidden in one of three boxes. Also, suppose that you managed to attribute her the goal of getting the ball as well as an open-ended uncertainty concerning the ball's location, meaning that she does not know in which box the ball is. If you have no idea which boxes she considers more or less likely hiding places out of the three, you have no principled way of predicting where she will search for it. In contrast, if you use the modal representational scheme, you can, for instance, attribute her the more specific uncertainty that the ball is in one of the three boxes with high probability. Thus, you can make the specific prediction that she will look for it in that box to be efficient. Alternatively, if you rely on logic instead of statistical probabilities, you may infer that she realized, for some reason, that the ball cannot be in two of the three boxes, and predict that by exclusion, she will look for it in the third box.

Nevertheless, epistemic goal ascription may also be possible in the open-ended scheme, at least retrospectively. That is, imagine you have ascribed the agent the instrumental goal of getting the ball, and you assume that (1) a crucial enabling condition of getting the ball is knowing where it is, and (2) you also infer that the agent cannot presently see where the ball is. Then, based on these premises, you may infer that it must be uncertain about its location. Still, given your representational constraints, you cannot attribute modal uncertainty over specific possible locations, just a general one. Then, given this attributed general uncertainty and the intuitive theory detailed in the previous section, upon observing the agent positioning itself to a location where it can check some potential hiding places, you may justify her behavior as a well-formed epistemic action targeting the ball's location given her previously inferred uncertainty. However, in this limited scheme, you cannot utilize the additional inductive constraint of efficiency since you cannot know what exact possibilities the agent entertains, and the value of information upon which efficiency depends can only be defined in relation to such possibilities.

So far, I discussed the attributional prerequisites of understanding uncertainty as a *cause* of epistemic actions. However, to recognize the intended *outcome* of epistemic actions, the causal understanding that seeing leads to knowing is also essential to grasp, otherwise alternative justifications of the observed behavior would be generated (e.g., that the agent's goal was only the establishment of the physical relation between its eyes and some object). Existing studies indicate that infants are sensitive to the gaze direction of others. They start to follow gaze spontaneously between 3 and 6 months of age, although this behavior is not precise in the very beginning when infants often simply look at the closest object in the followed direction (Del Bianco et al., 2018). In the presence of communicative cues, 8-month-olds have referential expectations when they follow gaze, showing signs of surprise when an expected referent is absent (Csibra & Volein, 2008). Finally, and most importantly, within their first year, infants also represent that what agents can see limits their possible goals and actions, as well as their epistemic states (Luo & Baillargeon, 2007; Luo & Johnson, 2008; Surian, Caldi, & Sperber, 2007), indicating a rich concept of seeing, the primary medium of active perception.

To sum up, I argued that epistemic actions are primarily driven by uncertainty rather than false/true beliefs or simple ignorance, and their intended outcomes are informed epistemic states. While there is evidence that infants understand that seeing leads to knowing, potentially allowing the recognition of epistemic outcomes, it is currently unknown whether and how infants conceptualize and attribute uncertainty, the cause of epistemic actions. I argued that infants could use either modal or open-ended representational schemes to reason about uncertain agents. Despite infants' demonstrated – though debated – ability to handle modal representations and manage uncertainties from their own point of view, the question of whether they can ascribe such modal contents to others is still open. Chapter 5 tries to address this issue.

1.4 Overview

The previous sections showcased some of the outstanding challenges of reasoning about epistemic behaviors in infancy and outlined the corresponding questions investigated in this dissertation. These will be addressed as follows.

Focusing on the hypothesized contents of infants' intuitive theory about epistemic actions, in Chapters 2–4, I present experiments targeting 14-month-olds' representation of the three proposed principles driving behaviors: instrumental relevance, interest in novelty, and efficiency. Chapter 2 concentrates on instrumental relevance, with two experiments testing whether infants recognize epistemic actions produced to disambiguate the optimal course of action to accomplish an instrumental goal. Chapter 3 examines infants' expectations about others' drive toward exploring novelty. Chapter 4 investigates whether infants expect agents to gather information in an efficient manner, balancing the cost of the action with the potential informational gain.

In Chapter 5, I discuss the possible representational structures underlying the attribution of uncertainty, and test whether infants and toddlers apply a modal representational scheme when interpreting and predicting uncertain agents' behavior. Chapter 6 evaluates how infants themselves assess expected information gain, considering whether they can predict the value of information before it is obtained. Finally, Chapter 7 provides a general discussion, integrating the findings from all chapters, and suggests future directions for research.

Chapter 2. **Principle I: Instrumental relevance**

One key role of information gathering is to change expectations regarding the optimal course of action to accomplish a goal. Since the optimal action is usually contingent upon a set of latent variables, reducing uncertainty about these variables is often a sub-goal of the main aim of behavior (e.g., Hoppe & Rothkopf, 2016, 2019; Schulz et al., 2019). As discussed in the previous chapter, representing this functional role of information-seeking actions might aid infants in recognizing them. Accordingly, in the two looking-time experiments presented in this chapter, we asked whether 14-month-olds attribute uncertainty to agents based on these agents' limited visual access to goal-relevant variables and use this inferred mental state to attribute information-seeking goals to these agents' subsequent actions.

2.1 **Experiment 1**

In Experiment 1, infants were presented with an action that was inefficient with respect to the observed agent's immediate instrumental goal of reaching an object. Specifically, instead of approaching one of two possible hiding locations of the object, the agent always moved to a distal point that afforded a view of these locations. However, in one (EFFICIENT EPISTEMIC ACTION) condition, the agent was uncertain regarding the object's true location; therefore, this action could be justified as a well-formed information-seeking behavior to locate it. In contrast, in the other (INEFFICIENT EPISTEMIC ACTION) condition, the same action could not be justified as the agent already had visual access to the object's location. We expected longer looks at the INEFFICIENT EPISTEMIC ACTION condition compared to the EFFICIENT EPISTEMIC ACTION condition.

The experiment consisted of two familiarization phases followed by two test trials (see Figure 1 for an overview). The first familiarization served to reveal the agent's goal, while the second was used to acquaint infants with the relevant perspectives determining the agent's visual access in the test phase. In the first familiarization phase, we demonstrated six times that a red, round-shaped agent prefers to selectively approach a target object (a duck), independent of its location, in contrast to a different, blue object. Based on previous findings (e.g., Woodward, 1998), we expected infants to infer, given these events, that the agent had the goal or preferred to approach the duck. Notably, the two objects had identical upper parts that played a role later in the test phase.

In the second familiarization phase, we showed two different situations, two times each. These phases aimed to familiarize infants with the various visual perspectives that can be taken in a novel environment from different positions at different levels. This environment (also used in the test phase) had two levels: on the upper level, there was a house and two tubes linking the two levels and ending in two boxes at the lower level. In addition, there was also a ladder connecting the upper level to a platform situated at the lower level from where the two locations could be visually inspected.

In one familiarization scenario, the agent came out of the house, watched the target object jumping into one of the two boxes through the respective tube, and then followed it by jumping into the same tube. We assumed that infants would interpret these events as the agent seeing the target's movement and subsequently approaching it, consistent with its previously shown preference in the first familiarization phase.

In the other situation, the agent was standing on a platform below, facing away from the tubes and boxes, while the non-target object appeared behind it on the first level and jumped to one of the boxes with a bumping sound. The agent turned around, looked at the non-target object in the box, then turned its back to it without performing any further action. This scenario was added to let infants recognize that the platform provides a perspective (also used in the test phase) for the agent to see the boxes' contents and check whether the target object is present (if not, no action is performed).

Finally, infants were shown one test trial per condition, in counterbalanced order, that differed from familiarization and also from each other in one critical aspect. In the INEFFICIENT EPISTEMIC ACTION condition, when the agent came out of the house, it had full visual access to both objects before these simultaneously jumped into the boxes, one in each box. We assumed this perspective could enable infants to attribute certainty to the agent regarding the target object's location.

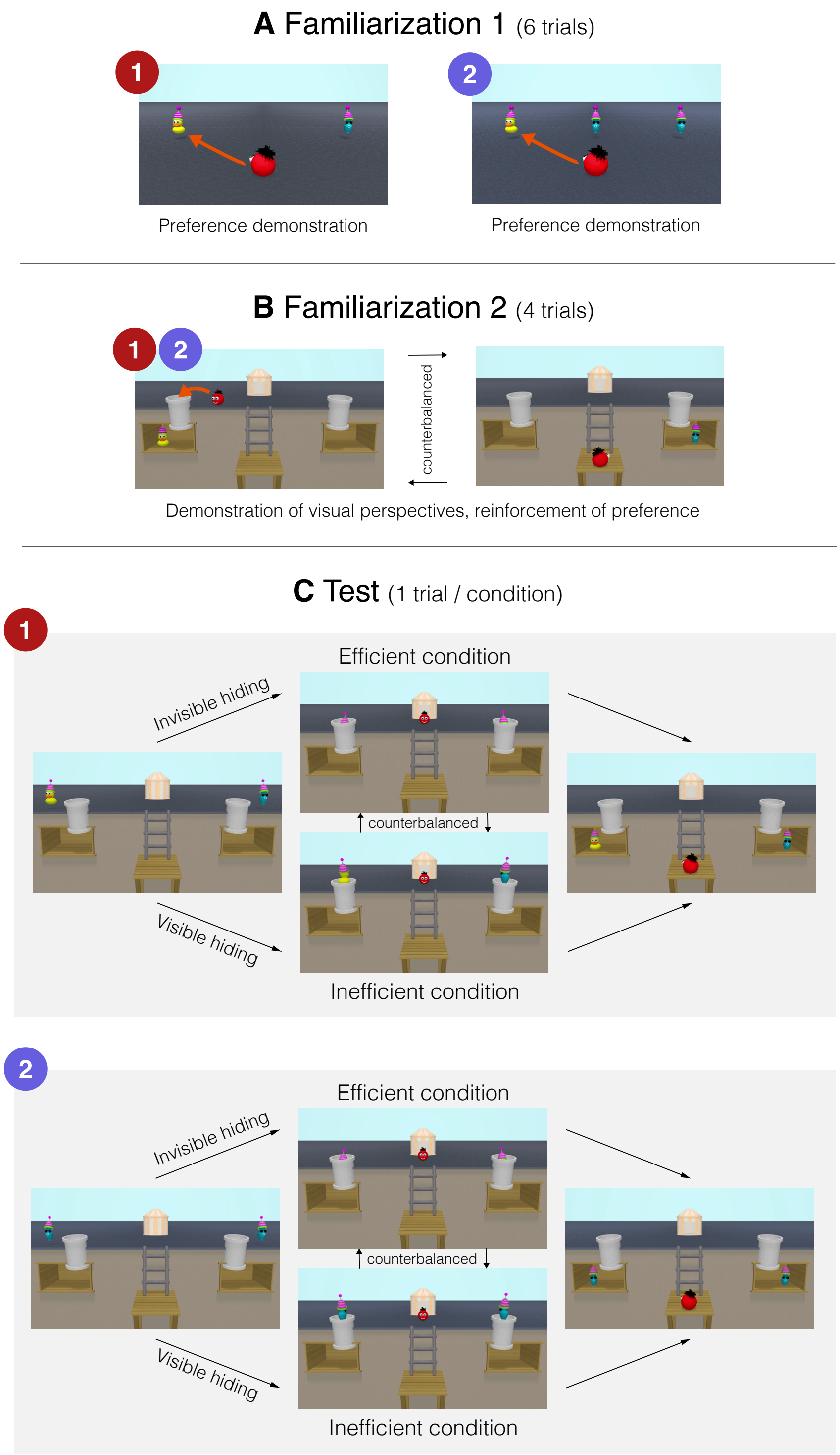

Figure 1. Structure of Experiments 1 to 2. Both experiments follow a three-phase structure: (A) *Familiarization 1*, (B) *Familiarization 2*, and (C) *Test*. The experiments are differentiated by color coding. **Experiment 1 (red)** focuses on infants' recognition of agents' information-seeking actions under instrumental uncertainty. In *Familiarization 1*, the agent shows preference by choosing a target object over a non-target independent of its location. *Familiarization 2* reinforces this preference and introduces the test environment and the agent's two main perspectives. The *Test* phase contrasts EFFICIENT (top path; objects are partially hidden from the agent, justifying information seeking) versus INEFFICIENT EPISTEMIC ACTIONS (bottom path; unnecessary detour as the target object's location is already known). **Experiment 2 (purple)** replicates the structure of Experiment 1 but introduces a control by adding an identical non-target object to *Familiarization 1* to test the specificity of infants' responses to the agent's information-seeking behavior. *Test* conditions are again divided into EFFICIENT and INEFFICIENT ACTIONS depending on the visibility of hiding the objects from the agent's perspective; however, the hidden objects are of the same kind.

Then, the agent took a detour: it ascended the ladder to the lower platform, taking the position where it had a clear view of which object was in which box. This detour, however, was hard to justify in this condition as the agent already had access to this information. In the other, EFFICIENT EPISTEMIC ACTION condition, when the agent came out of the house, only the identical, indiscernible parts of the objects were visible, after which the objects jumped to the boxes. That is, the location of the goal object fell within a range of possibilities from the perspective of the observed agent (location A or B) but not from the perspective of the infant (since before the agent exited the house, the infant could see which object is at which location, and had direct visual access to the objects on the lower level). This state of limited information made the two mutually exclusive instrumental actions available to the agent (go to location A or B) to have exactly the opposite utilities in the two equally probable states (target in A or target in B), rendering the average expected utilities ultimately equal for both possible actions. That is, given the agent's perspective, participants could attribute uncertainty to the agent regarding the target object's location. Then, again, the agent made the very same detour as in the INEFFICIENT EPISTEMIC ACTION condition; however, it was now justifiable by the agent's uncertainty.

We paused the animation at the end of the detour in both test conditions and measured infants' looking time. We expected longer looks in the INEFFICIENT EPISTEMIC ACTION condition, where the agent's uncertainty about the target object's location could not justify the detour.

2.1.1 PARTICIPANTS

24 full-term, typically developing 14-month-old infants were included in the analysis of the experiment (11 females, M_{age} = 14 months, 18.2 days, SD_{age} = 7.5 days). Additionally, 24 infants were excluded: 17 due to crying or fussiness, 5 due to experimental error, 2 due to equipment failure. Note that this experiment was run during the COVID-19 pandemic; therefore, we assume the situation might have been more stressful for parents, infants, and researchers, likely leading to high exclusion rates (given the safety measures, the masks, and the generic stress experienced). All infants were recruited in Budapest based on birth records. At the end of the experiment, infants received a small gift for their participation.

The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB), Budapest, Hungary, reference: 2019/13, and conducted in accordance with the principles defined in the Declaration of Helsinki. Prior to their participation in the research, all caregivers provided written informed consent.

2.1.2 Apparatus and stimuli

Animations were created and rendered in Maya®, a 3D computer animation software (Autodesk Inc.). The resulting stimuli were displayed on the inbuilt monitor of a Tobii T60 XL eye tracker device (51.1 x 32.3 cm, 1920 x 1200 px), controlled by PyHab (Kominsky, 2019), a stimulus presentation system for PsychoPy (Peirce et al., 2019). The stimuli, experiment control scripts, and data files are available at https://osf.io/uf2bn/.

2.1.3 Design and Procedure

The experiment took place in a soundproof room with dimmed lights. Participants were seated on their caregiver's lap, about 60 cm from the display. Caregivers wore opaque glasses that prevented them from seeing the stimuli. They were instructed to keep the child seated and not to interact with them. The experimenter was sitting behind a curtain, controlling the experiment and monitoring infants' behavior via a video camera. Infants watched two familiarization phases and a test phase.

In the first familiarization phase, infants saw six 8-second familiarization trials, the aim of which was to convey that the agent had the goal of approaching a specific object. In these trials, the agent approached the target object over a distractor, independent of its location. First, the agent, standing in the center of the scene, faced the infant, jumped twice to attract attention, then turned around and approached the target. Both the target and the distractor had identical upper parts. The target's location alternated across trials in a fixed order for all subjects (target on the Left or Right: LLRRRL).

In the second familiarization phase, infants viewed two 12-second trials. These trials aimed to familiarize infants with a new environment and the perspectives that the agent can take within it. The trials featured a two-level environment. On the upper level, there was a house and two tubes, each ending in a box on the lower level. Additionally, a ladder linked the upper level to a platform on the lower level, at the same elevation as the boxes, providing a vantage point for visual inspection of both boxes.

In trial type A, trials started with the target initially appearing in the upper level next to a tube (left in the first trial, right in the second). Then, the agent emerged from a house in the center, and they turned to face each other. Subsequently, the target's upper part (antenna) lit up twice (attracting the agent's attention), and it jumped to the tube, arriving in a box on the lower level. Then, the agent jumped into the same tube and reached the target in the respective box. In trial type B, the agent stood on the platform on the lower level, facing the infant. Then, the distractor object appeared next to a tube on the top level (right in the first trial, left in the second) and then jumped into it, emitting a sound and arriving at the corresponding box under the tube. Then, the agent turned to look at the distractor object in the box, then turned back without further action. Each trial was presented twice, in ABAB or ABBA order, counterbalanced between subjects.

In the test phase, infants were presented with one trial per condition (EFFICIENT and INEFFICIENT EPISTEMIC ACTION) that featured the same environment as the second familiarization phase. The trials were the same length (26 secs) and differed depending on condition. Initially, in both conditions, both the target and distractor objects were presented next to one of the tubes (target left and distractor right, or target right and distractor left). Then, they moved above (EFFICIENT condition) or partially inside (INEFFICIENT condition) their respective tubes. Finally, the objects submerged in (EFFICIENT condition) or emerged from (INEFFICIENT condition) the tubes; in the latter case, they stopped above them. In both conditions, the objects were seen in these two positions (above or within the tube) for the same amount of time; only the order of these positions differed.

When the objects moved to their second position, the agent came out of the house and looked toward the tubes, seeing only the identical upper parts (EFFICIENT condition) or the full objects (INEFFICIENT condition). Finally, the objects moved down the tubes to their respective boxes, and the agent moved down the ladder to the platform and looked at the distractor and then at the target. The presentation order of the two test conditions and the object locations were counterbalanced across subjects.

At the test trials, the last frame remained on, and looking times were recorded until the participants looked away for 2 consecutive seconds or looked at the screen for a total of 30 seconds. Attention-getters, depicting colorful geometrical figures on a gray background, were inserted between each trial, except in the first familiarization phase, where they were inserted after every second trial. The sessions were video-recorded. To extract exact looking time durations for data analysis, looking behavior was coded offline frame-by-frame.

2.1.4 RESULTS

Our main dependent measure for looking behavior was log-transformed looking time (based on Csibra et al., 2016), but descriptive statistics and plots feature raw values for ease of interpretation (see Figure 2). On average, infants looked longer in the INEFFICIENT EPISTEMIC ACTION condition (M = 18.22 s, SD = 9.57 s) compared to the EFFICIENT EPISTEMIC ACTION condition (M = 13.10 s, SD = 9.95 s). We conducted a mixed ANOVA to examine the effects of condition and condition presentation order on cumulative looking time. The analysis revealed a significant effect of condition ($F[1,22] = 9.926$, $p = .005$, $\eta^2_P = .311$) but not of presentation order ($F[1,22] = .016$, $p = .902$, $\eta^2_P = .0007$), nor an interaction ($F[1,22] = .452$, $p = .508$, $\eta^2_P = .020$). Thus, as predicted, infants looked significantly longer in the INEFFICIENT EPISTEMIC ACTION condition.

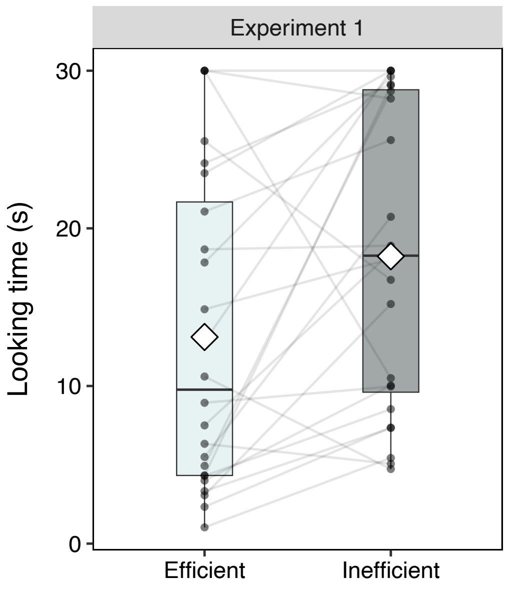

Figure 2. Boxplots of raw looking times (in seconds) for test trials in Experiment 1 (N = 24). White diamonds indicate means, bold horizontal lines indicate medians, boxes indicate interquartile ranges, and linked black dots indicate individual data points across conditions.

Although these results indicate 14-month-olds' ability to attribute information-seeking goals to uncertain agents, there may be an alternative interpretation of these looking time findings. Since the agent had initial knowledge of the target object's location in an INEFFICIENT EPISTEMIC ACTION condition, it is possible that infants had a specific expectation of a direct approach toward the target. This may yield relatively longer looking times if infants did not understand the scene in terms of information-seeking actions and formed no expectations in the EFFICIENT EPISTEMIC ACTION condition.

2.2 **Experiment 2**

To rule out the above alternative interpretation, specifically that infants' gaze was simply guided by expectations formed in the INEFFICIENT EPISTEMIC ACTION condition without understanding the information-seeking nature of the detour in the EFFICIENT EPISTEMIC ACTION condition, we conducted Experiment 2. We aimed to provide additional proof that infants understand information seeking while excluding the possible alternative explanation of Experiment 1.

Experiment 2 also consisted of two familiarization phases followed by two test trials. The main design was the same as in Experiment 1, except that we introduced one additional non-target object in the familiarization and test phases, identical to the non-target object in the first experiment. Specifically, in familiarization, we first demonstrated six times that the agent prefers to selectively approach the duck, independent of its location, in contrast to two other identical blue objects. The second phase of familiarization was the same as in Experiment 1, while the test trials only differed from the first experiment in terms of the identity of the objects: now only the two non-target objects appeared in both test conditions, hiding visibly or invisibly. Importantly, the top parts of all three objects were identical (see Figure 4).

The reasoning behind this manipulation was the following. In the condition where the two non-target objects are visible from the agent's perspective, infants can no longer expect the agent to directly approach either of them, as none of them is the target, unlike in Experiment 1. At the same time, the agent's subsequent detour in this condition remains inefficient: it cannot be justified as a well-formed epistemic action, as the agent already had information regarding the identity of the objects at both locations.

In contrast, in the condition with invisible hiding, the detour remains an efficient epistemic action, as, from the agent's point of view, the target could still be in one of the boxes. Therefore, under the hypothesis that infants interpret events in terms of epistemic actions in our paradigm, we predicted them to look longer in the former, INEFFICIENT EPISTEMIC ACTION condition compared to the EFFICIENT EPISTEMIC ACTION condition. Under the alternative hypothesis described above, however, such a difference could not be expected.

2.2.1 PARTICIPANTS

24 full-term, typically developing 14-month-old infants were included in the analysis of the experiment (13 females, M_{age} = 14 months, 15.37 days, SD_{age} = 7.68 days). Additionally, 6 infants were tested but excluded from the analysis: 3 due to crying or fussiness, 2 due to caregiver interaction, and 1 due to equipment failure. All infants were recruited in Budapest based on birth records. At the end of the experiment, infants received a small gift for their participation.

The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB), Budapest, Hungary, reference: 2019/13, and conducted in accordance with the principles defined in the Declaration of Helsinki. Prior to their participation in the research, all caregivers provided written informed consent.

2.2.2 APPARATUS AND STIMULI

As in the case of Experiment 1, animations were created and rendered in Maya®, a 3D computer animation software (Autodesk Inc.). The resulting stimuli were displayed on the inbuilt monitor of a Tobii T60 XL eye tracker device (51.1 x 32.3 cm, 1920 x 1200 px), controlled by PyHab (Kominsky, 2019), a stimulus presentation system for PsychoPy (Peirce et al., 2019). The stimuli, experiment control scripts, and data files are available at https://osf.io/uf2bn/.

2.2.3 DESIGN AND PROCEDURE

In general, the procedure was similar to Experiment 1. In the first familiarization phase, infants saw six familiarization trials (each 8 seconds long). This phase aimed to teach that the agent had the goal of approaching the target object. The agent approached the target object over two identical distractors, independent of its location. Trials started with the agent standing in the center of the scene while facing the infant and jumping twice to attract attention. Then, it turned around and approached the target, whose location depended on the trial type. Both the target and the distractors had identical upper parts. Within this first familiarization phase, the types of trials alternated in a fixed order (target on the Left, Center, or Right: LCRCRL). The second familiarization phase was identical to Experiment 1.

The test phase was identical to the one in Experiment 1, except that the target was replaced with the second distractor object, and these two identical distractors hid in the tubes visibly or invisibly to the agent (therefore, the location of these objects was not counterbalanced, as they were indistinguishable). We paused the animation at the end of the detour in both test conditions and measured infants' looking time. We expected longer looks in the INEFFICIENT EPISTEMIC ACTION condition, where the agent's uncertainty about the target object's location could not justify the detour.

2.2.4 RESULTS

As in Experiment 1, raw looking-time data was log-transformed before analysis. As predicted, infants looked longer again in the INEFFICIENT (M = 18.27 s, SD = 8.28 s) compared to the EFFICIENT EPISTEMIC ACTION (M = 11.87 s, SD = 8.37 s) condition. Moreover, this difference was statistically significant according to the mixed ANOVA conducted (Figure 3). The analysis revealed a significant effect of condition ($F[1,22] = 9.259$, $p = .006$, $\eta^2_P = .296$) but not of presentation order ($F[1,22] = 1.229$, $p = .280$, $\eta^2_P = .053$), nor the interaction of these factors ($F[1,22] = 1.655$, $p = .212$, $\eta^2_P = .070$).

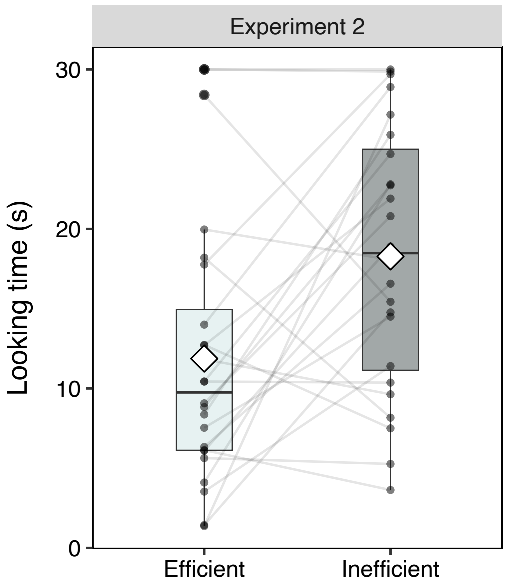

Figure 3. Boxplots of raw looking times (in seconds) for test trials in Experiment 2 (N = 24). White diamonds indicate means, bold horizontal lines indicate medians, boxes indicate interquartile ranges, and linked black dots indicate individual data points across conditions.

2.3 Discussion

The results from Experiments 1 and 2 indicate that 14-month-olds already recognize information-gathering behavior in the context of instrumental goal-directed actions. Infants observed actions that, though not directly efficient for achieving the actor's immediate goal of retrieving an object, facilitated visual access to this object's location. Infants looked less at this directly inefficient instrumental action when the actor was uncertain about the object's location than when the actor was knowledgeable (Experiment 1) or when the action could not yield relevant information from the agent's perspective (Experiment 2). Since looking-time differences are most often conceptualized as related to violation of expectations, this evidence is compatible with the hypothesis that infants could justify the action as information gathering in the former case but found it difficult to explain in the latter situations. Notably, this differentiation in our task required the combination of instrumental goal attribution, visual perspective taking, and the attribution of uncertainty.

Concerning the latter capacity, the specific content of the attributed uncertainty in this experiment is up for discussion. That is, in the EFFICIENT EPISTEMIC ACTION test condition, infants could have attributed uncertainty to the agent either about the location of the target object, the identity of the partially occluded ambiguous objects, or both. In fact, in our scenario, these two sources of uncertainty were not independent: although to accomplish the agent's goal of reaching the target, object location was most relevant, the potential object locations were highlighted by the ambiguous objects, as either could have been the target from the agent's perspective. That is, infants may have used the agent's uncertainty about object identities to attribute uncertainty about object locations. If so, infants' performance in this task also sheds new light on their grasp of the aspectuality of the mind.

Previous studies have indicated that infants understand how things appear to agents who are ignorant of their true nature. For instance, 15-month-old infants expect an agent looking for a blue-haired doll to choose a box displaying blue hair, despite knowing the box holds a different object (Song & Baillargeon, 2008; see Scott & Baillargeon, 2009 for a similar finding), and 17-month-olds anticipate that an agent will be misled by a thief who has replaced a desired object with a lookalike (Scott, Richman, & Baillargeon, 2015). These studies show that infants in their second year can appreciate that an agent may see an object that really is X as an object Y, where both X and Y *are properly defined*.

Appearances also played a role in our study, albeit they were not deceptive but ambiguous. In the test phase, we ensured that the agent saw some portion of the objects before its detour in all conditions, and we manipulated the visibility of those object parts that were relevant for *definite* identity disambiguation. In response, as indicated by the results, infants had different interpretations of the situation, reacting differently to the agent's subsequent detour. Thus, infants likely attributed to the agent different beliefs about the objects' identity in the two conditions, and the content of this attributed belief in the EFFICIENT EPISTEMIC ACTION CONDITION served as a justification for the agent's ensuing epistemic action. This content could not have been congruent with what infants perceived as reality – if it was, infants should have looked similarly in the two conditions. However, it is also unlikely that infants attributed the agent a definite, albeit false, belief about the object's identity, like in the studies above, since it is not evident how such a belief would justify information gathering. Instead, there is a third option that we suggest is compatible with the data: infants represented the agent as believing an object that really is X, as an object that is either X or Y. Chapter 5 will discuss the potential representational schemes infants could use to attribute such uncertain mental states.

Chapter 3. **Principle II: Drive toward novelty**

The results presented in Chapter 2 provide novel insights into how infants think of others' exploratory behavior, indicating, first, an early grasp of the causal relationship between the absence of sufficient access to goal-relevant variables and uncertainty and, second, an expectation that agents may engage in actions towards the sub-goal of reducing this uncertainty.

Task-related uncertainty, however, is not the only motivation behind humans' information gathering. We are continuously bombarded with information that may be relevant in the short or long term. It is often suggested that humans possess learning mechanisms that canalize learning by triggering information seeking when they stumble upon such learning opportunities. It is a longstanding observation that changes in the environment, in the form of novel or surprising stimuli, are likely to induce curiosity and subsequent exploratory behavior. This phenomenon was recognized already by Pavlov, who labeled it as the investigatory or "*What-is-it?*" reflex (Pavlov, 1927), but contemporary research also highlights that curiosity towards a stimulus depends on the degree of surprise and information gain it provides (e.g., Kidd, Piantadosi, & Aslin, 2012, 2014; Poli et al., 2020, 2022; Wu et al., 2022).

The reliable pattern between the detection of novelty and its consecutive investigation might also be used by infants in recognizing epistemic actions. In Experiments 1 and 2, based on the events of familiarization phases, infants could assume that the agent knew the two objects appearing during test trials. Therefore, it is unlikely that their differential looking time in the two conditions was driven by any attribution of novelty-induced curiosity. However, this does not exclude the hypothesis that infants can also make such attributions if provided with the appropriate cues.

3.1 Experiment 3

To test this hypothesis, we conducted Experiment 3. We used a similar paradigm as in previous experiments, keeping the test events from Experiment 1 but introducing slightly different familiarization phases. Importantly, events in these familiarization phases did not reveal any goal or knowledge of the agent related to the objects in the test phase, rendering these objects novel from the agent's perspective (see Figure 4).

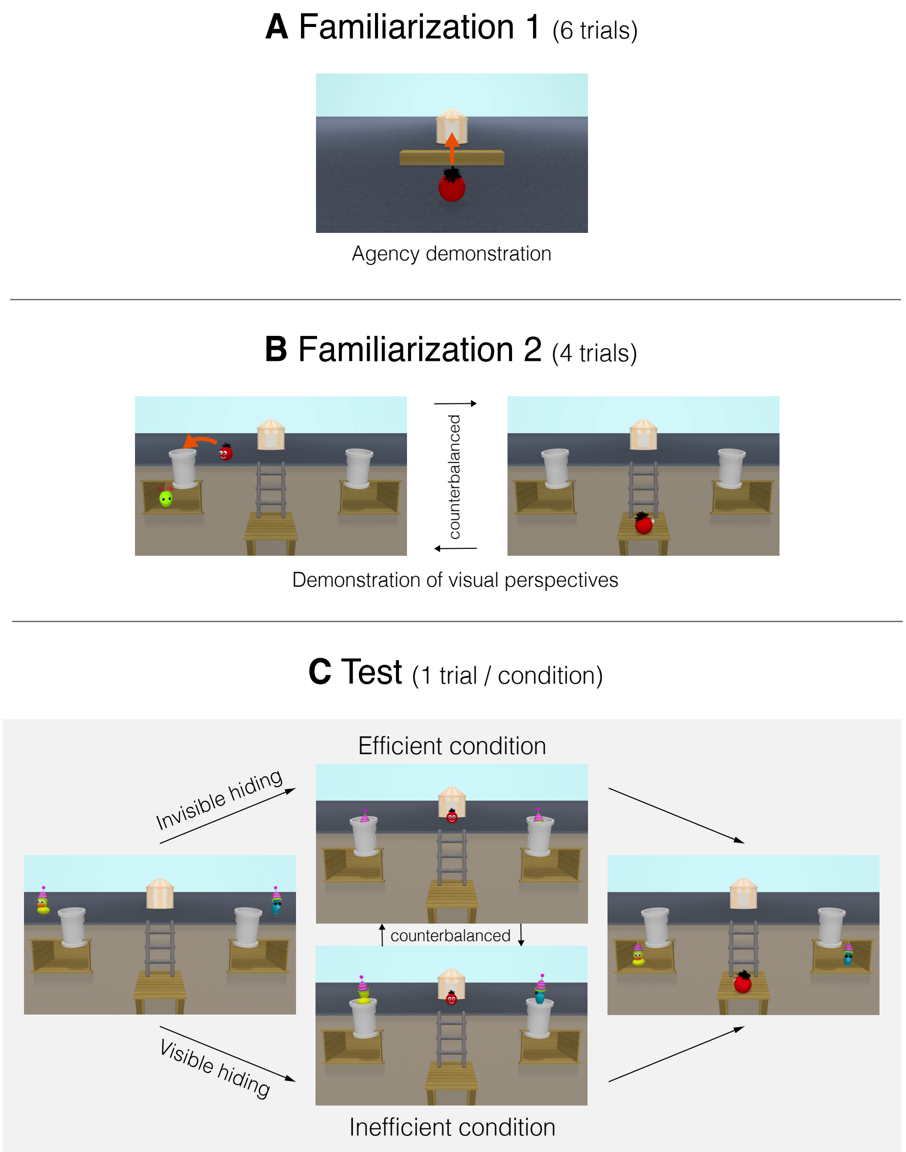

Figure 4. Structure of Experiment 3. The experiments follow a three-phase structure: (A) *Familiarization 1*, (B) *Familiarization 2*, and (C) *Test*. **Familiarization 1** depicts the agent acting instrumentally efficiently (approaching the house via jumping over the obstacle) to establish the agent's agentic nature without providing prior knowledge about its preference for later appearing test objects. **Familiarization 2** introduces the environment and the agent's possible perspectives, as in the previous experiments. The **Test** phase is identical to Experiment 1, but the agent is presented with the test objects (fully or partially) for the first time, introducing novelty from its perspective.

In the first familiarization phase, instead of the selective approach events (that induced a goal in the previous experiments), only the agentic nature of the agent was emphasized by showing a self-propelled, efficient goal-directed behavior: going to its house on a straight path when there was no obstacle on the way and jumping over an obstacle when there was one.

The second familiarization phase was similar to the one in Experiments 1 and 2, except that instead of the duck, a different object was shown as jumping into one of the tubes and being followed by the agent. Moreover, there was no object whatsoever in the scenario where the agent was standing on the platform below, looking towards the boxes. These changes were added to keep the temporal and familiarization structure of the previous experiments without teaching anything about any potential object-related goal of the agent.

Finally, infants were shown the same test conditions as in Experiment 1. However, this time, the two objects appearing during the test were novel to infants as well as to the observed agent. We reasoned that the partial occlusion of these now novel objects in the EFFICIENT EPISTEMIC ACTION condition could be sufficient for infants to attribute non-instrumental curiosity to the agent about the identity of these objects, justifying its subsequent detour as exploratory behavior. In contrast, in the INEFFICIENT EPISTEMIC ACTION condition, where there is no partial occlusion, and both objects could be observed by the agent before its detour, infants were predicted not to attribute such curiosity that could render the detour efficient and, therefore, to look longer compared to the EFFICIENT EPISTEMIC ACTION condition.

3.1.1 PARTICIPANTS

24 full-term, typically developing 14-month-old infants were included in the analysis of the experiment (14 females, M_{age} = 14 months, 17.25 days, SD_{age} = 6.51 days). In addition, 4 more infants were tested but not included in the final analysis, all due to crying or fussiness. All infants were recruited in Budapest based on birth records. At the end of the experiment, infants received a small gift for their participation.

The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB), Budapest, Hungary, reference: 2019/13, and conducted in accordance with the principles defined in the Declaration of Helsinki. Prior to their participation in the research, all caregivers provided written informed consent.

3.1.2 Apparatus and stimuli

As in Experiment 1 and 2, animations were created and rendered in Maya®, a 3D computer animation software (Autodesk Inc.). The resulting stimuli were displayed on the inbuilt monitor of a Tobii T60 XL eye tracker device (51.1 x 32.3 cm, 1920 x 1200 px), controlled by PyHab (Kominsky, 2019), a stimulus presentation system for PsychoPy (Peirce et al., 2019). The stimuli, experiment control scripts, and data files are available at https://osf.io/uf2bn/.

3.1.3 Design and procedure

The experiment took place in a soundproof room with dimmed lights. Participants were seated on their caregiver's lap, about 60 cm from the display. Caregivers wore opaque glasses that prevented them from seeing the stimuli. They were instructed to keep the child seated and not to interact with them. The experimenter was sitting behind a curtain, controlling the experiment and monitoring infants' behavior via a video camera. Infants watched two familiarization phases and a test phase.

In the first familiarization phase, infants saw six 8-second familiarization trials. These trials aimed to familiarize infants with the agent and convey information about its agentic nature. In these trials, the agent approached a house efficiently by jumping a barrier when it was present or directly when it was not. The trial types alternated in a fixed order for all subjects (Barrier or No barrier: BNNBBN).

In the second familiarization phase, infants were shown two 12-second trials. As in Experiments 1 and 2, this phase aimed to familiarize infants with the novel two-level environment where subsequent events took place. This phase was similar to that in the previous experiments, except that the target object was replaced with a novel object in trial type A, and the distractor was removed from the scene in trial type B. In trial type A, the fruit-like novel object was initially displayed next to a tube on the upper level (left in the first trial, right in the second). Then, as the agent emerged from a house in the center of the upper level, the object and the agent turned to face each other. Subsequently, the object jumped to the tube next to it, arriving at the box under the tube. The agent also jumped into this same tube and reached the object in the respective box. In trial type B, the agent stood on a platform below, facing the infant. Then, it turned to look at the boxes (which contained no objects) and then turned back without further action. Each trial was shown twice, counterbalanced between subjects in ABAB or ABBA order. The test phase was identical to Experiment 1.

We paused the animation at the end of the detour in both test conditions and measured infants' looking time. Looking times were recorded until the participants looked away for 2 consecutive seconds or had looked at the screen for a total of 30 seconds. The entire session was video recorded. To extract exact looking time durations for data analysis, looking behavior was coded offline frame-by-frame.

We expected longer looks in the INEFFICIENT EPISTEMIC ACTION condition, where the agent's uncertainty about the target object's location could not justify the detour. The preregistration of Experiment 3 is available at: https://doi.org/10.17605/OSF.IO/G7RM6.[11]

3.1.4 RESULTS

Infants looked longer in the INEFFICIENT ($M = 15.42$ s, $SD = 9.17$ s) compared to the EFFICIENT ($M = 11.28$ s, $SD = 8.39$ s) EPISTEMIC ACTION condition (Figure 5). The mixed ANOVA revealed a significant effect of condition ($F[1,22] = 5.568$, $p = .028$, $\eta^2_P = .202$), no main effect of presentation order ($F[1,22] = 1.081$, $p = .310$, $\eta^2_P = .047$), and an additional significant interaction of condition and presentation order ($F[1,22] = 5.448$, $p = .029$, $\eta^2_P = .198$).

Post-hoc analyses of this latter interaction (using pairwise t-tests with p-values adjusted via Bonferroni correction) indicated that the difference between the EFFICIENT and INEFFICIENT EPISTEMIC ACTION conditions was significant in the group that watched the INEFFICIENT

[11] There were two deviations from the preregistered analysis plan. First, while we initially planned to use t-tests to compare looking-times between conditions, we later recognized that presentation order introduced variability that a simple t-test could not account for. To address this, we employed mixed ANOVAs across all looking-time studies to more appropriately handle the data structure. Second, although we preregistered the use of Bayes Factors as an alternative to t-tests, this analysis was not conducted since we ultimately adopted the mixed ANOVA approach instead.

EPISTEMIC ACTION condition first (EFFICIENT M = 7.67 s, EFFICIENT SD = 5.02 s, INEFFICIENT M = 16.52 s, INEFFICIENT SD = 9.83 s, t[11] = -4.279, p = .002), but not in the group that watched the EFFICIENT condition first (EFFICIENT M = 14.90 s, EFFICIENT SD = 9.66 s, INEFFICIENT M = 14.32 s, INEFFICIENT SD = 8.76 s, t[11] = -.015, p =1.000).

This interaction in the current experiment is likely caused by infants being presented with two novel objects for the first time in the first test trials, leading to a predisposition to look at these trials longer, which interacted with the effect of condition. Importantly, however, this tendency to look longer on the first trial does not cancel out the significance of the main effect of action efficiency and cannot explain the data pattern itself (as infants did not look shorter when the INEFFICIENT CONDITION was presented second), but instead shows that the sequential presentation also had a general impact on engagement during the experimental session, as it is often found in looking-time studies (e.g., Baillargeon, 1987; Luo et al., 2003; Pietraszewski et al., 2017).

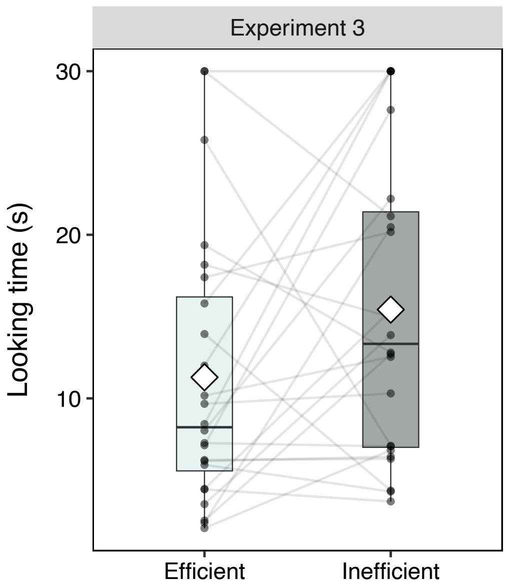

Figure 5. Boxplots of raw looking times (in seconds) for test trials in Experiment 3 (N = 24). White diamonds indicate means, bold horizontal lines indicate medians, boxes indicate interquartile ranges, and linked black dots indicate individual data points across conditions.

3.2 Discussion

In sum, Experiment 3 shows that infants look less at agents visually exploring objects when those objects are novel to the agent than when they are familiar. This suggests that in addition to being sensitive to the instrumentally relevant epistemic goals of others, infants can also attribute curiosity-driven goals to agents who face novel or ambiguous situations.

It is worth noting that although these results provide evidence for infants' expectation of exploration in response to novelty, it remains an open question how well-specified *novelty* is in infants'—and also in adults'—theory of epistemic actions. Consider the classical notion of the investigatory or orienting reflex, as described by Pavlov (1927) and Sokolov (1966): a detected sensory stimulus is compared to an existing memory of it, and if the two deviate above a certain threshold, head and eye movements are automatically directed towards the stimulus. That is, even in the earliest accounts of curious behavior, it is emphasized that there should be a threshold of novelty. This makes sense, as without such a limit, learners would be alert to the tiniest of changes and be prone to wasting their resources on nuisance variations, weakening their ability to acquire a generic understanding of their environment.

Later theoretical accounts abandoned the loose notion of '*novelty*' and argued that curiosity should be directed toward stimuli with optimal or intermediate levels of surprise (e.g., Berlyne, 1960; Hebb, 1955; Hunt, 1965). Kidd, Piantadosi, and Aslin (2012) tested a modern incarnation of this idea in an experiment with 7- to 8-month-old infants who were shown sequences of images with different predictability. Results indicated that participants' probability of looking away from the screen was highest to events of either very low or very high predictability, suggesting that events generating intermediate levels of surprise are perceived preferentially. A clear appeal of such a strategy is that it prevents learners from spending resources on overly predictable and unpredictable events.

However, avoiding completely novel or unpredictable stimuli also seems to be an untenable solution in the long term, as it leaves open the question of how learning can start off in the first place. In a recent proposal, Dubey and Griffiths (2020) argue that the degree of novelty that induces curiosity should be related to the predictability of the environment. In essence, they claim that a lower threshold of surprise is the optimal solution when the environment is unpredictable, that is when the agent can encounter any stimulus in the future, regardless of its previous occurrence. Whereas if environments are predictable, that is, stimuli more frequently encountered in the past are more likely to appear in the future, an inverse U-shaped curve describes the optimal relationship between curiosity and surprise.[12]

In sum, previous work suggests that it is not simply novelty that drives humans' interest; rather, it is the interplay between surprise and predictability that determines the optimal conditions for curiosity. However, since both of these quantities are relative to what a particular individual knows about her environment, it is difficult for an observer to approximate them from the third-person perspective.

Fortunately, observers may also rely on other sources to predict curiosity-driven information gathering. First, they might use their knowledge regarding the typicality of events and objects occurring within the perceptual field of an agent, and assuming that this knowledge is widely shared between individuals, they might expect curiosity about atypical stimuli. Much of the knowledge that children acquire daily is under the assumption that it will be common knowledge among others. Research indicates that preschoolers anticipate that generic information will be broadly known (Cimpian & Scott, 2012); moreover, they often find it challenging to comprehend that some individuals, such as infants, may not possess this generic knowledge about the world (Caza, Atance, & Bernstein, 2016). Thus, children may use this assumption of universality about generic knowledge when filtering what may be unexpected for others.

Second, observers could also pay attention to others' emotional expressions of surprise to infer interest. The ability to leverage others' emotional expressions seems to emerge early. Recently, Wu, Merrick, and Gweon (2024) demonstrated that 15-month-olds could use an experimenter's surprise expression to revise their expectations about statistically probable and improbable events. In this study, the experimenter sampled a ball from a box containing red and white balls, displaying either a surprised or unsurprised expression before revealing the

[12] These accounts, in one way or another, hypothesize that curiosity about stimuli depends on the degree of the immediate surprise they lead to. Yet another proposal, advocated by Kaplan & Oudeyer (2007), emphasizes how prediction errors change in time. According to their model, agents should act so that they optimize learning progress, avoiding predictable and unpredictable situations, and instead focusing on those that maximize prediction error reduction, thus circumventing unlearnable problems. Sometimes learning progress is quantified as information gain, that is, the overall reduction of uncertainty over a set of hypotheses (Poli et al., 2020).

ball to the infants. When the experimenter showed no surprise, infants looked longer at improbable outcomes, consistent with prior research. However, when the experimenter showed surprise, this typical pattern disappeared or reversed, suggesting that infants can use others' expressions of surprise to adjust their own expectations about events. It is at least conceivable that their surprise-based inferences also extend to the domain of action understanding, predicting interest in sources of surprise.

Third, observers might also use their own motivational or epistemic states to attribute curiosity. For instance, within our paradigm, the test objects were initially novel to both the infants and the observed agents, which yields an intriguing possibility: infants might have attributed interest to the agent based on their own initial curiosity, operating under the assumption that what sparks their interest would naturally be of interest to others as well. This would reflect an egocentric scaffolding of curiosity attribution, which, combined with encoding the agent's limited prior visual access to the objects, could have aided infants in interpreting the agent's behavior as epistemic in the EFFICIENT condition. Alternatively, infants could have adopted a more objective stance, hypothesizing that unknown objects attract exploration by default, independent of their own motivations towards these objects. It is a task for further experiments to elucidate whether infants also expect inquiry about entities that are highly familiar to them but novel to the observed agents.

Chapter 4. **Principle III: Efficiency**

Results from Chapters 2 and 3 converge to highlight infants' early capacity to discern and attribute information-related goals based on the epistemic states of the observed actors. Drawing upon this foundation, we designed Experiment 4 to explore further whether infants also anticipate the maximization of information gain relative to its cost.

4.1 **Experiment 4**

To test whether infants expect efficient information seeking, we measured their looking time in response to two different actions of an agent, both of which could be interpreted as information-directed, required the same effort, but provided different amounts of information. We reasoned that if infants expect others to maximize information gain, they should be surprised and look more at the action that provided less information than what would have been possible with equal cost (same physical cost).

The environment where the observed events played out was similar to the one in the previous experiments, with two notable differences. First, on the upper level, there were now three tubes leading to three boxes at the lower level. Second, two ladders connected the upper level to two different platforms situated at the lower level between the three boxes (Figure 6). Again, the experiment consisted of two familiarization phases and one test phase. In the first familiarization phase, across three trials, infants watched the target object jumping into one of the three boxes through a tube (this box was a different one in every trial according to our counterbalancing scheme), and the agent noticing this jump and following the target by jumping into the same box. In the second familiarization phase, in a further six trials, infants watched the agent initially standing in one of two platforms (A or B). From platform B, it could see the contents of two out of three boxes, while from platform A, it could see the content of only one box, as a wall obstructed the view of the remaining two boxes. Then, as before, the target jumped into one of three boxes. If the agent stood on the platform that afforded a view of the box where the target jumped, it followed the target by jumping into the same box. In contrast, if the agent stood on the platform which did not afford visual access to this box, it looked around and then remained idle. To balance the number of jumps to each box (while controlling for the number of times the agent successfully discovered the target from each platform), we added two filler trials simply showing the target jumping to the central box.

Finally, in the two conditions of the test phase, the agent and its target object first appeared in the center of the upper level; then, just as they made eye contact, an occluder appeared in front of the agent, blocking its sight entirely. Subsequently, the target jumped into one of the boxes (in the central box in each test) and left the scene altogether, unbeknownst to the agent. Afterward, the occluder disappeared, and the tubes flashed two times to highlight the possible hiding locations of the duck. Finally, the agent looked briefly at the flashing tubes, then moved to the platform where it could see the contents of two out of three boxes (Efficient epistemic action condition) or to the platform that provided a view to only one box (Inefficient epistemic action condition).

We paused the animation and measured looking time at the end of the detour (when the agent arrived at one of the platforms). We expected longer looks in the Inefficient epistemic action condition, where the agent gained less information about the target's location than possible from the other, equally distant platform.

4.1.1 Participants

24 full-term, typically developing 14-month-old infants were included in the analysis of the experiment (14 females, M_{age} = 14 months, 14.33 days, SD_{age} = 9.04 days). In addition, 6 more infants were tested but not included in the final analysis due to crying or fussiness. All infants were recruited in Budapest based on birth records. At the end of the experiment, infants received a small gift for their participation.

The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB), Budapest, Hungary, reference: 2019/13, and conducted in accordance with the principles defined in the Declaration of Helsinki. Prior to their participation in the research, all caregivers provided written informed consent.

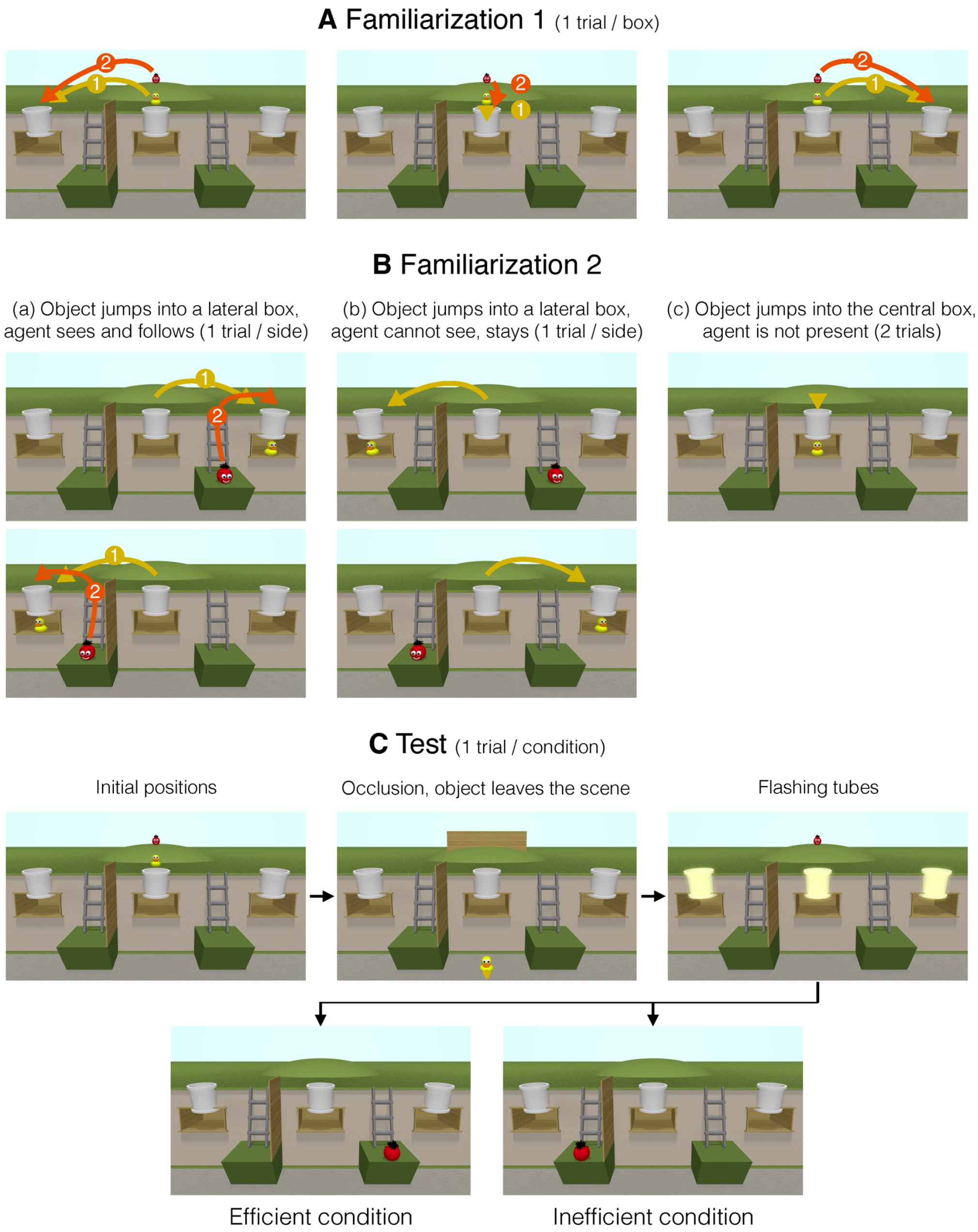

Figure 6. Structure of Experiment 4. Experiment 4 targets infants' expectations regarding the efficiency of information-seeking actions. It consists of three phases: (A) Familiarization 1, (B) Familiarization 2, and (C) Test. **Familiarization 1** involves the agent observing its target object jumping into one of three boxes and then following it (one trial per box). This phase is intended to familiarize infants with the agent's goal-directed behavior and the setup's spatial layout. **Familiarization 2** differentiates between scenarios where the target object's landing is either visible or not visible from the agent's perspective, further divided into (a) object in one lateral box is visible from the agent's platform, allowing the agent to follow it directly (one trial per side); (b) object is not visible from the agent's position on the opposite side, leading to no action from the agent due to the lack of visual access (one trial per side); (c) two additional trials show the object jumping to the central box, without the agent being present to balance the frequency of jumps to the three boxes. The **Test** phase examines infants' responses to efficient versus inefficient information-seeking actions. Initially, the agent and the object are at the scene's center. Then, an occluder blocks the agent's view as the object leaves, introducing a scenario where the agent lacks information about the object's new location (which jumps into the middle tube and then leaves the scene). Then, the tubes flash briefly to indicate the potential hiding locations of the object, highlighting the choice the agent faces. EFFICIENT EPISTEMIC ACTION CONDITION (left path): The agent chooses a platform that provides a view of two out of three boxes, maximizing information gain. INEFFICIENT EPISTEMIC ACTION CONDITION (right path): The agent opts for a platform with a view of only one box, a less efficient information-seeking action given the same effort.

4.1.2 APPARATUS AND STIMULI

As in Experiments 1 to 3, animations were created and rendered in Maya®, a 3D computer animation software (Autodesk Inc.). The resulting stimuli were displayed on the inbuilt monitor of a Tobii T60 XL eye tracker device (51.1 x 32.3 cm, 1920 x 1200 px), controlled by PyHab (Kominsky, 2019), a stimulus presentation system for PsychoPy (Peirce et al., 2019). The stimuli, experiment control scripts, and data files are available at https://osf.io/uf2bn/.

4.1.3 DESIGN AND PROCEDURE

The experiment took place in a soundproof room with dimmed lights. Participants were seated on their caregiver's lap, about 60 cm from the display. Caregivers wore opaque glasses that prevented them from seeing the stimuli. They were instructed to keep the child seated and not to interact with them. The experimenter was sitting behind a curtain, controlling the experiment and monitoring infants' behavior via a video camera.

In this experiment, we had a similar environment as in Experiments 1 to 3, with the main change that instead of two tubes and boxes, we had three, and instead of one ladder, we had two. As in the previous experiments, there were two levels. In this case, on the upper level, there was a low hill where the agent initially stood and three tubes, each ending in a box in the lower level. Additionally, two ladders linked the upper level to two corresponding platforms on the lower level. These platforms (and their respective ladders) were situated between the left and center and the center and right boxes. In addition, on the right side of the left platform, a wall was blocking the view to the center and right boxes. Thus, the left platform only provided a view to the left box, while the right platform only provided a view to the center and right boxes.

In the first familiarization phase, which aimed to convey that the agent's goal was to reach the target object, infants saw three 15-second familiarization trials depicting the target object jumping to one of the three tubes and their corresponding boxes once, followed by the agent. The trial types alternated in two orders (box Left, Center, or Right: CLR or LRC), counterbalanced between subjects. The second familiarization phase included six trials of three types (A, B, C). This phase aimed to familiarize infants with the agent's perspectives when standing on the two platforms.

In trial types A and B, the agent initially stood on one of two platforms on the lower level (one platform offering a view of two boxes and the other of only one box due to the view-blocking wall), and the target object jumped into one of the lateral tubes falling into the box underneath. In trial type A (19 seconds), the agent went up the ladder and followed the target to its box as its location was visible from its platform. In trial type B (15 seconds), the target was not visible to the agent, so the agent inspected the empty box/es and remained stationary. Both types of trials were presented once per agent position (left or right platform).

In two additional type C trials (7 seconds), the target jumped to the central box without the agent's presence. These C trials ensured an even distribution of the target's jumps to the three boxes. The six trials alternated in two types of order, counterbalanced between subjects (AL: type A and the agent on the left platform, AR: type A and the agent on the right platform, BL: type B and the agent on the left platform, BR: type B and the agent on the right platform, C: object jumps to the central box; orders: ALBRCCARBL or CBLARBRALC).

In the test phase, infants were presented with one trial per condition (EFFICIENT and INEFFICIENT EPISTEMIC ACTION). The trials were equal in length (28 seconds) but differed in content, depending on the condition. In both conditions, the target first stood in front of the agent, facing the infant. Then, the target turned around, looking at the agent. The agent jumped two times, and then an occluder landed between them. While the agent's view was occluded, the target left the scene, jumping to the lower level and going off-screen at the middle bottom part of the screen. After the occluder disappeared, the agent inspected the three tubes which were flashing (to highlight the possible locations of the target) and moved either to the left (Inefficient epistemic action condition) or the right (Efficient epistemic action condition) platform by climbing down the respective ladder. The order of presentation of the two test conditions was counterbalanced between subjects.

We paused the animation at the end of the detour in both test conditions and measured infants' looking time. Looking times were recorded until the participants looked away for 2 consecutive seconds or had looked at the screen for a total of 30 seconds. The entire session was video recorded. To extract exact looking time durations for data analysis, looking behavior was coded offline frame-by-frame.

We expected longer looks in the INEFFICIENT EPISTEMIC ACTION condition, where the detour was not the most efficient action means to gain information. The preregistration of Experiment 4 is available at: https://doi.org/10.17605/OSF.IO/7X6DN.[13]

4.1.4 RESULTS

Infants looked longer in the INEFFICIENT (M = 16.06 s, SD = 9.03 s) than in the EFFICIENT (M = 11.70 s, SD = 8.44 s) EPISTEMIC ACTION condition (Figure 7). The mixed ANOVA revealed a significant effect of condition ($F[1,22] = 4.719$, $p = .041$, $\eta^2_p = .177$), no main effect of presentation order ($F[1,22] = .737$, $p = .400$, $\eta^2_p = .032$), and an additional significant interaction of these factors ($F[1,22] = 4.947$, $p = .037$, $\eta^2_p = .184$). Further analysis of this interaction, using pairwise t-tests adjusted by Bonferroni correction, revealed a significant difference in the group that viewed the INEFFICIENT condition first (EFFICIENT M = 8.79 s, EFFICIENT SD = 7.01 s, INEFFICIENT M = 17.43 s, INEFFICIENT SD = 9.74 s, $t[11] = -3.929$, $p = .004$), but not in the group that viewed the EFFICIENT condition first (EFFICIENT M = 14.61 s, EFFICIENT SD = 9.01 s, INEFFICIENT M = 14.69 s, INEFFICIENT SD = 8.45 s, $t[11] = .031$, $p = 1.000$).

---

[13] There were three deviations from the preregistered analysis plan. First, although we preregistered exploratory analyses of pupil data during the critical phase when the agent made its detour, the data were too sparse and noisy to support reliable statistical inferences. Specifically, only 15 out of 24 participants had more than 50% valid pupil datapoints in the critical phase of both test conditions. This limitation primarily arose because Experiments 1 to 5 were tailored primarily to measure looking-time responses, not pupillometry. Each participant saw each test condition only once, and caregivers were instructed following standard looking-time protocols rather than procedures optimized for minimizing movement and luminance variability—both critical factors for reliable pupil measurements. While the availability of the Tobii eye-tracker made it feasible to preregister exploratory pupil analyses, the design was not optimized for this purpose. The preregistration reflected the possibility that, if looking-time measures proved insufficiently sensitive, preliminary differences in pupil dilation might emerge, which could then inform the design of future studies specifically tailored for pupillometric investigation. However, given the high degree of data loss and measurement noise, we did not pursue further analyses of pupil data in this experiment.

Second, while we initially planned to use t-tests to compare looking-times between conditions, we later recognized that presentation order introduced variability that a simple t-test could not account for. To address this, we employed mixed ANOVAs across all looking-time studies to more appropriately handle the data structure. Third, although we preregistered the use of Bayes Factors as an alternative to t-tests, this analysis was not conducted since we ultimately adopted the mixed ANOVA approach instead.

These findings suggest that similarly to Experiment 3, the presentation order influenced the effect of conditions, interacting with the impact of condition when the EFFICIENT EPISTEMIC ACTION condition was watched first. The interaction observed may be attributed to the fact that the target object disappeared in test trials, a novel event in the overall experimental session that might have been unexpected the first time infants saw it. Consequently, this could lead to longer overall looking times in the first test trials. However, importantly, the overall longer looks during the first trial cannot explain the main finding that action efficiency affected infants' engagement, specifically, that they looked longer when the agent took the path that led to comparatively less information than it would have possible if the other path had been taken, indicating an expectation of information-related efficiency.

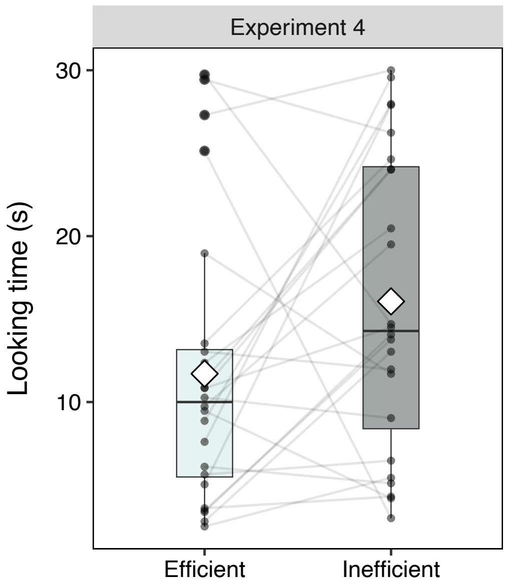

Figure 7. Boxplots of raw looking times (in seconds) for test trials in Experiment 4 (N = 24). White diamonds indicate means, bold horizontal lines indicate medians, boxes indicate interquartile ranges, and linked black dots indicate individual data points across conditions.

4.2 Discussion

Experiment 4 sheds light on an additional layer of infants' conception of epistemic actions: an assumption of efficiency in the domain of information seeking. The fact that infants looked longer when an agent chose a less informative path compared to a more informative one with the same cost suggests that they have some rudimentary expectation of information maximization in inquiry. It is, however, an open question what exactly this expectation amounts to.

First, infants might reason about information in the statistical sense. In this framework, the value of any particular piece of information an observation provides is determined by its ability to alter the probability of a set of hypotheses that the inquiring agent entertains. In our paradigm, the hypotheses would concern the target object's possible locations, and the agent could gain more information by choosing an observation that changes two rather than one location's probability. Given that both observation options had the same cost, infants could conclude that the agent should move to the place that provides the more informative observation in this sense.

Alternatively, infants might possess a less abstract concept of information maximization, such as the maximization of visibility or perceptual access to objects of interest. In the context of our study, this would mean that infants predict the agent's actions based on which position—observing one box from platform A or two from platform B—affords a greater likelihood of direct observation of the target. This concept seems rather intuitive but may not fully capture the nuances of human inquiry: it would be limited to valid interpretative inferences only in cases when the agent's target of interest is a discrete observable state (such as an object's present location, shape, or material). In these contexts, agents can indeed maximize information by maximizing the likelihood of perceiving the target itself.

However, humans' targets of interest are often not directly observable (they may be spatiotemporally distributed or distal, abstract, or entirely hypothetical), such that inquiry about them is limited to perceptually assessing states that are informative only indirectly. For instance, when a person examines footprints in the snow, her actual interest likely lies in identifying and locating the unseen entity that made them, and it would be premature to conclude that her ultimate epistemic goal is to learn about the footprints per se, just because her actions are increasing the chances of visually perceiving the footprints themselves. Future work should uncover whether infants' understanding of information gathering also extends to such indirectly informative observations.

Whichever path of reasoning they follow, infants should compute possibilities and corresponding likelihoods. Using the first, richer concept of information would require infants to attribute representations of probability distributions over a range of possible states to other agents (in our task, the possible hiding locations). As reviewed in Chapter 1, it is currently debated whether infants themselves can represent such sets of possibilities. If they can, it may be hypothesized that they can also attribute them as contents of others' mental states. This issue will be explored in the next chapter.

On the other hand, using the second, leaner concept limited to perceptual access is not necessarily less demanding cognitively. For instance, to apply it in our paradigm, infants should have assessed the likelihood of the target being in either of the two boxes visible from platform A or in the single box visible from platform B and decide which platform to visit, all from the agent's perspective—as the target, from their own perspective, was not even in the scene. Alternatively, instead of attributing the whole planning process to the agent, infants could also rely on the processes involved in their own decision-making, effectively taking the agent's inferred mental state contents as the input of these processes instead of their own (Schulz, 2011). This approach would sidestep the need to attribute reasoning about the likelihood of action outcomes, relying instead on their first-person reasoning processes.

So far, the results of Experiment 4 were interpreted as evidence that infants expect agents to gather information in line with a richer or linear concept of information maximization. However, another interpretation suggests that the observed pattern of looking times might not require infants to represent informational efficiency at all.

According to this alternative, infants may have construed the wall in the experimental scenario not only as an occluder of visual access but also as a physical barrier restricting movement. From this perspective, infants might have simply tracked the last known location of the target and expected the agent to act instrumentally efficiently — that is, to choose the shortest and least obstructed path to reach the location where the object was last seen. Given that the object jumped into the central box, and that this location appeared to be more easily accessible from the right side (unblocked by the wall), it would have been surprising for the agent to detour to the left side, which was partially obstructed.

On this account, infants' longer looking times in the Inefficient condition can be explained by a violation of expectations about physical path efficiency toward a known goal state, without requiring any attribution of uncertainty or reasoning about information gathering. In other words, infants could show this pattern of looking by simply assuming that the agent knew where the object was and should have chosen the most direct, least obstructed route to that location.

While this is a theoretically plausible interpretation, several aspects of the experimental design reduce its likelihood. First, the critical event where the target object exited through the central box was visible to the infant but not to the agent, whose view remained occluded at that point. Thus, infants had clear grounds to attribute ignorance to the agent regarding the duck's final location. If infants had instead assumed that the agent believed the duck was in the central box, both test conditions should have appeared equally surprising, as the agent should have directly approached the central box in both cases. This expectation is further supported by the familiarization trials, where infants repeatedly saw the agent approach the target in the central box directly and efficiently.

Moreover, although the wall may have visually separated the left and right parts of the scene, the agent's prior movements during Familiarization 1 demonstrated that it was fully capable of moving freely between sides on the upper part of the environment. This repeated exposure to unimpeded cross-scene movement should have mitigated any strong inference that the wall imposed a physical constraint on the agent's actions.

Nevertheless, fully ruling out this interpretation would require further empirical testing. A future study could modify the test scenario by having the target enter and leave the scene through the left box. If infants truly represent agents' information-seeking behavior as guided by considerations of informational efficiency, they should still find it surprising when the agent chooses the less informative path, even when that path appears physically less restrictive. Conversely, if the present results can be explained solely by tracking instrumental efficiency toward a last known object location, this reversal should lead to a different pattern of looking times. Such future studies would help clarify whether infants' expectations in this scenario genuinely involve reasoning about epistemic efficiency or can be accounted for by simpler interpretations based on physical access and efficient instrumental goal pursuit.

Chapter 5. **Attribution of uncertainty**

As discussed in Chapter 1, a productive intuitive model of epistemic action would need to be causal. That is, it should specify the kind of mental states that cause agents to engage in information gathering. In our adult, full-fledged understanding, this motivating mental state is assumed to be the acting agent's uncertainty about particular instrumental or non-instrumental variables.

In Section 1.3, two potential representational schemes were hypothesized that observers may use to attribute uncertainty to other individuals. The difference of these schemes lies in the content of the attributed mental state:

- **Modal scheme:** uncertainty represented as an attitude towards a set of possibilities. Example: *THE BALL IS AT* X_1, *OR* X_2, *OR* X_3.
- **Open-ended scheme:** uncertainty represented as an attitude towards an open proposition, one or more arguments of which are empty, noted by a placeholder that specifies what is common in all possibilities (e.g., a location, or, more specifically, a box). Example: *THE BALL IS AT X.*

Each experiment presented in Chapters 2 to 4 contained some ambiguity from the perspective of the observed agent, but it is not evident which representational scheme infants used to attribute uncertainty to her. In fact, this question of representation equally applies to other results indicating early capacity to reason about others' lack of goal-relevant knowledge (e.g., Liszkowski, Carpenter, & Tomasello, 2008; Tauzin & Gergely, 2018).

One way to discriminate between the above hypotheses is to test whether subjects are willing to attribute certainty to an agent based on the agent's *inferential* access to critical variables. Specifically, only the modal scheme allows infants to attribute (or simulate) productive disjunctive inferences to other agents who are faced with negative evidence about a variable in question.

For instance, if an observed agent is shown that its target object is not at A and B from three possible locations (A, B, C), the modal scheme would allow infants to attribute the agent a belief that the object is at location C, by exclusion. While if only location A could be excluded from the agents' perspective, she should be represented as remaining uncertain between options B and C. In contrast, the open-ended scheme does not allow such inferences, as the possibilities that the agent entertains are not attributed in the first place. Consequently, as the two representation schemes license different inferences regarding an observed agent's knowledge, they should allow different inferences about her subsequent actions as well, similar to the previous experiments.

5.1 **Experiment 5**

Following this logic, we conducted Experiment 5 with 14-month-old infants. Participants observed actions that appeared inefficient concerning the immediate goal of the agent reaching its target object and could either be justified as appropriate epistemic actions or not, depending on whether the agent could *infer* the target object's location (either at location A, B, or C) prior to this action. When such justification was absent, we anticipated comparatively longer looking times, under the hypothesis that they use the modal representation scheme for uncertainty attribution. The setup (Figure 8) was akin to that of Experiment 4. At the upper level, three tubes led to three different boxes on the lower level, connected by two ladders to two platforms between the boxes. The experiment included two familiarization phases and a test phase.

The first familiarization phase aimed to acquaint infants with the scenario and the three potential hiding places (boxes A, B, or C) for the target object (a duck). Over three trials, we demonstrated the target jumping into one of the three boxes through its corresponding tube in each trial, varying the box per our counterbalancing scheme. A distinctive sound accompanied the target's jump. In the second familiarization phase, in a further six trials, infants watched the agent initially standing in one of two platforms (left or right, counterbalanced order). Then, the target, starting from the center of the upper level, jumped into one of three boxes. In each case, the agent turned around and could see where the target jumped, and it followed the target by jumping into the same box. The role of this phase was to familiarize infants with the agent's goal of reaching the target object as well as the various perspectives it can take in the environment.

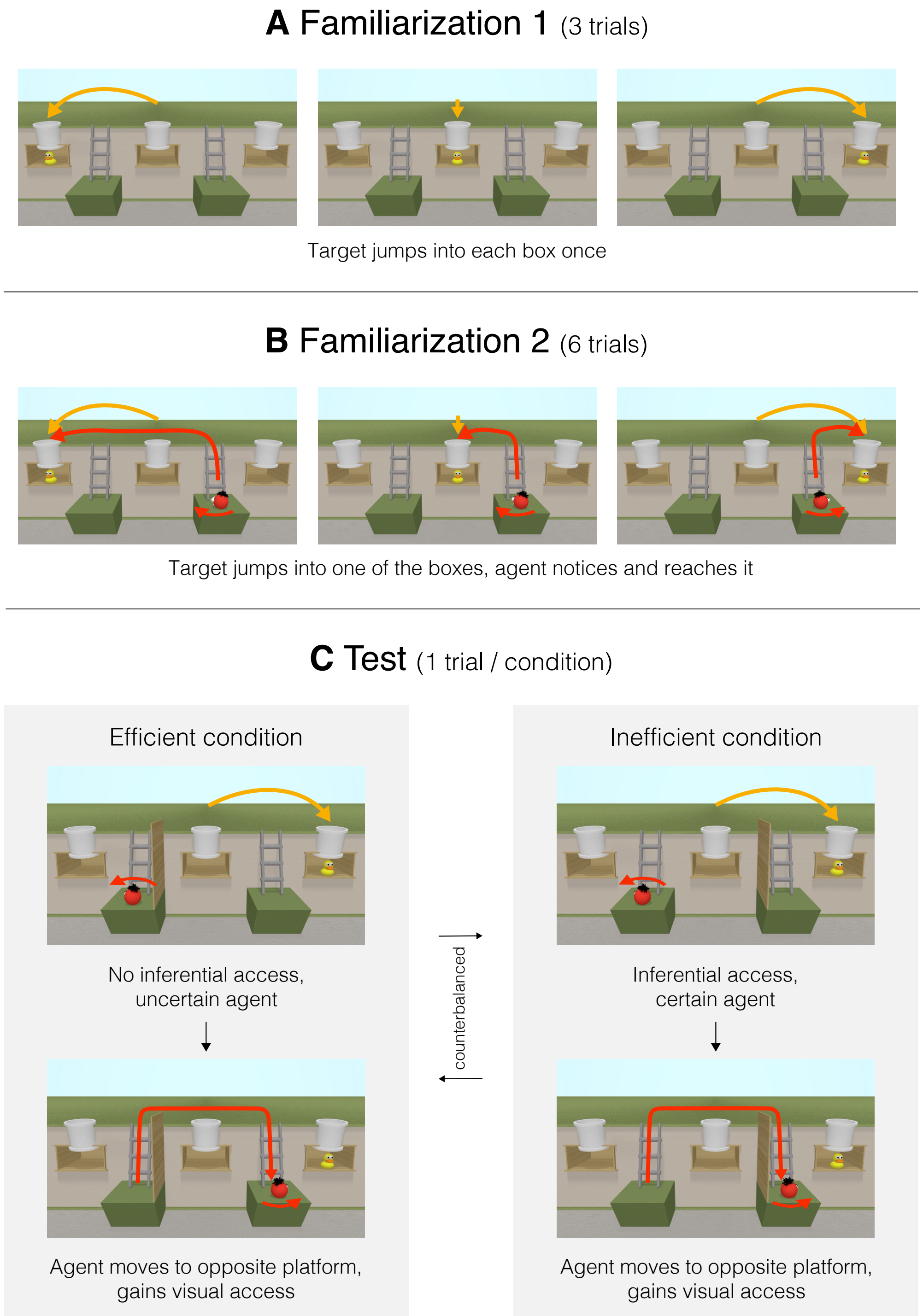

Figure 8. Structure of Experiment 5. Experiment 5 targeted infants' understanding of inference-based knowledge. It consisted of three phases: (A) Familiarization 1, (B) Familiarization 2, and (C) Test. **Familiarization 1** introduced infants to three potential hiding spots for a duck target, showing the target jumping into each box via a tube across three trials. In the six trials of **Familiarization 2**, infants observed the agent on one of two platforms (left or right, in a counterbalanced order), tracking the duck as it jumped into one of the three boxes, thereby familiarizing them with the agent's goal and the various perspectives within the environment. The **Test** phase presented two conditions, each starting with the agent on a platform and the target at the center of the upper level. In the INEFFICIENT EPISTEMIC ACTION condition, an occluder blocked the agent's view of the opposite lateral box (but not the other two), leading the agent to redundantly check the box where the duck hid, despite its inferential access to its location. In contrast, the EFFICIENT EPISTEMIC ACTION condition had the occluder also block the center box, meaning the duck's hiding place was uncertain, thus justifying the agent's identical detour to gain a visual access.

In the test phase, infants were shown two conditions (one trial per condition), both starting with the agent on one of the platforms (counterbalanced between subjects) and the target at the center of the upper level. In the first condition, an occluder blocked the agent's view of box C opposite its platform (the rightmost box from the left platform and vice versa) but not the other two boxes A and B. The target hid in this occluded box C. This situation allowed infants to deduce that the agent might know by inference that the target was at location C, as locations A and B were visibly empty from her perspective. Despite this, the agent then moved to the opposite platform for a redundant view of the occluded box C. We labeled this scenario as the INEFFICIENT EPISTEMIC ACTION condition due to the unnecessary nature of this detour, similar to the previous experiments.

In the EFFICIENT EPISTEMIC ACTION condition, the occluder blocked the agent's view to the center (B) and the lateral box (C) as well, after which the target hid in the lateral box. This scenario allowed for multiple possibilities regarding the target's location from the agent's perspective, as only one box was visibly empty from her perspective, justifying the same detour to the opposite platform that was redundant in the INEFFICIENT EPISTEMIC ACTION condition as an information-gathering action.

We paused the animation to measure the infants' looking time at the end of the detour, predicting longer durations in the INEFFICIENT EPISTEMIC ACTION condition compared to the EFFICIENT EPISTEMIC ACTION condition, where the detour lacked justification due to the agent's prior inferential knowledge of the target's location.

5.1.1 PARTICIPANTS

A total of 24 full-term, typically developing 14-month-old infants were included in the analysis (N = 24; 11 girls; M_{age} = 14 months, 9.70 days; SD_{age} = 6.33 days). All infants were recruited in Budapest for research purposes, based on birth records. 5 additional infants were tested but not included due to crying or fussiness (4) and equipment failure (1). Infants received a small gift for their participation at the end of the experiment.

The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB), Budapest, Hungary (reference: 2019/13), and conducted in accordance with the principles defined in the Declaration of Helsinki. Prior to their participation in the research, all caregivers provided written informed consent.

5.1.2 APPARATUS AND STIMULI

As in Experiments 1 to 4, animations were created and rendered in Maya®, a 3D computer animation software (Autodesk Inc.). The resulting stimuli were displayed on the inbuilt monitor of a Tobii T60 XL eye tracker device (51.1 x 32.3 cm, 1920 x 1200 px), controlled by PyHab (Kominsky, 2019), a stimulus presentation system for PsychoPy (Peirce et al., 2019). The stimuli, experiment control scripts, and data files are available at https://osf.io/uf2bn/.

5.1.3 DESIGN AND PROCEDURE

The experiment took place in a soundproof room with dimmed lights. Participants were seated on their caregiver's lap, about 60 cm from the display. Caregivers wore opaque glasses that prevented them from seeing the stimuli. They were instructed to keep the child seated and not to interact with them. The experimenter was sitting behind a curtain, controlling the experiment, and monitoring infants' behavior via a video camera.

In the first familiarization phase, infants saw three 14 seconds long familiarization trials depicting the target object jumping to one of the three tubes and their corresponding boxes once. The second familiarization phase included six trials (19 seconds each) of the agent, initially standing in one of the platforms, noticing the target jumping into one of the boxes. We counterbalanced the box of hiding in the overall 9 familiarization trials (3 levels of order: C-B-A-A-B-B-C-A-C, A-C-B-B-C-C-A-B-A, B-A-C-C-A-A-B-C-B) and the direction where the agent started to turn before noticing the goal-object in the second familiarization phase (2 levels: right-right-left-right-left-left or left-left-right-left-right-right).

In the test phase, infants were presented with one trial per condition (EFFICIENT and INEFFICIENT EPISTEMIC ACTION). The trials were equal in length (29 seconds) but differed in content, depending on the condition. In both conditions, the agent first stood on one of the platforms, while the target was displayed in the center of the upper level, both facing the infant. Then, an occluder ascended from above, and occluded the agent's view either only to the opposite lateral box (INEFFICIENT condition), or also to the center box (EFFICIENT condition). Then, the target jumped into the lateral box, emitting a sound. The agent turned around, observing the visually accessible box(es), then moved to the opposite platform, observing the boxes available from that position, when we started measuring looking time. The order of the test conditions (2 levels: EFFICIENT-INEFFICIENT, INEFFICIENT-EFFICIENT), and the initial position of the agent (2 levels: right, left) were counterbalanced.

Looking times were recorded until the participants looked away for 2 consecutive seconds or have looked at the screen for a total of 30 seconds. The entire session was video recorded. To extract exact looking time durations for data analysis, looking behavior was coded offline frame-by-frame. The preregistration of the experiment is available at: https://doi.org/10.17605/OSF.IO/7PQBC.[14]

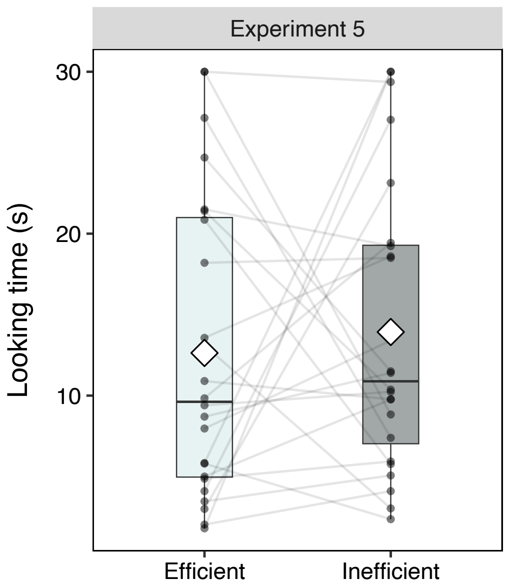

Figure 9. Boxplots of raw looking times (in seconds) for test trials in Experiment 5 (N = 24). White diamonds indicate means, bold horizontal lines indicate medians, boxes indicate interquartile ranges, and linked black dots indicate individual data points across conditions.

[14] There were three deviations from the preregistered analysis plan. First, although exploratory pupil analyses were preregistered for the same reasons outlined in relation to Experiment 4, the study design was similarly optimized for looking-time measures, resulting in sparse and noisy pupil data that precluded reliable statistical inference. Specifically, only 18 out of 24 participants had more than 50% valid pupil datapoints in the critical phase of both test conditions. Second, instead of the preregistered t-tests, we used mixed ANOVAs to account for variability introduced by presentation order. Third, although Bayes Factors were preregistered as an alternative to t-tests, we did not conduct this analysis after adopting the mixed ANOVA approach.

5.1.4 RESULTS

On average, infants looked slightly longer in the INEFFICIENT (M = 13.92 s, SD = 8.87 s) condition compared to the EFFICIENT (M = 12.63 s, SD = 9.24 s) condition (Figure 9). However, the mixed ANOVA revealed a no significant effect of condition ($F[1,22] = .695$, $p = .413$), no main effect of presentation order ($F[1,22] = .265$, $p = .612$), and no significant interaction of these factors ($F[1,22] = 3.519$, $p = .074$).

5.1.5 DISCUSSION

The goal of Experiment 5 was to probe the content of mental states that 14-month-olds attribute to uncertain agents. One hypothesis we put forward was that infants grasp others' uncertainty using a modal representational scheme, attributing mental states with mutually exclusive possibilities as contents. To test this hypothesis, we asked whether 14-month-olds can infer an agent's knowledge concerning the location of an object if she has enough evidence to infer it via disjunction. Specifically, if an agent is faced with sufficient negative evidence with regards to two out of three possible hiding places of an object (it is not in A and B from A, B, or C possible locations), she should be attributed the belief that the target location must be the last possible one (C). Crucially, only beliefs with modal content have the appropriate format for this inference. Therefore, we reasoned that if infants can attribute knowledge based on negative evidence, they can only do it by attributing such modal content. However, our results fail to show that this is the case. We found no statistically significant difference between infants' looking time in response to an agent's epistemic action depending on this agent's prior inferential knowledge.

This null-result could mean that 14-month-olds cannot attribute modal contents to other agents. However, in light of the recent evidence that infants by this age can already reason about mutually exclusive possibilities (Cesana-Arlotti et al., 2018; Cesana-Arlotti, Kovács, & Téglás, 2020; but for a contrasting account, see Leahy & Carey, 2020; Feiman, Mody, & Carey, 2022), it is worth examining other potential factors contributing to the observed null result.

One possibility is that infants' limitations lie not in their capacity to attribute modal representational contents to observed agents, but in their failure to update these contents based on the negative evidence (in this case, the absence of the target in two out of three boxes) available to these agents—for example, due to not being able to integrate well inferential abilities in their Theory of Mind at this young age.

The current consensus is that children do not recognize inference as a source of others' knowledge until the age of 6 or later (Sodian & Wimmer, 1987; Miller, Hardin, & Montgomery, 2003; Pillow, 1999; Pillow et al., 2000). In the seminal study of Sodian and Wimmer (1987), children were presented with an experimenter who was aware of two key facts: all objects in a specific container were identical, and one of these objects had been transferred to an opaque bag. If children grasped that knowledge of these premises allows the correct inference concerning the bag's contents, they would recognize that the experimenter knew what was in the bag (by the rule of universal instantiation: all Xs are Ys, the transferred object is an X, therefore it must be a Y). To ensure that children themselves could infer the critical conclusion from the premises, their evaluation of the experimenter's knowledge of the bag's contents was compared to their assessment of their own knowledge when they knew the premises but did not see the transfer. Moreover, to control for egocentric answering (attributing knowledge due to their own knowledge), other conditions were added where the container had two types of balls, so the experimenter's knowledge depended on directly observing the transfer. Analysis of children's answers revealed three response patterns: the correct pattern (affirmative responses to knowledge questions with one type of ball and affirmative responses based on perceptual access with two types of balls), neglect of inference (attributing ignorance when the experimenter could infer but did not observe object identities), and the yes-bias pattern (affirmative answers to all questions). Most 4-year-olds showed incorrect patterns, while most 6-year-olds showed correct assessments, and the primary error in younger children was neglecting inferential access, and this error persisted even if they shared the experimenter's visual perspective.

More closely related to the present question, Pillow (1999) conducted a version of this paradigm that tested whether children attribute knowledge to a puppet based on inference via elimination of possibilities. Children were introduced to a scenario where they had to assess both their own knowledge and the puppet's knowledge regarding the colors of hidden toys. Each trial began with the children being introduced to a puppet, followed by the presentation of two different-colored marbles. After engaging in a discussion about the marbles' colors, both were hidden in separate cans. Depending on the trial, either the child or the puppet peeked into one of the cans. Subsequently, the children were asked to answer whether they or the puppet knows the color of the marble in the other can. While 4-year-olds succeeded above chance only in assessing their own knowledge, six-year-olds showed high accuracy in both self and other questions across all scenarios.

However, one should be cautious in drawing definite conclusions from such evidence to the prelinguistic Theory of Mind of young children. These studies mainly show whether children map others' inference-based mental state kinds to the noun *"knowledge"* or the verb *"know"*. Participants are asked to determine if an observed agent *knows* something based on what she could logically infer, rather than to specify *what* the agent actually infers, the final conclusion of her reasoning. Therefore, low performance in these tasks does not necessarily imply that children fail to make the correct content attribution; rather, it means they do not associate the epistemic state derived through inference with the appropriate linguistic terms. In fact, in a control experiment of Sodian and Wimmer (1987), in response to the question *"Did he/she/you know this or did he/ she/you guess?"* children judged even the experimenter's explicit verbal statement of the correct conclusion as a *"guess"*, rather than *"knowledge"*, and similarly, most of them considered themselves as having guessed the bag's content in the case of their *own* correct inferences. This result aligns with research suggesting that a mature comprehension of propositional attitude verbs develops around the age of 6 or even later (Abbeduto & Rosenberg, 1985; Johnson & Maratsos, 1977; Macnamara, Baker, & Olson, 1976; Mitchell & Robinsion, 1992; Moore, Bryant, & Furrow, 1989; Moore & Davidge, 1989).

To assess whether children can attribute, rather than linguistically categorize, inference-based epistemic states, paradigms that test their ability to use these attributions, for instance, to predict behavior, would be more suitable. To the best of my knowledge, there has been only one study that avoided directly questioning children about mental state categories and instead measured their predictions in social interactions. Building on the concept of mutual exclusivity in word learning, Rai and Mitchell (2007) explored whether 5- to 6-year-olds could predict which item another child would associate with a new label, given their knowledge of the labels for the other items. Mutual exclusivity (or referent disambiguation) studies involve showing a child two objects, one familiar and one novel, and asking them to find the object corresponding to a novel label (e.g., "*dax*"). Many studies have demonstrated success in these tasks at about 18 months of age, with children typically choosing the novel object as the referent (Bion, Borovsky, & Fernald, 2012; Halberda, 2003; Markman & Wachtel, 1988). In their adaptation of this paradigm, Rai and Mitchell presented two children, Child 1 and Child 2, with three cartoon characters, two of which were named either before or after Child 2 had left the scene. The key question was whether Child 1 could predict if Child 2, upon their return, would be able to correctly identify a newly mentioned character named using a novel label based on whether Child 2 had heard the earlier naming episodes. Both 5- and 6-year-olds were significantly more likely to predict that Child 2 would point to the target on trials where Child 2 had premise information sufficient to make an inference than when Child 2 did not have such premise information.

One interpretation of successful performance in referent disambiguation tasks is that the child eliminates the object with a known label, operating under the assumption that a single object cannot possess multiple labels. As the novel label clashes with the known one, the child disregards that object with the known label as a possible referent and selects the other object instead. On the other hand, according to alternative explanations, children succeed because they are inclined to assign new names to unknown objects by default, rather than deliberating and discarding objects with known labels (Feiman, Mody & Carey, 2022; Golinkoff, Mervis, & Hirsh-Pasek, 1994). If children also expect others to map new names to new objects and this heuristic is what drives their success in the study of Rai and Mitchell, then their results provide little evidence for inference-based knowledge attribution requiring a modal representation scheme. Therefore, other paradigms are required to provide stronger evidence for this capacity, where successful performance cannot be explained by the application of heuristics.

A second possibility behind the null result of Experiment 5 is that infants did not have a problem with attributing inference-based knowledge to the agent, but they either had low confidence in this attribution judgment or attributed low certainty to the agent herself. When considering how children typically judge others' epistemic sources, their confidence in knowledge attributions based on perception in general exceed that of attributions based on inference. If children notice that an agent has a clear and observable line of sight to an object, they can confidently ascribe a correct belief about the object's position to the agent, thanks to the dependable link between seeing and knowing. Studies show that not only young infants (e.g., Luo & Johnson, 2009), but primates (Hare et al., 2000) and scrub jays (Emery & Clayton,

2001) also rely on this causal relationship to the extent that they make use of cues to others' visual access. In contrast, the perceived reliability of disjunctive reasoning and the certainty with which infants or young children assess others' abilities to make such inferences seems to be lower. Children's confidence in their *own* deductive inferences relative to their own inductive inferences and guesses develops between 4 and 6 years of age (Pillow, 2002; Pillow & Anderson, 2006), and they do not consistently consider others certainty based on inference higher than based on guesses in the preschool years (Pillow et al., 2000).

Therefore, it is plausible that infants were more ready to accept information seeking as justified in the INEFFICIENT condition in Experiment 5 than in Experiments 1 to 4, because they judged the agent's certainty only slightly more probable than her uncertainty. Unfortunately, violation of expectation (VoE) paradigms are ill-suited to indicate such uncertain attribution judgments, assuming that looking-time patterns depend on the infants' certainty in their predictions. This holds true even under the postdictive processing account of VoE paradigms, which suggests that infants' differential responses to unexpected events can be explained by their success in reconciling these events with their prior theories after they occur, rather than by their prediction errors (Wellman, 2011). If infants assigned only a slightly higher probability to the hypothesis that the agent correctly identified the target's location than to the hypothesis that the agent remained uncertain, the latter hypothesis provided sufficient postdictive explanatory power for the agent's detour without overwhelming the infants cognitively, leading to shorter looking times.

To summarize, the inconclusive results of Experiment 5 could be due to factors other than infants' inability to apply a modal representation scheme. One potential explanation is infants' difficulty to update attributed modal representational contents based on negative evidence available to the observed agent, due to a developmental lag in including inferential capabilities in their Theory of Mind. However, most existing studies on older children's understanding of inference as a knowledge source may not accurately reflect their ability to attribute inference-based epistemic states, focusing instead on their comprehension of the language describing these states. The sole paradigm that attempts to assess children's use of these attributions for predicting behavior could be influenced by heuristic reasoning. Moreover, the uncertainty of inferential knowledge compared to observational knowledge might lead infants to ascribe less certainty to the former, affecting their responses in tasks that do not sufficiently account for subtle variations in certainty.

5.2 Experiment 6

Therefore, in Experiment 6, we sought to address these issues by developing a new paradigm to test preschoolers' capacity to attribute inference-based epistemic states to others. We designed a task requiring nonverbal responses from participants, which could be clearly evaluated on a binary scale, thus potentially making nuanced attribution judgments more explicit. Additionally, we aimed to ensure that the conclusions drawn from disjunctive inference could be attributed to agents only through eliminating attributed possibilities within a modal representational scheme. We focused on preschoolers because, as per the literature review above, existing studies on inference-based knowledge attribution have not conclusively determined their actual capabilities in this domain, leaving us without a clear benchmark for starting investigations targeting infants' abilities.

In the experiment, 4-year-old children engaged in a tablet-based, interactive game, where each round required them to feed one of three specific animals hidden in a box with its preferred food, which they had previously learned. To succeed in feeding, participants first had to select one of two agents to pass the correct food to them that they could then give to the animal. Although the participants themselves always knew which animal was in the box, the knowledgeability of the agents—either through direct vision or disjunctive inference—was manipulated. We were interested in whether the children would consistently seek assistance from the agent who was in a position to know the animal's identity, based on the type of knowledge source available.

The experiment was structured into three phases: two familiarization stages followed by a test phase, all revolving around a three-level box setup displayed on the tablet (Figure 10ab). Each level acted as a hiding spot for an animal (pig, dog, bear), with each type of animal's hiding place fixed across trials. To underscore this fixed arrangement, symbols corresponding to each animal were displayed next to their respective levels.

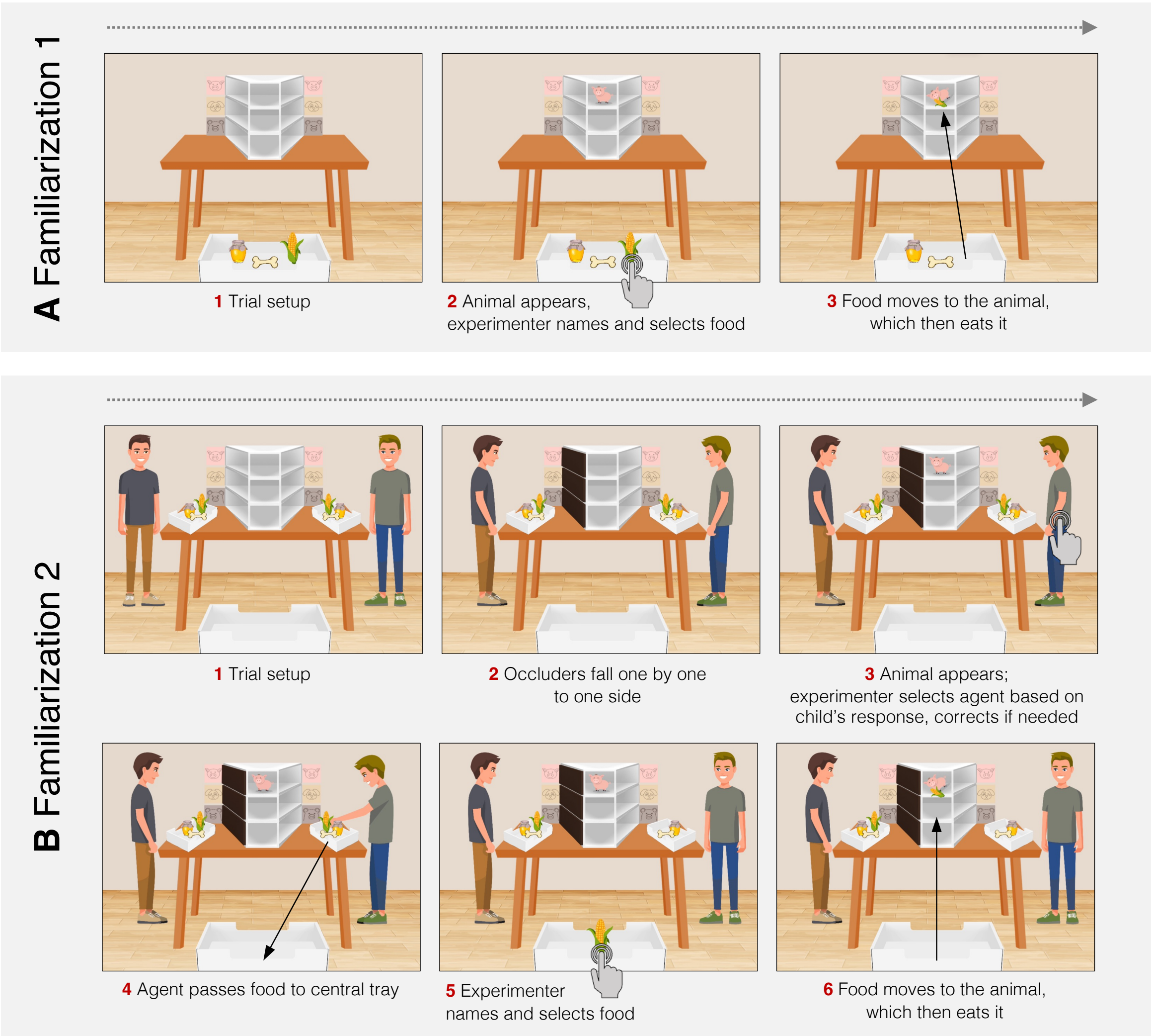

Figure 10a. Structure of Experiment 6. The experiment consists of two Familiarization phases and a Test phase. **Familiarization 1 (A):** The first familiarization phase begins with the basic setting, where the experimental setup is presented to the participants. In three trials, an animal appears in the box, and the experimenter names the animal and selects the appropriate food for it. The selected food then moves to the animal, and the animal consumes it. **Familiarization 2 (B):** Each 6 trials of the second familiarization phase starts with the basic setup. Then, occluders fall one by one to one side of the box, limiting the visual access to the animal from that side. An animal appears, and the child is prompted to select an agent to assist in feeding. The experimenter provides corrective feedback if the child selects the incorrect agent. Once the correct agent is chosen, he passes the food to the central tray. The experimenter names and taps on the food, which then moves to the animal, and the animal eats it.

The first familiarization phase involved simple demonstrations where the experimenter fed the animals directly, making explicit which animal ate which food and at which level they resided to ensure that all participants started with the same basic understanding of the game.

The second phase introduced two agents displayed at the two sides of the central box and required the children to choose the agent who was knowledgeable of the animal hiding in the box in the given trial to assist the experimenter in feeding.

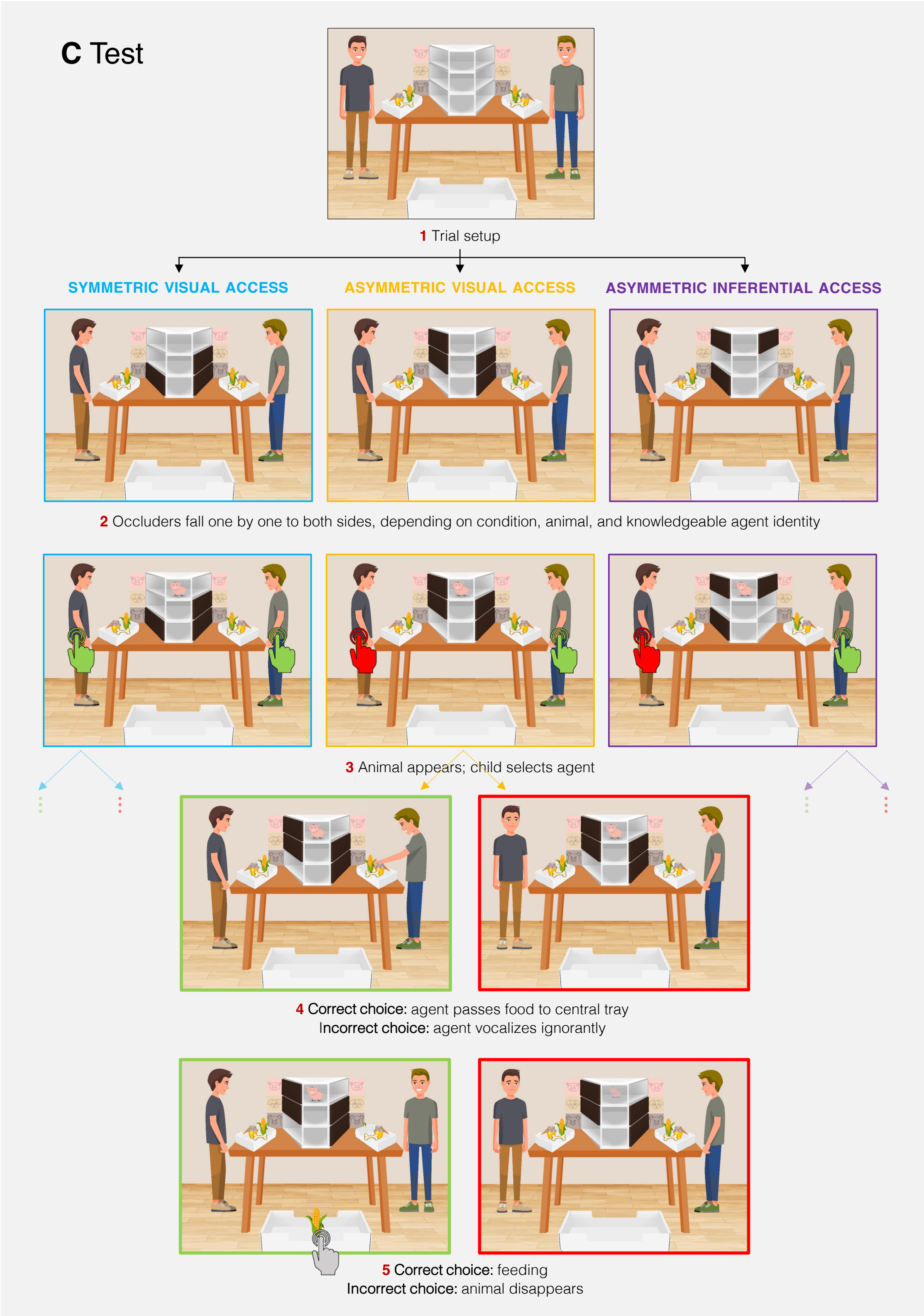

Figure 10b. Structure of Experiment 6 (*continued*). Test phase (C): The test phase involves 18 trials designed to assess the children's own decision-making under three conditions. The trial setup is presented, followed by the occluders falling one by one to both sides of the box, varying according to the condition—symmetric visual access, asymmetric visual access, or asymmetric inferential access. An animal appears (the pig, in this example), and the child selects an agent by tapping on the tablet screen. If the correct agent is selected, he passes the food to the central tray, and the food then can be moved to the animal for feeding. If the incorrect agent is chosen, the agent vocalizes ignorance, and the animal disappears (the correct and incorrect outcomes are only shown for the asymmetric visual access condition, but they are similar for all conditions).

From here on, occluders played a crucial role in manipulating the two agents' knowledgeability. While the front of the box, facing the participants, was always open, the lateral sides could be blocked or left visible using occluders for each level.

With the occluders, we could manipulate visibility of the content of each level from the perspective of the two agents. In this second familiarization phase, one agent always had full visual access, while the other had none, guiding children to discern who was knowledgeable about the box's contents purely through visual access. Above chance performance in this phase was used later as an inclusion criterion.

In the test phase, participants were asked to select the agent they believed knew the correct food by tapping on the agent's icon on the tablet screen. In each trial, they decided which agent to select under three different conditions, varying according to the agents' knowledge of the animal in the box. In the SYMMETRIC VISUAL ACCESS condition, both agents could see the animal inside the box. In the ASYMMETRIC VISUAL ACCESS condition, only one agent had a clear view of the animal. In the ASYMMETRIC INFERENTIAL ACCESS condition, neither agent could see the animal directly. However, the knowledgeable agent could deduce the animal's identity by elimination, given that only the level with the animal was hidden from his view, and the other two levels were visible.

We recorded participants selection in each test trial and were mainly interested in whether they chose the correct agent in the asymmetric inferential access condition, where this agent could only be attributed knowledge about the hidden animal's identity within a modal representational scheme.

5.2.1 PARTICIPANTS

To estimate our target sample size, we employed a mixed-effects logistic regression model to simulate scenarios using the LME4 (Bates et al., 2015) and SIMR packages (Green & MacLeod, 2016) in R (R Core Team, 2023), testing sample sizes ranging from 10 to 100 participants. The model incorporated condition as a fixed effect and participant as a random effect, assuming a moderate effect size with an odds ratio of 2 (when comparing the critical ASYMMETRIC INFERENTIAL ACCESS to hypothetical chance performance). The simulations suggested a sample size of 24 participants to achieve an 83.3% power level at a 0.05 significance threshold.

Therefore, a total of 24 typically developing 4-year-olds were included in the analysis (N = 24; 12 girls; M_{months} = 5.79; SD_{months} = 3.07). 10 additional children were tested but not included, 1 due to equipment error during the session, 4 due to chance (or below chance) performance in the second familiarization, and 5 due to completing less then 2/3 of the test trials.

We conducted tests in both labs of the CEU Cognitive Development Center, located in Vienna and Budapest, due to the availability of children. Despite the different locations, testing conditions were very similar, and we did not anticipate any differences between the children tested in the two labs. Children received a small gift for their participation at the end of the experiment. The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB), Budapest, Hungary (reference number: 2022-69), as well as the Psychological Research Ethics Board (PREBO) of CEU PU Austria, Vienna (reference number: 2024-17), and conducted in accordance with the principles defined in the Declaration of Helsinki. Prior to their participation in the research, all caregivers provided written informed consent, while the children provided verbal informed consent.

5.2.2 Apparatus and stimuli

The experiment was designed, run, and controlled in the Gorilla Experiment Builder (Anwyl-Irvine, Massonnié, Flitton, Kirkham & Evershed, 2020) on a tablet (12.4 inches display with 1752x2800 px resolution, 16:10 aspect ratio). Cartoon characters and objects were downloaded from the Adobe Stock service and further manipulated in Adobe Illustrator (Adobe Inc.) if necessary. The stimuli, experiment control scripts, and data files are available at https://osf.io/uf2bn/.

5.2.3 Design and procedure

The experimental session took place in an empty room, lasting about 12 to 15 minutes, depending on the onset of responses. Participants were seated in front of a tablet next to the experimenter. Caregivers were present but positioned so that they could observe without intervening. Caregivers were asked not to talk to their child, remain passive, and generally not to influence the child's behavior beyond ensuring they pay attention to the experimenter.

The experiment was structured into two familiarization phases followed by a test phase. Each phase unfolded within a setup presented on the tablet, featuring a central table with a three-level box placed on it. Each level served as a hiding spot for an animal. The hiding places of the animals were fixed for all participants: pig in the upper level, dog in the middle level, bear in the lower level. To ensure that participants could learn that the mapping of animals to the levels were fixed, additional symbols corresponding to the animal in the level were displayed on both sides of the level. The front side of the box, facing the participant and the experimenter, remained open, allowing the contents—specifically, the animal hidden during a trial—to be visible to them. However, occluders could be placed on the left and right sides of the box, blocking the view from those angles. Occluders could cover one, two, or all three levels, so we could manipulate the agents' visual access to the contents of each level. Each trial started with these occluders falling into their place, one by one (500 ms per occluder), followed by the appearance of the animal, which moved from side to side so that it was obvious from which angle it is visible (4500 ms).

Two trays were positioned on the table, one to the left and one to the right of the box. During the initial familiarization phase, these trays remained empty, as no agents were involved at this stage. In the second familiarization phase and the test phase, these trays held three types of food—corn, bone, and honey—corresponding to the preferences of the animals. Additionally, a central tray was placed closer to the participant's perspective and away from the table. During the second familiarization and test phases, when an agent was selected by tapping, he either correctly passed the appropriate food from the side trays to the central tray if he knew the content of the box (1000 ms) or, if he did not know, expressed ignorance with a sad facial expression and vocalization, leading to the animal disappearing and the next trial beginning (7500 ms). If the correct food was passed to the central tray, the participant could then tap on it to feed the animal, which responded with happy vocalizations as the trial concluded (4500 ms).

The session began with the first familiarization phase (3 trials), where the experimenter introduced the feeding game to the child ("*Pay attention, animals will appear in this box. To make* *sure they are well-fed, we need to provide them their favorite food.*"). This phase involved three trials where the experimenter directly fed each animal (pig, dog, bear; one animal per trial) with the correct food item (corn, bone, honey) while verbally reinforcing which animal liked which food and in which level of the box they "lived" (*"Here lives the [animal], it eats [food]. Let's feed it.*").

In the second familiarization phase (6 trials), the experimenter introduced the two agents who would help in feeding the animals. The child was required to request help from the correct agent based on the agent's knowledge of the animal in the box. This was done by asking the child which agent should be asked for help ("*Well, we don't have the food here. From now on, these two people will help with the feeding. We need to ask them for the animals' food*.", "*Here is the [animal]. Who should we ask for help?*"). This interaction was guided by the experimenter, who tapped on the appropriate agent on the tablet if the child gave the correct response ("*Yes, he knows that the [animal] is in the box*."; the experimenter taps on the chosen agent; "*He gave us the [food]. Let's feed the [animal] with the [food].*") and provided corrective feedback if the child made an error ("*No, he doesn't know what's in the box right now. The other person knows.*"; experimenter presses on the correct agent; "*He gave the [food]. Let's feed the [animal] with the [food].*"), thereby teaching how to play the game and reinforcing its objective. In the trials of this phase, one agent had full visual access to all levels of the box, while the other had no access whatsoever, as all levels were occluded from his side. Thus, the correct response was predetermined, based on who could know the content via visual access.

In the test phase (18 trials), the participants themselves were asked to make their choice by tapping on the agent they believed knew the correct food on the tablet screen ("*Now try to do it yourself.*"). They decided which agent to ask for help in each trial under three different conditions that varied according to the agents' knowledge of the animal in the box. We used a within-subjects design, where each participant was exposed to all three conditions, and we measured their decision on each trial (tap on the left or right agent).

In the SYMMETRIC VISUAL ACCESS condition, both agents could see the animal inside the box. However, two empty levels were concealed from one side, whereas from the opposite side, only one empty level was occluded. This uneven distribution of occluded levels was applied to measure a possible baseline bias in favor of choosing the agent who could see more levels.

In the ASYMMETRIC VISUAL ACCESS condition, only one agent had visual access to the animal. Both sides of the box had two occluders, but for the agent who could see the animal, only the empty levels were occluded. For the ignorant agent, one of the occluded levels included the level where the displayed animal resided. This uniform occlusion across both sides was intended to prevent participants from learning the heuristic that agents with fewer occluders are more likely to know the contents of the box (as this could potentially confound our results in the next condition)

In the ASYMMETRIC INFERENTIAL ACCESS condition, neither agent directly saw the animal. However, the knowledgeable agent could deduce which animal it is through elimination, as only the level with the animal was concealed from his viewpoint, and the visibility of the other two empty levels allowed him to infer which animal was hidden. For the ignorant agent, both the level with the animal and an additional level was obscured, hindering his ability to infer the correct animal.

Feedback was provided by the experimenter after each selection to reinforce learning outcomes—praise for correct choices ("*Yes, well done*.") and gentle correction for incorrect ones ("*No, he doesn't know what's in the box right now*.").

The overall 18 test trials were divided into 3 blocks of 6 trials each. Within these blocks, trials were assigned pseudo-randomly for each participant, with the constraint that each block had to contain two instances of the three conditions, the three animals, and knowledgeable agent identity (*left, right, or both*) values. Additionally, no condition, animal, or knowledgeable agent could repeat consecutively twice within or across block transitions. The look of the two agents were fixed per participant but counterbalanced between subjects *(right agent: black hair, brown trousers; left agent: blonde hair, blue trousers*; or vice versa). The trials of the first familiarization phase were presented in the following fixed order for all participants: *pig, dog bear*. The trials of the second familiarization phase were also presented in a fixed order for all participants (animal / knowledgeable agent): *bear / right, pig / left, bear / left, dog / right, pig / right, dog / left*.

The preregistration of the experiment is available at https://osf.io/f8qz7.

5.2.4 RESULTS

Our main variable of interest was participants rate of correct responses only in the two asymmetric access conditions – since in the SYMMETRIC VISUAL ACCESS condition either response (left or right agent) was correct in each trial, as both agents had sufficient knowledge to assist participants correctly. In the ASYMMETRIC VISUAL ACCESS condition, one agent had clear visual access while the other did not, and we expected a significant preference for the knowledgeable agent. In the ASYMMETRIC INFERENTIAL ACCESS condition, neither agent had direct visual access; however, one could infer the correct answer. Here, we expected that participants would select the inferentially knowledgeable agent more often than would be expected by chance.

Descriptive statistics indicated that the aggregated mean proportion of correct responses in the ASYMMETRIC VISUAL ACCESS condition was 0.903 (SD = 0.21) and 0.708 (SD = 0.354) in the ASYMMETRIC INFERENTIAL ACCESS condition (Figure 11).

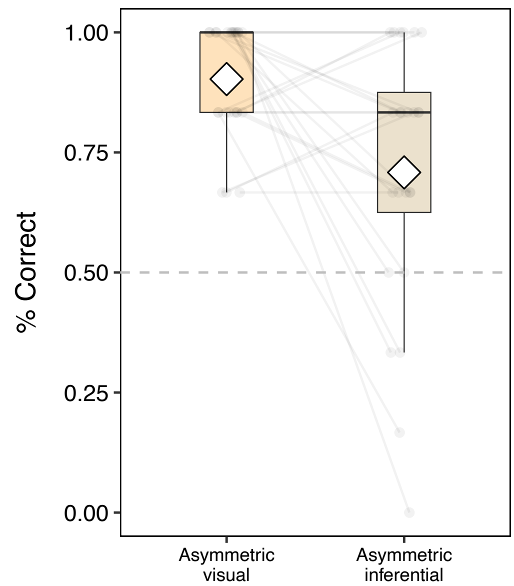

Figure 11. Boxplots of aggregate mean proportion correct scores in Experiment 6 (N = 24). White diamonds indicate means, bold horizontal lines indicate medians, boxes indicate interquartile ranges, and linked black dots indicate individual data points across conditions.

In order to find out whether participants were successful in choosing the knowledgeable agents, a generalized linear mixed model (GLMM) was employed to assess the probability of correct responses in the two asymmetric conditions while accounting for individual variability among participants. The model was specified with the following formula: `Response ~ 1 + Condition + (1 + Condition | Block) + (1 | Participant),` using the binomial family with a logit link function.[15] This model targeted the binary outcome variable `Response` (indicating whether a response was correct or incorrect) and included a fixed effect for `Condition` to capture systematic differences across conditions. Additionally, random effects were specified for `Block` and `Participant`, with random slopes and intercepts for `Condition` within `Block`, allowing the model to capture potential changes in the effect of `Condition` due to learning effects, as participants received feedback after each trial. The model was fitted using the LME4 package (Bates et al., 2015) in R (R Core Team, 2023), and maximum likelihood estimation was used for parameter estimation.

The analysis revealed that the likelihood of correctly responding in the ASYMMETRIC VISUAL ACCESS condition was significantly higher than what would be expected by chance (Intercept = 2.3673, SE = 0.3253, $z = 7.278$, $p < 0.0001$). Additionally, there was a significantly lower probability of correct responses in the ASYMMETRIC INFERENTIAL ACCESS condition compared to the ASYMMETRIC VISUAL ACCESS condition ($\beta = -1.4084$, SE = 0.3443, $z = -4.090$, $p < 0.0001$). While considerable variability was observed across participants ($\sigma^2 = 0.3542$), the variance associated with blocks was negligible, suggesting minimal block-level variability. To further examine potential learning effects, a second GLMM was also employed where `Block` was added as a fixed effect. However, no block coefficient proved to be significant ($p > 0.05$), indicating that block effects did not contribute significantly to the variability in correct responses. When random intercepts for testing location (Budapest or Vienna) were also added to the model, the results remained the same and the variance due to testing location was minimal ($\sigma^2 = 0.0042$), indicating a negligible impact of lab location on response variability.

Importantly, predictions from the model indicated that children performed significantly above chance in both conditions. Specifically, in the ASYMMETRIC INFERENTIAL ACCESS condition, children demonstrated a correct response probability of 72.29% ($z = 4.30$, $p < 0.0001$), significantly surpassing the chance expectation. Performance was even more

[15] For all GLMM analyses reported here, models of the form `y ~ 1 + Condition + (1 + Condition | Block) + (1 + Condition + Block | Participant)` were also fit to the data. These models allowed for random slopes for both `Condition` and `Block` at the participant level, in addition to the random intercepts already included in the primary models. While these more complex models provided similar quantitative results and the same qualitative conclusions regarding the effects of `Condition`, they resulted in poorer model fit as indicated by BIC values. Therefore, the simpler models were retained for reporting.

pronounced in the ASYMMETRIC VISUAL ACCESS condition, with a correct response probability of 91.43% ($z = 7.49$, $p < 0.0001$).

Traditional binomial tests were also employed to further evaluate performance against a chance expectation of 50% correct responses. The aggregated results by block indicated significant deviations from chance in all blocks for both conditions (Table 1). The overall aggregated proportions confirmed significant differences from chance in both conditions, with consistently higher performance in the ASYMMETRIC VISUAL ACCESS condition (Table 2). These binomial test results corroborate the findings from the GLMM, reinforcing that performance in both the ASYMMETRIC VISUAL and INFERENTIAL ACCESS conditions was significantly above chance levels, while performance was comparatively higher in the ASYMMETRIC VISUAL ACCESS condition across all measures.

Condition	Block	Mean	SD	SE	Binomial p-value
ASYMMETRIC VISUAL ACCESS	1	0.875	0.221	0.0451	<0.0001
ASYMMETRIC VISUAL ACCESS	2	0.938	0.224	0.0458	<0.0001
ASYMMETRIC VISUAL ACCESS	3	0.896	0.207	0.0423	<0.0001
ASYMMETRIC INFERENTIAL ACCESS	1	0.708	0.359	0.0732	0.0055
ASYMMETRIC INFERENTIAL ACCESS	2	0.688	0.355	0.0726	0.0133
ASYMMETRIC INFERENTIAL ACCESS	3	0.729	0.361	0.0736	0.0021

Table 1. Binomial test results for correct responses by condition and block.

Condition	Mean	SD	SE	Binomial p-value
ASYMMETRIC VISUAL ACCESS	0.903	0.216	0.0255	<0.0001
ASYMMETRIC INFERENTIAL ACCESS	0.708	0.354	0.0417	<0.0001

Table 2. Overall aggregated binomial test results for correct responses by condition.

We also evaluated the potential bias of choosing agents with less occluded levels in front of them, that could confound interpretations of the above chance performance in the asymmetric inferential condition, where the knowledgeable agent was always the one with less occluders in front of him. We conducted analyses focusing on two conditions—SYMMETRIC VISUAL and ASYMMETRIC INFERENTIAL. Both conditions were characterized by an uneven number of occluders, however, in the SYMMETRIC VISUAL condition, the number of occluders was independent from the knowledgeable agent (as both agents were always knowledgeable in this condition) that allowed us to assess whether participants exhibited a baseline preference for agents with fewer occluders in front of them.[16]

First, a GLMM was applied to investigate whether there was a systematic preference for agents with fewer occluders. As in the previous GLMM, the model incorporated condition as a fixed effect, with random intercepts for participants and random intercepts and slopes for condition within each block, to account for potential learning effects. The fixed effects analysis revealed no significant difference in the choice of agents with fewer occluders between the two conditions ($z = 0.281$, $p = 0.779$). Specifically, the estimated log odds of choosing the agent with fewer occluders were not significantly higher in the ASYMMETRIC INFERENTIAL condition compared to the SYMMETRIC VISUAL condition ($\beta = 0.07727$, $SE = 0.27538$). The model's intercept, indicating the baseline log odds of selecting the agent with fewer occluders in the SYMMETRIC VISUAL condition, was significant (Estimate = 0.97585, $SE = 0.27824$, $z = 3.507$, $p < 0.001$). This suggests a general propensity towards choosing agents with fewer occluders across conditions. The random effects analysis showed significant variability between participants ($\sigma^2 = 0.8739$) but negligible variability across blocks, indicating consistent performance across the test phase.

Second, traditional binomial tests were conducted to compare the proportion of trials in which agents with fewer occluders were chosen against a chance level of 0.5. These tests revealed that in both conditions, participants selected the agent with fewer occluders in front of him at rates significantly above chance ($p < 0.05$ across all blocks and also aggregated over blocks; displayed in Figure 12).

Specifically, in the SYMMETRIC VISUAL condition, the proportion of choosing the agent with fewer occluders averaged 0.694 ($SD = 0.268$), while in the ASYMMETRIC INFERENTIAL condition, it averaged 0.708 ($SD = 0.279$). These results suggest a general bias towards selecting agents with fewer occluders, which persisted across both examined conditions, with the effect of this bias not significantly different in them.

[16] These analyses were not preregistered, as this alternative explanation emerged post-preregistration.

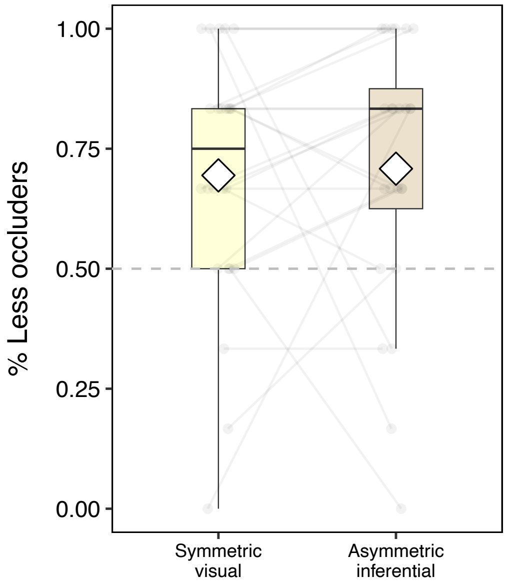

Figure 12. Boxplots of aggregate mean proportion of choosing an agent with less occluders in front of him in Experiment 6 (N = 24). White diamonds indicate means, bold horizontal lines indicate medians, boxes indicate interquartile ranges, and linked dots indicate individual data points across conditions.

Third, to further explore whether the observed bias towards agents facing fewer occluders influenced performance in the ASYMMETRIC INFERENTIAL condition, a beta regression analysis was performed. This analysis examined whether the tendency to select agents with fewer occluders in the SYMMETRIC VISUAL condition predicted correct responses in the ASYMMETRIC INFERENTIAL condition.

The beta regression model, which accounted for individual and block-level variability, did not show a significant effect of the proportion of selecting agents with fewer occluders in the SYMMETRIC VISUAL condition on performance in the ASYMMETRIC INFERENTIAL condition (Intercept = 0.381, SE = 0.388, z = 1.002, p = 0.316; β = 0.4366, SE = 0.481, z = 0.907, p = 0.364). Variability was minimal at the block level, but more pronounced among participants (σ^2 = 0.112), consistent with the participant-specific differences observed in the GLMM.

These findings indicate that while a general bias towards choosing agents with fewer occluders existed, this bias did not significantly predict the ability of participants to correctly identify the knowledgeable agent in the ASYMMETRIC INFERENTIAL condition.

Finally, we checked whether participants showed a side bias in their selection of choosing the left or right agent. The GLMM analysis showed a marginal trend towards choosing the right agent (Intercept = -0.207, SE = 0.109, z = -1.895, p = 0.0581). However, there was no evidence that side selection was associated with the accuracy of participants' responses (β = 0.079, SE = 0.322, z = 0.246, p = 0.806). This suggests that while there might have been a slight preference for selecting the agent on the right side, this did not confound response accuracy.

5.2.5 DISCUSSION

The objective of Experiment 6 was to address the limitations of Experiment 5 by developing a novel paradigm to test preschoolers' ability to attribute disjunctive-inference-based epistemic states to others. We designed a tablet-based interactive game requiring nonverbal responses, aiming to evaluate if preschoolers could identify the knowledgeable agent based on inferential reasoning.

Our results indicated that children performed significantly above chance in the ASYMMETRIC INFERENTIAL ACCESS condition, where the knowledgeable agent could only infer the animal’s identity through elimination but could not see it directly. This finding supports the hypothesis that 4-year-olds can attribute knowledge based on disjunctive inference. Additionally, the comparison with the ASYMMETRIC VISUAL ACCESS condition, where the knowledgeable agent had direct visual access, showed that children were more accurate when the agent's knowledge was based on perception. This difference suggests that while children can reason about inferential knowledge, they still exhibit higher confidence in perceptual knowledge.

Importantly, our analyses accounted for the potential bias of preferring to choose agents with fewer occluders in front of them, ensuring that the observed performance was not driven by the simple heuristic of assuming more knowledge from those who can see more. While this analysis suggests that a generic bias to choose agents with fewer occluded boxes (or conversely, more visible box contents) cannot explain our results, it is possible that participants would use this heuristic strategy selectively when faced with the ASYMMETRIC VISUAL ACCESS condition. Future studies should control for this aspect more rigorously for stronger results.

5.3 General discussion

The representational schemes hypothesized earlier—modal and open-ended—provide distinct means for how uncertainty is attributed. The modal scheme involves representing uncertainty as a set of possibilities (e.g., the ball could be in location A, B, or C), while the open-ended scheme represents it as an attitude towards an open proposition (e.g., the ball is somewhere). The aim of Experiments 5 and 6 was to explore whether and at what age children attribute modal representational content to uncertain agents by testing their ability to infer an agent's knowledge that is based on the elimination of possibilities. In both experiments, we investigated whether participants could deduce which proposition an observed agent takes to be true from three possibilities (concerning locations or identities) when she is given evidence about the falsity of all but one of these possibilities. Our hypothesis was that if participants can use a modal representation scheme to represent others uncertainty, they should attribute certainty to the agent when only one possibility remains plausible, by disjunction.

While the null results from Experiment 5 left it unclear whether 14-month-old infants struggle with applying the modal representational scheme or with updating already attributed modal contents, Experiment 6 successfully demonstrated that preschoolers can attribute knowledge based on disjunctive inference, indicating that children at this age possess a modal representation scheme. Future studies should aim to bridge the gap between the ages tested in Experiments 5 and 6, employing paradigms that are sensitive to the cognitive abilities of younger children while retaining the nonverbal response format effective with preschoolers. Given that the paradigm developed for Experiment 6 did not require verbal responses, it could potentially be adapted to suit 3- or even 2-year-olds to further explore the development of this representational capacity.

Chapter 6. **Evaluation of expected information gain**

As discussed in earlier chapters, to arrive at accurate epistemic goal attributions, it is useful for individuals to presume that people collect information in a cost-effective manner, weighing the costs against the potential benefits of the information gained. Most normative models suggest this can be achieved by balancing expected information gains with the costs associated with gathering that information (Chater et al., 1998; Coenen et al., 2018; Nelson, 2008). Although findings from Experiment 4 (Chapter 4) suggest that infants anticipate others to pursue more informative rather than less informative inquiries at equal costs, there are at least two limitations of those results in this respect. First, they could simply reflect infants' expectation that others aim to maximize the likelihood of perceiving their target of interest. Second, in contrast to the scenario presented in that experiment, in real life, the objects of human interest are often not immediately perceptible, limiting inquiries about them to assessing states that are informative only indirectly. However, our current understanding is limited concerning whether infants are capable of evaluating and making inquiries to maximize information about such latent variables.

In recent years, infants' active search for information has received a renewed focus, with research exploring infants' role in their knowledge construction (Aguirre et al., 2022; Begus & Bonawitz, 2020; Begus & Southgate, 2012; Goupil, Romand-Monnier, & Kouider, 2016; Kidd, Piantadosi, & Aslin, 2012, 2014; Kim, Sodian, & Proust, 2020; Kovács et al., 2014; Poli et al., 2020, 2023; Raz & Saxe, 2020; Twomey & Westermann, 2018). A prominent line of these studies focus on processes that canalize infants' learning by triggering information seeking when learning opportunities have been detected. Sensitivity to learning opportunities can be assumed to stem from computations about how particular observed states relate to their prior representations. Considerable progress has been made regarding the nature of this relation in the case of infants. Kidd, Piantadosi, and Aslin (2012) have shown that when 7- to 8-month-old infants are presented with probabilistic sequences of images, their likelihood of looking away from the screen is highest to events of either very low or very high surprise, suggesting that events generating intermediate levels of surprise (negative log-likelihood) are perceived preferentially. A clear appeal of such a strategy is that it prevents learners from wasting resources on overly predictable and unpredictable events. Moreover, similar results were found in the auditory domain (Kidd, Piantadosi, & Aslin, 2012), with young children (Cubit et al., 2021), and most recently, with macaques (Wu et al., 2022), indicating the generality of this phenomenon.

Another proposal emphasizes how learning changes over time (Kaplan & Oudeyer, 2007; Ten et al., 2020). The suggestion is that agents should act to optimize learning progress, not only avoiding predictable and unpredictable situations but focusing on those opportunities that maximize prediction error reduction, thus circumventing unlearnable problems. Poli et al. (2020) hypothesized that in the case of infants, such maximization of learning progress is may be achieved by modulating attention towards stimuli according to the information gain they previously led to. In this study, a set of probabilistic cue-target sequences were displayed to 8-month-olds, where first, a cue appeared in the middle of the screen, then a target reappeared in one of four quadrants around the cue location. The strongest predictor of infants' probability of looking away was the experienced information gain offered by the stimulus (infants were more likely to look away with the reduction of information gain); moreover, infants spent more time looking at a stimulus with higher information gain.

Based on these results, Altmann, Bazhydai, and Westermann (2021) have proposed that infant active learning and curiosity follow the following process: first, learners randomly choose stimuli to explore until encountering a familiar entity (for instance, a second stimulus from a previously encountered category). This stimulus will then be investigated as long as the learning progress stays above a given "boredom threshold", then, they will return to random exploration. This account points to the more general issue of the factors behind query selection (Liquin, Callaway, & Lombrozo, 2020, 2021) in early active learning. In the reviewed studies, the active selection component exhausted in paying or not paying further attention to *already* encountered stimuli. The active selection of new queries in infant exploration is often left unexplained, presumed to be entirely random (as in the theory of Altmann et al., 2021), or based on learned, backward-looking associations between internal states of uncertainty and action routines, reinforced by previously experienced information gain as a reward (Carruthers, 2018, 2019, 2024; Carruthers & Williams, 2022; Perner, 2012).

Nevertheless, targeting stimuli in a forward-looking, prospective manner, in proportion to *expected* information gain given the current model of the environment, could provide a much more efficient means of learning than random or habitual exploration, and there are developmental findings that are not straightforward to explain without it. For example, in the study of Stahl & Feigenson (2015), 11-month-old infants, when shown events that violated core conceptions of support and solidity (e.g., a car appeared to roll over gaps without falling or passed through walls), explored in ways specific to the violated aspects of naïve physics, not randomly (they banged objects that violated expectations of solidity and dropped objects that violated expectations of support). Also, when 11-month-old infants were exposed to adults providing information about novel objects, their EEG oscillatory activity in the theta frequency range was modulated by the informative potential of the adult: it was higher in anticipation of a communicative utterance from a previously informative than from an uninformative adult (Begus, Gliga, & Southgate, 2016). Similarly, recent findings show that 8-month-olds exhibit different pupil dilation responses to visual stimuli that they have learned to associate with informative versus uninformative cues (Ghilardi et al., 2024).

Although these results make the proposal of random selection unlikely, one might account for them with model-free learning of associations between actions and information (e.g., infants might have learned previously that banging provides information rewards when there is ambiguity about solidity). However, it is debatable whether directing exploration only through such pre-learned habits scales up as an efficient solution to infants' real-world learning situations, who are constantly faced with novel domains of uncertainty. A more beneficial approach would be to proactively target inputs anticipated to be informative in the current situation.

Accounts favoring habitual query selection may be motivated by the concern that prospective, forward-looking selection mechanisms would require metarepresentations, as infants would need to represent a future mental state to evaluate it according to its expected informativity. That is, to evaluate the representation of some variable (e.g., the temperature) as more or less informative in relation to a particular question (e.g., what to wear today) the mental state of representing that variable should be re-represented – at least intuitively. However, it is unclear whether a metarepresentational capacity is indeed required to evaluate information prospectively. According to normative formulations (Lindley, 1956), computations of expected information gain necessitate access to different variables' marginal and conditional probability distributions. The amount of information an observable variable Y could provide about a target variable X depends on the average information gain over each possible value of Y about X. The information gain of one such possible value Yi is given by the difference between the uncertainty (entropy) of the marginal distribution P(X) and the uncertainty of the conditional distribution $P(X|Y_i)$. In this picture, there is no reference to future or present mental states. Instead, algorithmic-level solutions would require access to knowledge of environmental statistics, which infants learn about early on (Aslin, 2017; Saffran & Kirkham, 2018).

In summary, despite the ongoing research into early exploratory behavior, the developmental origins of our ability to proactively search for relevant information are unclear. Although there is a growing body of research on how infants' observational behavior may be sensitive to the information they gained in the past, it is seldom addressed – possibly due to overestimating its (meta)representational demands – whether their behavior is also directed by forward-looking evaluation of information.

6.1 **Experiment 7**

We developed a paradigm to investigate 12-month-olds' ability to actively use information sources by controlling on-screen events with their gaze. Our approach was based on two assumptions. First, that infants can learn statistical contingencies, as ample evidence indicates (Aslin, 2017; Saffran & Kirkham, 2018). Second, that infants are motivated to obtain advance information about future events and latent parameters. The inclination to gain information as soon as possible, even if it is costly, has been studied extensively in non-human animals and human adults. For instance, if the delay between choice and reward is sufficiently long, even pigeons prefer less rewarding targets (e.g., 20% chance of reward) immediately followed by an uncertainty-resolving cue over more rewarding (e.g., 50% chance of reward) targets without such cues (Stagner & Zentall, 2010). In some situations, human adults are also willing to pay for advance information on the likelihood that their decision is suboptimal (Tversky & Edwards, 1966; Eliaz & Schotter, 2010; Bennett et al., 2016).

Based on these assumptions, we hypothesized that when uncertain about future events and unobservable variables, infants are motivated to actively seek out cues they judge to be informative about them relative to their current knowledge. To investigate this, we designed a paradigm where infants controlled on-screen events through their gaze. The experiment involved presenting infants with a set of three boxes, one of which contained a hidden character that would be revealed at the end of each trial. By fixating on laterally positioned buttons, infants could "shake" two different, partially overlapping pairs of boxes, thereby gathering auditory cues about these boxes' possible contents before they were revealed.

The experiment was divided into familiarization, training, and test phases. In the familiarization phase, infants passively observed animations where boxes containing objects produced sounds when shaken, while empty boxes did not. This phase aimed to teach the infants the contingency between the shaking action and the presence of an object within a box.

Following the familiarization phase, infants participated in 18 training trials where they were presented with the opportunity to 'shake' two out of three boxes by directing their gaze towards one of two lateral buttons, each linked to a specific pair of boxes. This setup enabled the infants to use auditory feedback (the presence of shaking sounds) from the boxes to revise their hypotheses about the location of the character. Each button could shake two boxes simultaneously. One button would shake boxes A and B together, while the other would shake boxes A and C. If infants heard a shaking sound, they could deduce that one of the two shaken boxes contained the hidden character, but not which one specifically. If no sound was heard, they might infer by elimination that the character must be in the remaining unshaken box. During the training phase, the character could be in any box, with an equal probability of ⅓ for each (A: ⅓, B: ⅓, C: ⅓). We assumed that by the end of the training phase, infants would have approximated this probability distribution.

In the test phase, at the beginning of each trial, one box was revealed to be empty, suddenly altering the probabilities of the character's location in the remaining boxes. When either box B or C was shown to be empty, the probabilities shifted such that shaking one box pair was more informative about the character's actual location (see Figure 13). Specifically, when box B was excluded, shaking boxes A and B provided more information, while when box C was revealed to be empty, checking the content of boxes A and C was more informative. However, if box A was the excluded one, this asymmetry in information gain was absent.

That is, in the first two cases, the lateral buttons, previously offering equal information, now provided different degrees of expected information gain, challenging infants to adapt their decision-making strategies based on (1) the previously learned probability distribution of the boxes, (2) the learned contingency between buttons and box pairs, and (3) the new information about the empty box. We hypothesized that if infants are motivated to maximize their knowledge about the hidden character's location even before it appears, and are capable of evaluating expected information gain, then they would strategically choose to fixate on the button that offered more information gain in the test trial, even though both buttons had been equally informative in the past, during the training phase.

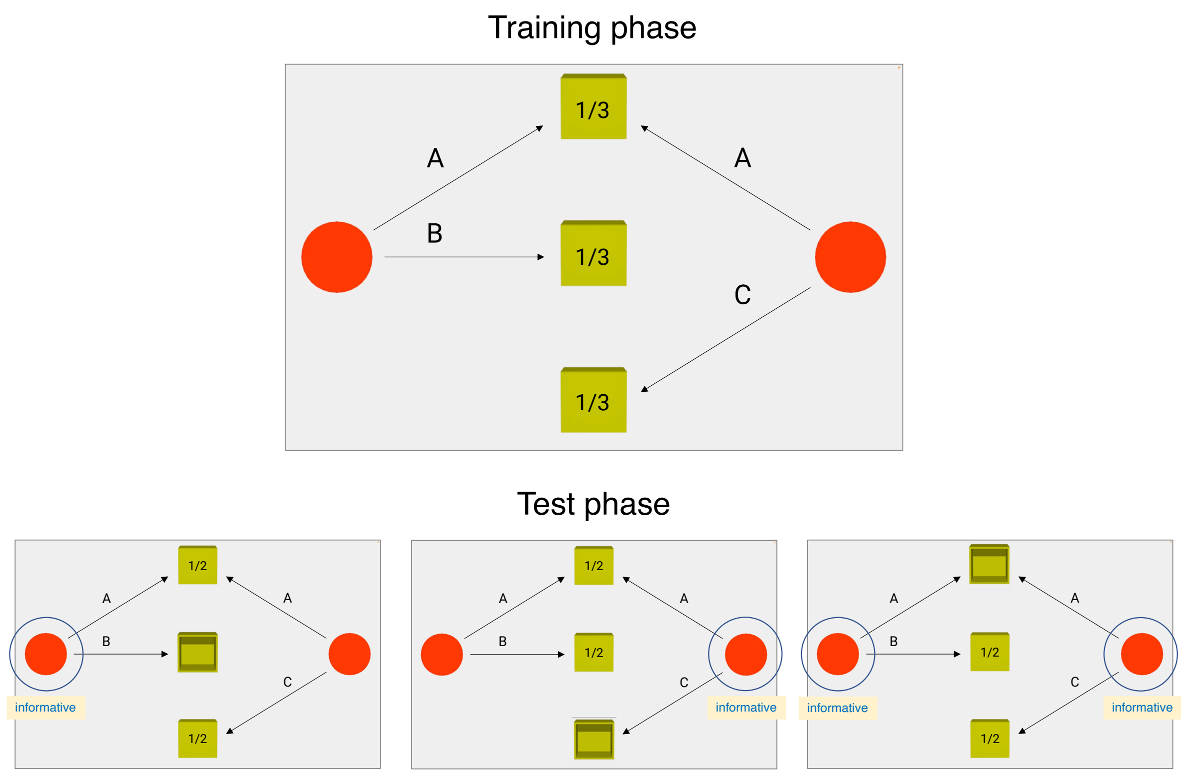

Figure 13. The logic of Experiment 7. Each button triggered the simultaneous shaking of two boxes: one button shook boxes A and B, while the other shook boxes A and C. If a shaking sound was heard, infants could infer that the character was in one of the two shaken boxes, though not precisely which one. Conversely, the absence of sound suggested that the character was in the remaining unshaken box. During the **training** phase, the character's location was equally probable across all three boxes (A: ⅓, B: ⅓, C: ⅓). By the **test** phase, the probabilities were altered at the beginning of each trial by revealing one box as empty. When box B or C was shown to be empty, the probability distribution shifted, making the shaking of specific box pairs more informative about the character's location. Shaking boxes A and B was more informative when box B was excluded (*left*), and shaking boxes A and C was more informative when box C was excluded (*center*). This asymmetry in information gain did not occur when box A was the excluded one (*right*).

6.1.1 Participants

A total of 50 full-term, typically developing 12-month-old infants were tested in the experiment. All infants were recruited in Budapest for research purposes, based on birth records. 10 participants were excluded from the final analysis due to fussiness, and 4 more due to contributing less than 66 of the training trials (trials with more than 50% of potential valid gaze samples during the decision phase were considered valid). Thus, 36 infants' data were included in the analysis (N = 36; 20 girls, M_{age} = 12 months, 12.44 days; SD_{age} = 7.72 days). Our counterbalancing scheme required at least 12 participants for the experiment. When setting the sample size, we reasoned that the sample size of 36 is within the order of magnitude that previous studies of infants' information-sensitive looking behavior have used (e.g., Kidd, Piantadosi, & Aslin, 2012; Poli et al., 2020).

The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB), Budapest, Hungary (reference: 2019/13), and conducted in accordance with the principles defined in the Declaration of Helsinki. Prior to their participation in the research, all caregivers provided written informed consent.

6.1.2 Apparatus and stimuli

Animations were created and rendered in Maya®, a 3D computer animation software (Autodesk Inc.). The resulting stimuli were displayed on the inbuilt monitor of a Tobii T60 XL eye tracker device (51.1 x 32.3 cm, 1920 x 1200 px), which recorded participants' gaze and pupil diameter. The experiment was controlled by a custom script written in Python 3.8 (Van Rossum & Drake, 2009). The stimuli, experiment control scripts, and data files are available at https://osf.io/uf2bn/ and the preregistration at https://doi.org/10.17605/OSF.IO/A639V.[17]

[17] The preregistered design included two additional "nonresponsive" trials at the end of both Experiment 7 and 8, where fixating on the buttons did not trigger any event. These were intended to assess whether infants had learned the contingency between fixating on the buttons and the shaking events. However, very few participants completed all 24 trials to reach these nonresponsive trials before fussing out, resulting in insufficient data for meaningful analysis. Consequently, these trials were not analyzed. Moreover, although exploratory analyses of pupil size were preregistered for exploratory purposes, these analyses were not conducted because the design did not include a reliable baseline period for measuring pupil dilation.

6.1.3 DESIGN AND PROCEDURE

The experiment took place in a soundproof room with dimmed lights. Participants were seated on their caregiver's lap, about 60 cm from the display. Caregivers wore opaque glasses that prevented them from seeing the stimuli. They were instructed to keep the child seated and not to interact with them. The experimenter was sitting behind a curtain, controlling the experiment, and monitoring infants' behavior via a video camera. Before the experiment, participants underwent a 5-point calibration procedure. The entire session was video-recorded. The experiment consisted of a *familiarization* phase, followed by 18 *training* and 6 *test* trials. I first describe the *training* trials' general structure to clarify our paradigm.

In each trial, infants were shown three boxes arranged vertically (box A on top, box B in the middle, and box C on the bottom) placed at the center of a screen. One of the boxes always contained a child-friendly, animated character (a blue elephant, a yellow duck, a green bug, a brown hamster, a toy plane, or a red spinning top, each making its own sound and characteristic movement when appearing). At the beginning of the trial, the boxes were closed, so infants could not know which boxes were empty and which one hid the character.

In the decision phase, before the contents of all three boxes were revealed, infants had maximum 10 seconds to shake two out of the three boxes by directing their gaze to one of two salient 'buttons' (red discs) positioned laterally on the two sides of the screen. Fixating one of the buttons allowed infants to shake boxes A and B simultaneously, while fixating the other allowed them to shake boxes A and C at the same time. That is, both buttons shook box A, but shaking boxes B and C was exclusive to only one of the buttons. This relationship between buttons and boxes was highlighted by an animation that followed subjects' fixation on the buttons: two lines ("arms") spread out of the buttons to the corresponding boxes, appearing to shake the boxes by physically being connected to them. The gaze-contingent nature of the events was emphasized by providing infants visual feedback on their gaze in this decision phase: a small blue dot (cursor) was displayed to follow participants' gaze with minimal delay (this delay also allowed us to smooth the cursor movements by averaging its x-y positions in the last 150 ms).

Shaking a box that contained a character produced a sound. Therefore, after shaking the boxes and hearing (or not hearing) a sound, infants' belief about the character's location in the current trial could be revised through inference. Crucially, because the shakes of the two boxes happened at the same time, if there was a shaking sound, subjects could not know which of the two shaken boxes contained an object, but if they emitted no sound, they could infer, by disjunction, that it is in the third, unshaken box (assuming that there is a character hidden in each trial and only one of the boxes can contain a character).

The overall temporal structure of training trials (Figure 14) was the following: (1) decision phase, maximum 10 secs (participants could start using the buttons after the first 0.5 sec and must have fixated on them for at least 0.3 s to initiate shaking; we set these timings to lower the probability of accidental triggering of shaking due to fixating at buttons inadvertently); (2) shaking event, in which two boxes were shaken (if chosen), 3.5 secs; (3) pause, 0.5 sec; (4) box opening event, 4 secs (when all boxes when opened and the hidden character appeared); (5) an inter-trial-interval (0.5 sec). In test trials, a box exclusion event (4 secs) was added before the decision phase, revealing one of the boxes to be empty.

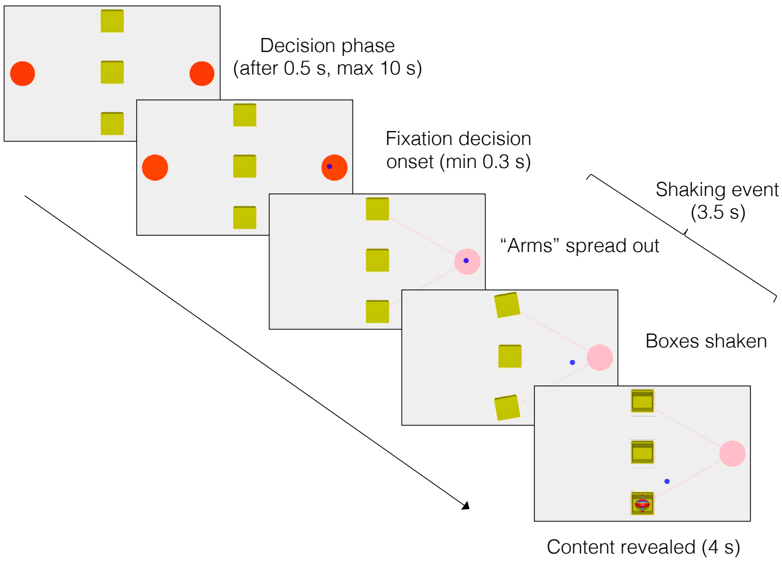

Figure 14. The structure of familiarization trials in Experiment 7.

In the *training* trials, all boxes could potentially contain characters. The actual probability distribution of character locations was uniform (A: ⅓, B: ⅓, C: ⅓). However, in the trials of the *test* phase, one of the boxes was always first revealed to be empty, leaving only two boxes as possible hiding locations. When box B or box C were eliminated, the probabilities of character locations changed such that shaking one of the box pairs would suddenly become more informative regarding the actual location of the character (if box A was eliminated, this difference did not arise). That is, if box B was eliminated, that button was more informative that shook boxes A and B. The temporal structure of the test and training trials was the same, except that test trials started with a 4-seconds-long elimination phase, where one of the empty boxes opened. During the elimination phase, the buttons were not displayed yet.

In addition to the *training* and *test* phases, the experiment started with a *familiarization* phase where we aimed to teach infants the relationship between shaking sounds (or the absence thereof) and whether the boxes are full (or empty). This phase was not gaze-contingent, rather, infants passively observed animations with the following events: (1) a single, open box containing a blue ball was shaken, such that the ball also moved within the box, colliding with its walls, accompanied by shaking sounds contingent with the collisions; (2) the same box was closed and shaken again, accompanied by the same sounds; (3) the box opened again, revealing the ball inside; (4) the previous box disappeared, and another, empty box appeared, then the events (1-2) were repeated, but with this empty box, without any sounds (as there was no ball, and therefore no collision to produce them). This familiarization was 20 seconds long, and it was played twice.

The following variables were counterbalanced between subjects: 2 possible mappings between circles and box pairs (left button shakes boxes A and B, and right button shakes boxes A and C, or vice versa), 6 possible box locations (box A: top, box B: center, box C: bottom, etc.).

Since the experiment was dependent on infants' response, we aimed to make sure that by the time they arrived at the *test* phase, they would (1) have shaken boxes that were full and empty approximately equally often, (2) have shaken full boxes with both buttons approximately equally often. For this reason, at the beginning of each trial in the *training* phase, the control script determined which box would store the object based on a set of constraints and conditions. First, we defined two kinds of rates. *Reinforcement rates* corresponded to each button and quantified, based on the previous trials, the proportion of times when the given button was fixated on and shook the full box (the box with the character in it). *Evidence rates* quantified the overall proportion of shaking empty boxes (negative evidence) versus filled boxes (positive evidence) in the previous trials. Finally, the hiding locations were drawn from a set of locations that contained 6 instances of each of the three locations.

Our control script first checked whether the reinforcement rate of button B was higher or lower than the rate of button A. If the rate of one button was higher than the other, the script aimed to counterbalance it by prioritizing an unshaken location for the high-rate button and a shaken location for the low-rate button. If the difference between the reinforcement rates of the two buttons was within the threshold of 0.2, the script checked the evidence rate, specifically, whether there was more positive or negative evidence during the trials so far (above the threshold of 0.2). If there was more positive evidence, the script prioritized unshaken locations for both buttons, and in case of more negative evidence, it prioritized shaken locations. The script used all available boxes as potential hiding locations if the difference between reinforcement rates and the evidence rate were outside their respective thresholds. The prioritization process involved checking whether preferred locations were available in the list and whether they would be third repetitions. If there were valid locations that passed these criteria, one of them was randomly chosen. If none of the preferred locations passed the criteria, the remaining locations were checked for availability and repetition.

The areas of interest (AOIs) for triggering events on the display were defined based on specific pixel coordinates. The two lateral buttons spanned a horizontal range of 420 pixels each, with the right button AOI covering the area from 540 to 960 pixels right from the center, and the left button AOI covering from -960 to -540 pixels. Both button AOIs also spanned a vertical range of 400 pixels, extending 200 pixels above and 200 pixels below the center of the screen. The boxes' AOIs each occupied a 400 × 400 pixel area. The upper box AOI spanned from -200 to 200 pixels horizontally and from 200 to 600 pixels vertically. The center box AOI covered the same horizontal range but spanned from -200 to 200 pixels vertically. Similarly, the lower box AOI also covered the -200 to 200 pixel horizontal range, but extended from -600 to -200 pixels vertically.

6.1.4 RESULTS

We included trials with more than 50% valid gaze samples during the decision and the shaking phase (8% of the training trials and 29% of the test trials were excluded) in the final analysis. Participants were included in the final sample if they contributed at least 2/3 (12 out of 18) of the training trials (training trials $M_{training} = 16.47$, $SD_{training} = 1.68$; test trials $M_{test} = 4.75$,

$SD_{test} = 1.36$). On average, infants provided fixation decisions on 91.4% of training trials (SD = 0.10) and 88.7% of test trials (SD = 0.18).

We were interested whether participants actually learned that looking at the buttons triggered boxes' shaking movements. For this reason, we coded when participants first looked at the boxes after making a decision between the two buttons. Since the actual movement (shaking) of the boxes only started 750 ms after the gaze decision, if the calculated reaction time was below this threshold, we coded the look as anticipatory. As a possible signature of acquiring this contingency, we applied a Generalized Linear Mixed Model (GLMM) to see whether the likelihood of such anticipatory looks increased as a function of time (trial number). The model used a binomial distribution with a logit link function to predict the binary outcome of whether a look occurred before the 750 ms threshold and included random intercepts for participants. The intercept was significant (Intercept = −3.4293, SE = 0.4228, z = −8.111z, $p < 0.0001$), indicating that at the beginning of the experiment (i.e., at trial 1), the probability of making an anticipatory look was very low (3.13%). The fixed effect of trial number was found to be statistically significant ($\beta = 0.1016$, SE = 0.0287, z = 3.539, p = 0.000402), indicating that as the trial number increased, participants were more likely to make

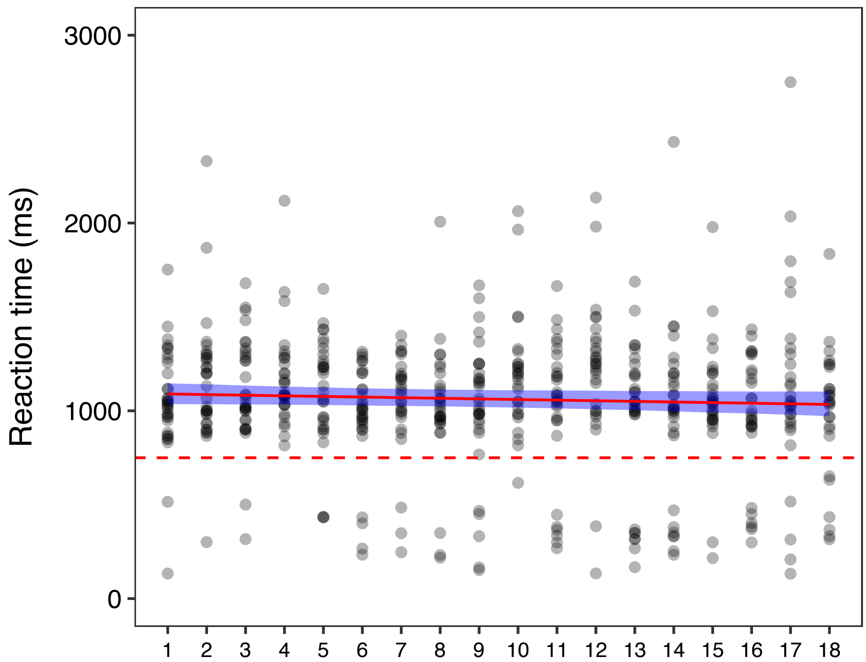

Figure 15. The timing of first looks at the boxes after decision onset throughout the training phase of Experiment 7. Points represent subjects' reaction times, the red line indicates the model-predicted reaction time for each trial, the blue-shaded region captures the 95% confidence interval. The red dashed line at 750 ms indicates the anticipation threshold.

anticipatory looks. Specifically, the odds of making an anticipatory look increased by approximately 10.7% with each successive trial (Figure 15). This suggests that over time, participants learned to anticipate the movement of the boxes more accurately.

As the second indicator of learning, we analyzed the relationship between trial number and fixation decision onset in training trials using a linear mixed model (Figure 16). To address the skewness in the reaction time data, the fixation decision onset times were log-transformed. The model included trial number as a fixed effect and random intercepts and slopes for participants. The fixed effect of trial number on the log-transformed choice time was significant (β = 0.0173, SE = 0.0062, df = 88.99, t = 2.761, p = 0.007), indicating that, on average, the log-transformed fixation decision onset increased with each successive trial. The significant positive effect suggests that as the experiment progressed, participants took slightly longer to make their fixation decisions, which could indicate both increased deliberation and decreased engagement over time.

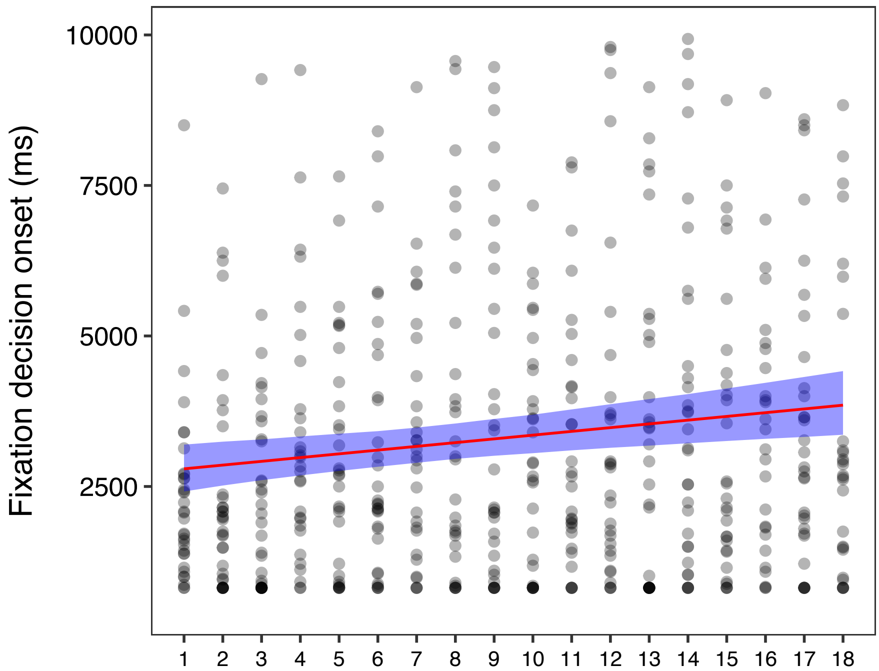

Figure 16. Fixation decision onsets throughout the training phase of Experiment 7. Points represent subjects' decision onset, the red line indicates the model-predicted choice onset for each trial, the blue-shaded region captures the 95% confidence interval.

To check our central hypothesis, whether infants looked more at the informative or uninformative buttons at test trials, we fit a GLMM with a binomial distribution and logit link function to assess the probability of choosing the informative button. The model included random intercepts and random slopes for trial number at the participant level, allowing for individual differences in the overall likelihood of choosing the informative button and in how this likelihood changed across trials. In this analysis, we only considered trials where one of the buttons provided more information than the other. The model yielded a significant intercept of -0.5349 (SE = 0.1967, z = -2.72, p = 0.00653), indicating that the log-odds of success were negative. This corresponds to an estimated probability of choosing the informative button at approximately 36.9%, significantly lower than the 50% chance level. That is, in contrast to our hypothesis, the average of these scores indicated that choosing the *uninformative* button was more likely than choosing the informative one (M = 0.37, SD = 0.27; see Figure 17).

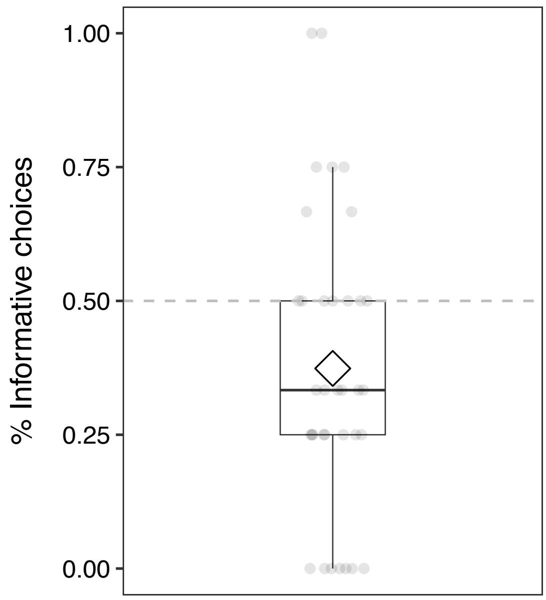

Figure 17. Boxplot of aggregate mean proportion of choosing the informative versus the uninformative button in Experiment 7 (N = 36). The white diamond indicates the mean, the bold horizontal line indicates the median, the box indicates the interquartile range, dots indicate individual aggregated proportions, the dashed line indicates chance performance.

As a possible indicator of the different cognitive demands driving these decisions, we were also interested in whether fixation decision onset time differed between informative and uninformative decisions. To explore this, we used a linear mixed model with decision time as a dependent variable and decision type as the fixed predictor, with random intercepts and random slopes for trial number at the participant level. The decision onset times were log-transformed to address skewness and meet the assumptions of the model. The results indicated that, on average, informative decisions were associated with shorter fixation decision onset times compared to unsuccessful trials (β = -0.3159, SD = 0.1495, t = -2.114). This difference was statistically significant (p = 0.0375), suggesting that participants made informative decisions quicker (M = 2607.34 ms, SD = 1942.68 ms) than noninformative decisions (M = 3798.82 ms, SD = 2396.167 ms).

We also examined whether participants' decision times differed between trials where both decision options provided information ("balanced" trials) and those where only one option was informative ("unbalanced" trials). A linear mixed model was fitted with the log-transformed decision time as the dependent variable and trial type (balanced vs. unbalanced) as the predictor, including with random intercepts and random slopes for trial number at the participant level. The results indicated a slight increase in log-transformed decision time in unbalanced trials by $\beta = 0.1315$ (SD = 0.1159), though this difference was not statistically significant ($t = 1.108$, $p = 0.27$). This suggests that decision times did not significantly differ between balanced (M = 3043.68 ms, SD = 1991.69 ms) and unbalanced (M = 3422.13 ms, SD = 1946.43 ms) trials.

Finally, to check whether there was a side bias in participants' fixation decisions during the trials, a GLMM with a binomial distribution and logit link function was employed, considering the binary outcome of choosing a particular side as the dependent variable (the model included random intercepts and random slopes for trial number at the participant level). The fixed effect intercept, representing the overall log-odds of choosing one side over the other, was estimated at $\beta = -0.0923$ (SE = 0.1119, $z = -0.825$, $p = 0.409$) indicating no bias toward any side.

6.1.5 DISCUSSION

In order to investigate 12-month-olds’ ability to evaluate expected information gain, we developed a gaze-contingent paradigm that allowed infants to gather information with their visual fixations on the display. The structure of this paradigm was designed to allow subjects to learn the probabilities of different events (the probability of a box revealing an interesting object), as well as the means to use two information sources (buttons that shook different but overlapping pairs out of the three boxes) providing advance information about these events in each trial. In the first part of the experiment, given that infants learned the uniform probability of these events, both sources provided equal amount of information to them. However, in test trials, one event was excluded as possibility (one box was revealed to be empty), swiftly changing the event probabilities and consequently also the expected information (in unbalanced test trials) that could be gained from the two sources.

We hypothesized that if infants aim to gain advanced information about upcoming events, they will use the information sources to do so, that is, they will fixate on the two lateral buttons on the screen strategically. Moreover, we presumed that if they can evaluate expected information gain, they will use the informative rather than the uninformative source in test trials, fixating on the button that was learned to shake only one of the remaining boxes, not both, thus maximizing expected information gain. The results, however, point in the opposite direction: infants were more likely to consult the uninformative sources, that is, they tended to choose the button that shook both remaining boxes. There are various possible explanations for this result.

First, infants may have acted on the learned association between the remaining pair of boxes and the button that previously shook them. That is, they may have fixated on the button they associated with the remaining pair, without any information-gathering goal in mind. Such an explanation would not even require that infants learned the contingency between their fixations on the screen and the shaking of boxes. However, we found that the likelihood of anticipating box movements after fixation onset was increasing throughout the training phase, suggesting that infants did learn about this contingency. Moreover, it is known that infants are fast learners of associations between their motor response and contingent stimuli. For instance, in the mobile conjugate reinforcement paradigm (Rovee & Rovee, 1969), 10-weeks-old infants' response rate triples within 6 minutes if their foot kicks are contingently followed by the movement of a dangled mobile in front of them, but not when these movements are independent of their kicks. Even 6- and 8-month-olds can learn to make fixations on the screen to trigger the appearance of new stimuli, and they anticipate the consequences of these fixations after only 3 trials (Wang et al., 2012). In comparison, in our experiment, 12-month-olds had 18 trials to learn about gaze-contingency until the test phase.

In addition, the result that infants were faster to choose the informative button than the noninformative one suggests that if anything, low-level visual associations played a role in influencing infants' behavior toward the other direction. That is, when a box was eliminated, infants were quicker to fixate on the button that they associated with shaking this eliminated box in the past. In contrast, when they chose the uninformative button, their decisions were slower, indicating that these choices might have been the result of a relatively more deliberative process.

A second possibility is that although infants learned that they were in control of shaking box pairs via fixating on the lateral buttons, they did not use them to gain information about the upcoming events but rather to generate a shaking sound that accompanied the shaking of the full box. In this case, their aim could have been to maximize the likelihood of shaking the full box, which could have been achieved in test trials by choosing the uninformative button which shook both remaining boxes.

Finally, it is also possible that infants learned the relevant contingencies and also used the buttons to gain information, but they were engaged in confirmatory hypothesis testing. Looking for confirmatory evidence is a broadly studied phenomenon in adults (see McKenzie, 2004 for a review). It has been argued (Lapidow & Walker, 2020b) that repeated positive testing could sometimes be a reasonable strategy to learn about the stability of a hypothesis in novel circumstances – and indeed, our test trials were novel relative to the training trials in that they included only two closed boxed and one open one, possibly cueing a changed context. In such a context, infants could have been interested whether one of the remaining boxes still contains a character, and the most effective to test this was to shake both of them.

6.2 **Experiment 8**

In Experiment 7, we aimed to separate the phase of acquiring a representation of the probability distribution generating trial outcomes from having to rely on this representation to maximize information gain in a new context. In the training phase, infants could learn that the character may be in any of three boxes with equal probability, and thus in this phase the buttons provided equal expected information gain. However, during the test phase, this probabilistic structure was suddenly altered by eliminating one box as a possible hiding place. This manipulation ensured that if infants chose the informative button, they were relying on a forward-looking measure of expected information gain rather than a backward-looking strategy, such as selecting the button that had previously provided more information.

However, this abrupt change might have led infants to perceive a shift in the underlying task structure, prompting them to engage in confirmatory hypothesis testing to verify if the previous structure still applied.

Therefore, in designing Experiment 8, we were interested in how infants' behavior would change if we eliminated this sudden shift in the task's probabilistic structure. Instead of creating asymmetry in the buttons' informativeness through a discrete exclusion phase, we allowed infants to gradually learn a skewed probability distribution of character locations. We assumed that as infants incrementally learn the underlying probability distribution of character locations, based on evidence from prior trials. Consequently, the information gain they could expect from shaking the box pairs would also change over time, as it depends on the probability distribution. Given a stationary probabilistic distribution and sufficient experience, learning about the distribution itself becomes less crucial, and infants' representation of it can be used to either maximize expected information gain about the character's location or maximize the likelihood of shaking the full box (a potential strategy that infants might have used in the previous experiment, although our paradigm, it is not possible to discriminate this from confirmatory hypothesis testing).

To formalize these assumptions, we developed an ideal observer model (detailed in the next section) capable of learning and making decisions according to these strategies within a simulated version of our paradigm. In addition to modeling infants' responses, we used the model's simulation results to inform the design of our experiment. Specifically, we used the model to generate a skewed probability distribution of character locations and a series of trial outcomes under which the predicted choices of the two main strategies would diverge the most. Throughout a simulated experiment, the ideal observer agent maintains a belief about the true underlying probability distribution of the character being in each box. The agent's belief about this distribution is updated using Bayesian inference as new locations are observed at each trial, starting with a uniform prior (assuming that the character can be in any of the three boxes with equal probability). In each trial, the agent selects which pair of boxes to shake by maximizing a utility function, which depends on its objective. The agent can have two different strategies: (a) minimizing uncertainty about the hiding location in the current trial (that is, maximizing expected information gain) or (b) maximizing the probability of shaking a box that contains the character.

Based on the ideal agent's simulated choice patterns, we derived a skewed probability distribution of character locations (A: 11/24, B: 9/24, C: 4/24) and a corresponding sequence of trial outcomes that sufficiently differentiated between these two model objectives. The simulations indicated that this distribution is learned with relatively high certainty approximately halfway through the experiment. Afterward, the agent consistently shakes one box pair (A and C) under the information gain strategy and the other box pair (A and B) under the full-box shaking strategy (see Figure 18).

Our interest was in determining which of these two strategies best captures infants' behavior when they are given the opportunity to gradually learn the skewed probability distribution of character locations, as opposed to the sudden adaptation required in the previous experiment. However, this gradual design introduces a fundamental limitation: simpler, backward-looking versions of the two hypothesized strategies could also lead to similar behavior. Specifically, since the past and future were correlated in terms of probabilistic structure, relying on how informative past choices were regarding trial outcomes (backward-looking version of strategy a) and how rewarding past choices were in terms of producing shaking sounds (backward-looking version of strategy b) could also result in similarly divergent choices in the latter part of the session, with appropriate learning rates. To determine whether forward- or backward-looking strategies best describe infants' choices, we also modeled these latter two strategies and conducted a model comparison based on the data.

6.2.1 MODEL SPECIFICATION

We utilized an ideal observer model, that we call as the "Ideal Shaker," to simulate optimal decision-making in our paradigm. The Ideal Shaker maintains and updates its beliefs about the probabilities of the character hiding in one of three boxes (A, B, and C) over a series of trials. These beliefs are modeled using a Dirichlet distribution, which represents a probability distribution over the possible sets of probabilities for the character's location. The Dirichlet distribution serves as a conjugate prior to the multinomial distribution that governs the character's actual locations, allowing the Ideal Shaker to update its beliefs as new evidence is observed. Initially, the agent's belief about the probability distribution of the character being in each box is represented by a Dirichlet distribution parameterized by $\alpha = (\alpha_A, \alpha_B, \alpha_C)$

where all $\alpha_i = 1$, indicating a uniform prior. This distribution is updated as new evidence (character locations) is observed. The Dirichlet distribution is given by:

$$P(\theta \mid \alpha) = \frac{1}{B(\alpha)} \prod_{i=1}^{3} \theta_i^{\alpha_i - 1}$$

where $\theta = (\theta_A, \theta_B, \theta_C)$ represents the probability vector of the character being in each of the three boxes, and $B(\alpha)$ is the normalization constant (the Beta function for the Dirichlet).

Each time the agent observes the character in one of the boxes, it updates the corresponding α parameter: $\alpha_i \leftarrow \alpha_i + 1$ for the observed box i. From the Dirichlet parameters, the agent can derive the approximated probabilities of the character's location in each box. The approximated probability p_i for each box i is:

$$p_i = \frac{\alpha_i}{\sum_{j=1}^{3} \alpha_j}$$

The Ideal Shaker was designed to pursue various objectives during the experiment. Under objective A, the agent aims to minimize its uncertainty about the character's location in the current trial via maximizing expected information gain (EIG). EIG is computed by comparing the entropy of the agent's currently represented distribution $\boldsymbol{p}$ of location probabilities in current tial with the expected entropy after shaking a pair of boxes (i, j):

$$\text{EIG}_{ij} = H(\boldsymbol{p}) - E\left[H\left(\boldsymbol{p} \middle| S_{ij}\right)\right]$$

where S_{ij} represents the sound outcome (whether a sound is heard or not) when shaking boxes i and j. The entropy $H(\boldsymbol{p})$ of the distribution $\boldsymbol{p}$ is given by:

$$H(\boldsymbol{p}) = -\sum_{i=1}^{3} p_i \log p_i$$

While the expected entropy $E\left[H\left(\boldsymbol{p} \middle| S_{ij}\right)\right]$ is calculated as the weighted sum of the entropies associated with the possible outcomes of shaking, where the weights are the probabilities of those outcomes occurring:

$$E\left[H\left(\boldsymbol{p} \middle| S_{ij}\right)\right] = P\left(S_{ij} = 1\right) \cdot H\left(\boldsymbol{p} \middle| S_{ij} = 1\right) + P\left(S_{ij} = 0\right) \cdot H\left(\boldsymbol{p} \middle| S_{ij} = 0\right)$$

where $P\left(S_{ij} = 1\right)$ is the probability of hearing a sound when shaking boxes i and j. $H\left(\boldsymbol{p} \middle| S_{ij} = 1\right)$ is the entropy of the updated belief distribution given that a sound was heard.

$P(S_{ij} = 0)$ is the probability of not hearing a sound when shaking boxes i and j. Finally, $H(\boldsymbol{p}|S_{ij} = 0)$ is the entropy of the updated belief distribution given that no sound was heard.

Alternatively, under objective B, the agent may focus on maximizing the immediate reward of shaking a box that contains the character (thus inducing the shaking sound and confirmatory evidence), without regard to the informational value of the action regarding the trial outcome. The expected value (EV) for this objective when shaking boxes i and j is the sum of the probabilities of these boxes containing the character:

$$\mathrm{EV}_{ij} = p_i + p_j$$

Under both objectives, the agent uses a softmax decision rule to choose which pair of boxes to shake. The utility values calculated for each pair of boxes, according to the chosen objective, are transformed into a probability distribution over actions using the softmax function:

$$P(\text{Shake } (i,j)) = \frac{\exp\left(\frac{U_{ij}}{\tau}\right)}{\exp\left(\frac{U_{AB}}{\tau}\right) + \exp\left(\frac{U_{AC}}{\tau}\right)}$$

where U_{ij} is the utility for shaking boxes i and j (either EIG_{ij} or EV_{ij} depending on the objective). In this function, the nominator is the exponentiated utility for the specific pair (i, j), while denominator is the sum of the exponentiated utilities for the two possible pairs, AB and AC. τ is a temperature parameter that controls stochasticity in choices. In our simulations, τ was set to 0.1, a relatively lower value, meaning the agent was more likely to choose the pair of boxes with the higher utility value.

Under these two objectives, the Ideal Shaker's behavior was simulated over a series of trials with the true underlying probability distribution of the character's location set to $\theta = (11/24, 9/24, 4/24)$ for boxes A, B, and C, respectively. As the agent accumulated evidence over trials, its Dirichlet parameters α were updated, refining its belief about the true distribution of the hiding location. The simulation results, as illustrated in Figures 18 and 19, demonstrated a clear divergence in the models' behavior over the course of the trials. Specifically, after an initial period of fluctuation, the difference in the probabilities assigned by the softmax function to the two choice options (ΔP) started a steady increase across trials for both objectives around the midpoint of the simulated experimental session.

Under objective A (maximizing expected information gain) the agent predominantly chose to shake boxes *AC*. Under objective B (maximizing the probability of shaking the full box), the agent preferred shaking boxes *AB*. This growing disparity indicated that as the model accumulated evidence about the distribution, it became increasingly confident in its choices, leading to more deterministic decisions. Concurrently, the change in entropy of the Dirichlet distribution, which represents the model's uncertainty about the true probability distribution, decreased steadily across trials. This decline in entropy change reflects the model's learning process, where the reduction in uncertainty about the underlying distribution corresponds with the divergence in choice probabilities.

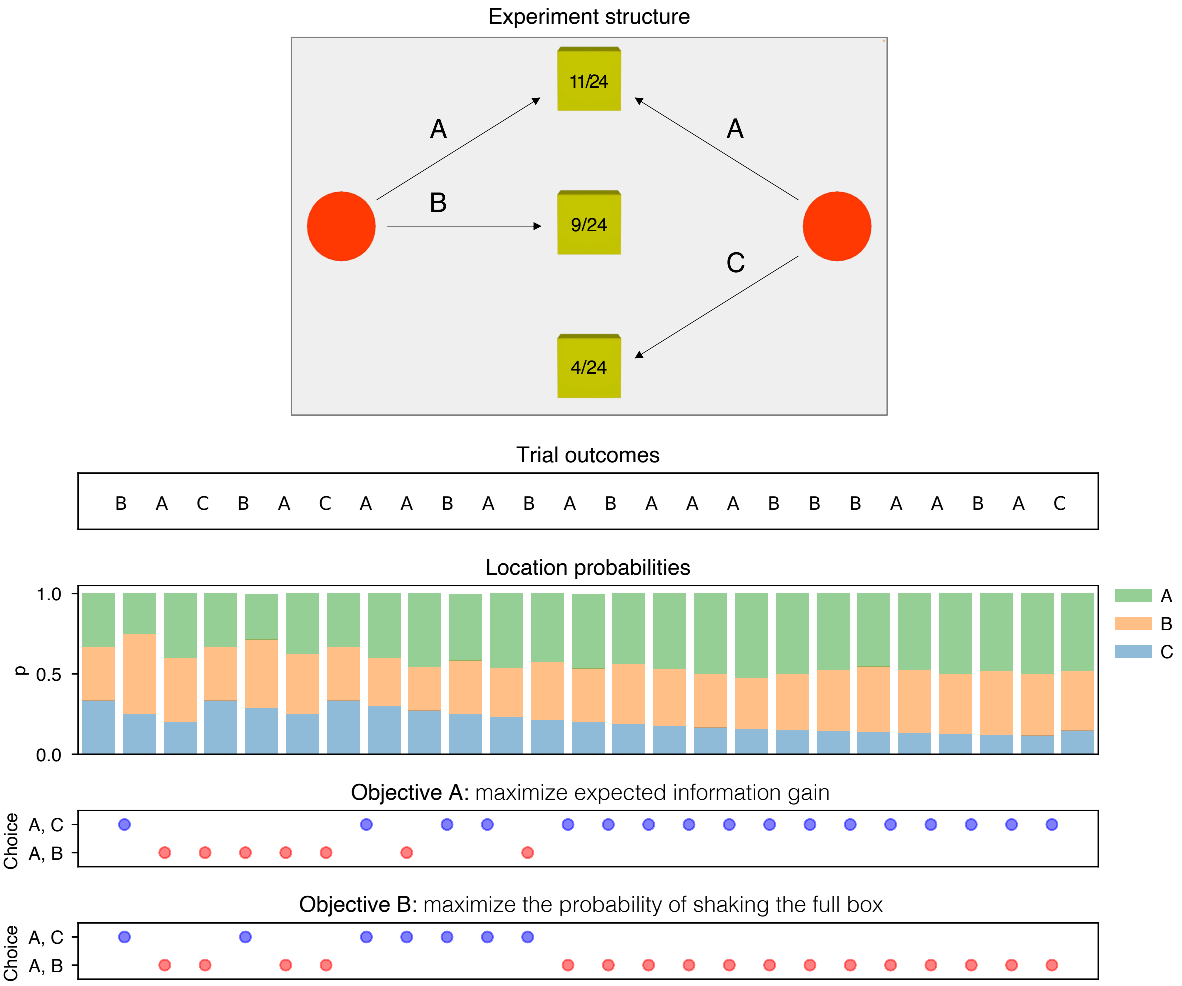

Figure 18. Simulation results of the ideal observer under the objectives A and B. As the agent observes the trial outcomes (first two panels), its representation about location probabilities changes (third panel). The agent makes choices for shaking box pairs either by maximizing expected information gain (Objective A) or by maximizing the probability of shaking a box that contains the character (Objective B). By the second part of the experiment, under Objective A, the agent predominantly chooses to shake the A and C pair, as this configuration provides the highest expected information gain. Conversely, under Objective B, the agent favors the A and B pair, prioritizing the likelihood of shaking a box with the character. For simulations, the τ parameter was fixed at 0.1.

In addition to these forward-looking objectives, we extended the model to include two additional backward-looking objectives. These objectives were implemented as simpler, model-free Q-learning policies, allowing the agent to make decisions based on the outcomes of previous choices rather than expected future information gains or approximated box probabilities.

Under objective C, the Ideal Shaker's choices are driven by the information gained from past choices. Instead of evaluating the expected information gain from potential future choices, the agent only considers the information gains achieved in previous trials when making decisions. Specifically, the agent tracks the cumulative information gain for each pair of boxes shaken in the past and uses this cumulative gain to guide current choices. The utility for this objective when shaking boxes *i* and *j* is updated incrementally as:

$$Q_{ij}(t+1) = Q_{ij}(t) + \eta\left[\mathrm{IG}_{ij}(t) - Q_{ij}(t)\right]$$

where $Q_{ij}(t)$ is the Q-value for the pair of boxes *i* and *j* at trial *t*, η is the learning rate, and $\mathrm{IG}_{\mathrm{ij}}(t)$ represents the information gain from shaking boxes *i* and *j* at trial *t*. The learning rate η controls how quickly the agent updates its Q-values in response to new information. A higher η means the agent gives more weight to the most recent trial, leading to faster updates but potentially more volatile Q-values. A lower η results in slower, more stable updates, as the agent gradually adjusts its Q-values based on cumulative experience.

Finally, under objective D, the focus is on the past success of shaking actions, specifically on the rewards obtained from shaking the boxes that contained the character. Here, the agent accumulates the rewards (shaking sounds) obtained from choosing each pair of boxes over all past trials and updates its current estimate of the utility for each action using a similar incremental Q-value update rule as in Objective C:

$$Q_{ij}(t+1) = Q_{ij}(t) + \eta\left[\mathrm{R}_{ij}(t) - Q_{ij}(t)\right]$$

where $Q_{ij}(t)$ is the Q-value for the pair of boxes *i* and *j* at trial *t*, η is the learning rate, and $\mathrm{R}_{ij}(t)$ represents the reward (whether a sound was heard or not) from shaking boxes *i* and *j* at trial *t*.

For both objectives C and D, the agent still employs the softmax decision rule to choose which pair of boxes to shake. To encourage exploration early in the learning process, we applied optimistic initialization to the Q-values. That is, we set $Q_{ij}(t = 0)$ to 1, encouraging the agent to explore different actions in the beginning rather than settling on a single option prematurely (by time, the Q-values converge to more accurate estimates, and the agent's choices become more deterministic).

Importantly, these objectives rely heavily on past choices, unlike Objectives A and B, which consider the learned probability distribution of box outcomes, irrespective of previous choices. Therefore, their characteristic choice patterns can be better investigated by running multiple simulations. Figure 20 illustrates that with an appropriate learning rate parameter (we used $\eta = 0.5$, an intermediate value) these backward-looking objectives can also converge on opposite choices in the second part of the experiment. Under objective C (choosing box pairs by maximum past information gain), similar to but less consistently than under objective A, the choice of box pair AC tends to dominate by the second half, while under objective D

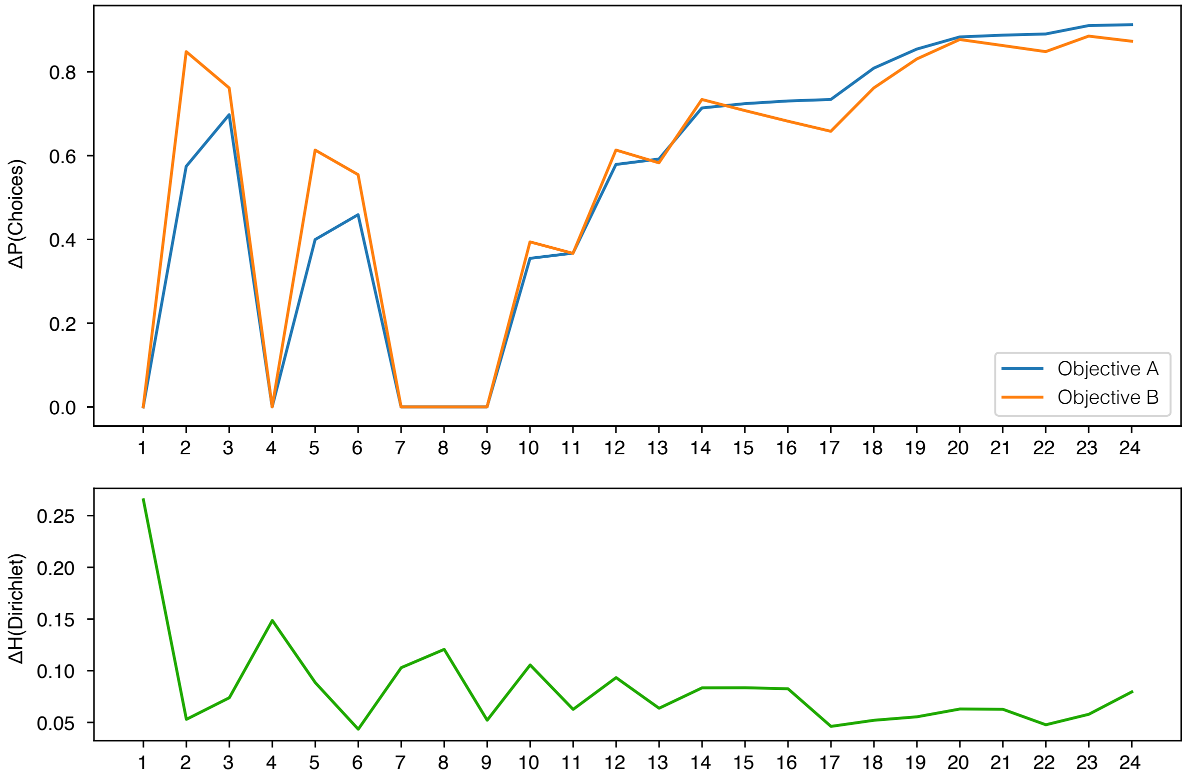

Figure 19. Simulated divergence in choice probabilities and change in entropy across trials. The top panel depicts the difference in probabilities assigned by the softmax function to the two choice options (box pairs) across trials for two objectives: maximizing expected information gain (objective A) and maximizing the probability of shaking a box containing the character (objective B). Over time, the difference between these probabilities increases, indicating that the model's choices become more deterministic as it approximates the true probability distribution of hiding location. The bottom panel shows the changes in the entropy of the Dirichlet distribution as a result of learning trial outcomes, quantifying the reduction in the model's uncertainty associated with the distribution over the probabilities of hiding locations. For simulations, the τ parameter was fixed at 0.1.

(choosing box pairs by maximum past shaking sound reward), the choice of box pair *AB* is preferred.

Notably, there is a distinct difference in how objectives A and C, the forward-looking and backward-looking information-directed strategies, respond to a sequence of box A outcomes in trials 15 to 17, following frequent box B outcomes. The agent operating under objective C, which relies solely on the experienced information gain from past choices, tends to assign a higher value to shaking box pair *AC* when previous trials frequently resulted in

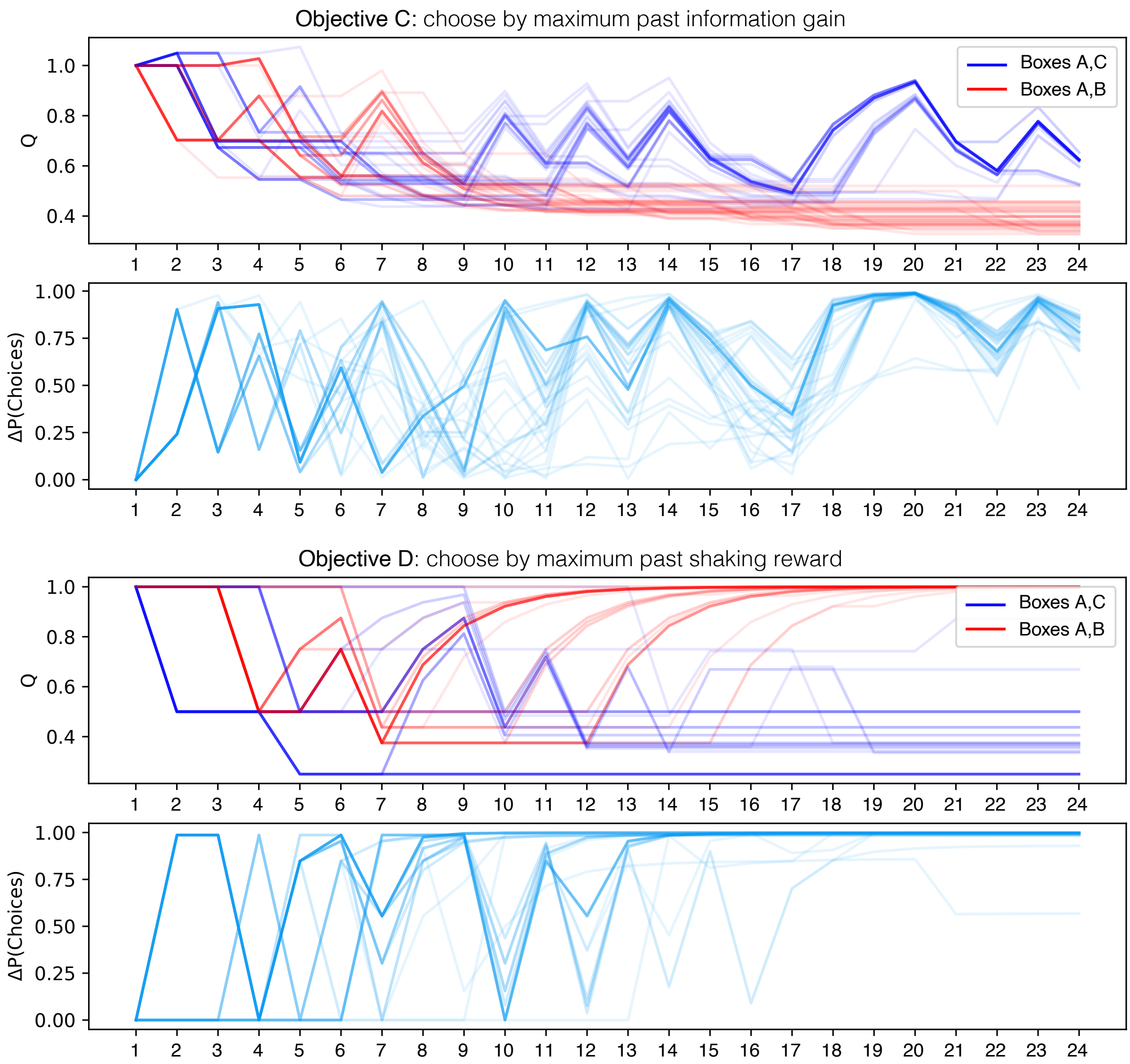

Figure 20. Simulated divergence in Q-values and choice probabilities across trials under objective C and D (simulated N = 36). The top two panels depict the Q values and the difference in choice probabilities assigned to the two options (box pairs) across trials for objective C (choosing pairs by maximum past information gain) in multiple simulation runs, while the bottom two panels depict the same for objective D (choosing pairs by maximum past shaking reward). Over time, the difference between choice probabilities tend to increase under both objectives, aligning with one of the choice options becoming more preferred (boxes AC for objective C and boxes AB for objective D). For these simulations, parameters were fixed at $\tau = 0.1$ and $\eta = 0.5$.

box B outcomes. This is because such a sequence typically provides greater information gain when the *AC* pair is chosen. However, when the agent suddenly encounters a string of trials where box A is the outcome, shaking the *AC* pair leads to less informative results. Consequently, this agent devalues this pair from trial to trial, and the probability of choosing the *AC* pair decreases, approaching chance levels. This sensitivity to recent evidence in the backward-looking model is modulated by the learning rate in the Q-value update function: a higher learning rate causes the agent to adjust more rapidly to recent outcomes, leading to quicker devaluation of the *AC* pair after uninformative results, whereas a lower learning rate would result in a more gradual adjustment.

In contrast, the agent using objective A bases its decisions on the actual probability distributions of the outcomes rather than on past information gains. For this agent, even a series of box A outcomes does not diminish the value of choosing the *AC* pair because, given the low probability of a box C outcome, the *AC* pair still offers the highest expected information gain. That is, the forward-looking strategy is more resilient to recent, uninformative outcomes compared to the backward-looking strategy, which is more reactive to immediate history of experience information gain, especially with high learning rate.

Overall, these simulated patterns of behavior informed the design of Experiment 8, providing normative criteria for when infants' choices could clearly reflect different underlying strategies. Our primary interest lay in the decisions made by infants during the second part of the experiment, from the 13th to the 24th trial, as this is the phase where the model predicted a steady divergence in choice probabilities under all four objectives. We reasoned that if infants would align their choices with any of the objectives, this critical second phase was when such behavior was most likely to emerge. Therefore, in addition to fitting the models to trial-by-trial behavior throughout the entire experiment, our analysis also focused on identifying patterns (specifically whether choices box pairs *AB* or *AC* dominates above chance) suggesting a shift toward one of the strategies during this key phase.

Moreover, based on the discussed difference in sensitivity to recent experience between forward- and backward-looking models, we paid particular attention to how infants modified their choices in response to the repeated A outcomes in trials 15 to 17. This sequence of trials served as a potential cue to their underlying strategy: a reliance on forward-looking assessment would likely maintain consistent AC choices aligned with expected information gain, while a backward-looking approach would lead to a transient shift in behavior to equally favoring the AB and AC choices.

6.2.2 PARTICIPANTS

A total of 84 full-term, typically developing 12-month-old infants were tested in the experiment. All infants were recruited in Budapest for research purposes, based on birth records. 23 participants were not included in the analysis: 17 due to fussiness, 6 due to their caregiver's interaction, and 3 due to equipment failure during the session. In addition, 22 more of the remaining participants failed to meet the inclusion criteria (detailed in the results section). Thus, 36 infants' data were included in the analysis (N = 36; 14 females, M_{age} = 12 months, 17.38 days; SD_{age} = 6.81 days), similar to Experiment 7. The resulting exclusion rate was 56%, which is significantly higher than the 30% exclusion rate in the previous experiment. This difference can be attributed to the longer overall experimental session and the additional criteria of providing 50% valid trials in the second, critical part of the experiment.

The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB), Budapest, Hungary, reference: 2019/13, and conducted in accordance with the principles defined in the Declaration of Helsinki. Prior to their participation in the research, all caregivers provided written informed consent.

6.2.3 APPARATUS AND STIMULI

Animations were created and rendered in Maya®, a 3D computer animation software (Autodesk Inc.). The resulting stimuli were displayed on the inbuilt monitor of a Tobii T60 XL eye tracker device (51.1 x 32.3 cm, 1920 x 1200 px), which recorded participants' gaze and pupil diameter. The experiment was controlled by a custom script written in Python 3.8 (Van Rossum & Drake, 2009). The stimuli, experiment control scripts, and data files are available at https://osf.io/uf2bn/ and the preregistration at https://doi.org/10.17605/OSF.IO/A639V.[18] The model's implementation, the simulations and the fitting procedure were also written and conducted in Python, these scripts are available at: https://osf.io/uf2bn/

6.2.4 Design and procedure

The experiment took place in a soundproof room with dimmed lights. Participants were seated on their caregiver's lap, about 60 cm from the display. Caregivers wore opaque glasses that prevented them from seeing the stimuli. They were instructed to keep the child seated and not to interact with them. The experimenter was sitting behind a curtain, controlling the experiment, and monitoring infants' behavior via a video camera. Before the experiment, participants underwent a 5-point calibration procedure. The entire session was video-recorded. The experiment consisted of a familiarization phase, followed by 24 trials.

The procedure differed from that of Experiment 7 in four aspects. First, the experiment included no trials with an elimination phase. Second, the distribution of the characters' hiding location was not uniform, but skewed (box A: 11/24, box B: 9/24, box C: 4/24). Third, the trial outcomes were fixed for all participants (as derived from prior simulations) and were not

---

[18] Experiments 7 and 8 were preregistered together. However, in response to a high exclusion rate due to infant fussiness observed in Experiment 7 and the unexpected results of Experiment 8, we implemented a set of deviations from the preregistration plan.

First, although we preregistered that Experiment 8 would include six test trials following 24 training trials – mirroring the structure of Experiment 7 (which included six test trials after 18 training trials) – we observed a high exclusion rate in Experiment 7 despite the shorter procedure. To mitigate this issue and reduce infant fatigue, we opted not to include additional test trials in Experiment 8, particularly since these trials were not critical for evaluating the explanations proposed for the findings of Experiment 7.

Second, we initially preregistered three objectives for the Ideal Shaker model but implemented only one of these (Objective A). We chose to omit the originally planned objective of minimizing uncertainty about the parameters of the underlying probability distribution and a related mixed model. Instead, we implemented three alternative objectives to better account for the observed behavioral patterns. This decision was motivated by two considerations. First, the unexpected results of Experiment 7 suggested that infants might have employed any alternative strategy: favoring full boxes either to confirm hypotheses or to maximize the likelihood of hearing the shaking sound. We therefore prioritized modeling this alternative (Objective B). Second, after preregistration, we recognized that the response patterns predicted by a forward-looking strategy of maximizing expected information gain might be similar to those predicted by a backward-looking reliance on cumulative past information gain. To address this, we introduced an additional model objective to capture this latter strategy (Objective C) and also included a backward-looking variant of Objective B (Objective D).

contingent on participants previous choices. The outcome sequence was *B, A, C, B, A, C, A, A, B, A, B, A, B, A, A, A, B, B, B, A, A, B, A, C*.

Finally, while in Experiment 7 the relationship between two buttons and the boxes they shook was highlighted by two lines ("arms") spreading out of the buttons to the corresponding boxes after a decision was made, here, the two lines were already connected to the respective boxes from the beginning of the decision phase. This manipulation was introduced to provide visual cues to the contingencies between buttons and box pairs thus decreasing the task demand of recalling them from memory before each decision. Thus, the shaking phase was shorter by 250 ms, as the "arm spreading" animation was not included. Beyond these modifications, the procedure was the same as in Experiment 7.

As in Experiment 7, infants were given visual feedback of their gaze position during the decision phase, in the form of a small blue dot, tracking gaze with minimal delay, with its movement smoothed by averaging its x-y positions over 150 milliseconds.

The following variables were counterbalanced between subjects: 2 possible mappings between circles and box pairs (left button shakes boxes A and B, and right button shakes boxes A and C, or vice versa), 6 possible box locations (box A: top, box B: center, box C: bottom, etc.).

6.2.5 **RESULTS**

As in Experiment 7, we included trials in the data analysis with more than 50% valid gaze samples during the decision and the shaking phase (9.24% of the trials were excluded). Participants were included in the final sample if they contributed at least 2/3 (16 out of 24) of the trials. In addition, to maximize the probability that the included participants were engaged with the task in the phase where we expected choices to diverge based on the hypothesized strategies, we included only those infants who provided valid trials on at least 50% (6 out of 12 trials) of the critical second half of the experiment (N of overall valid trials M = 21.80 , SD = 2.29, N of critical valid trials M = 10.41, SD = 1.90).

We first examined whether participants exhibited an increasing trend in anticipatory looks at the boxes throughout the experiment, specifically after making a fixation decision. The actual movement of the boxes began 500 ms after the gaze decision (250 ms sooner than in the previous study due to the omission of the arm's spreading animation). Therefore, if the participant's first look at the boxes occurred within this 500 ms window, we coded it as anticipatory.

To assess whether the likelihood of anticipatory looks increased over time (i.e., across trials), we applied a generalized linear mixed model (GLMM) with a binomial distribution and logit link function. Based on visual inspection of reaction time patterns (Figure 21), we included "session half" (first vs. second half) as a fixed factor, as anticipations appeared to increase during the first half but plateau in the second half. The model predicted the binary outcome of whether a look occurred before the 500 ms threshold and included trial number, session half, and their interaction as fixed effects, with participant identity as a random effect. The random effects structure allowed both intercepts and slopes for trial number to vary by participant.

The model's intercept was significantly negative (Estimate = -3.09, SE = 0.58, z = -5.33, p < 0.001), indicating that the likelihood of anticipatory looks was low at the beginning of the experiment. While the effect of trial number was positive (β = 0.09, SE = 0.07), it did not reach statistical significance (z = 1.25, p = 0.211), suggesting no strong evidence for a consistent

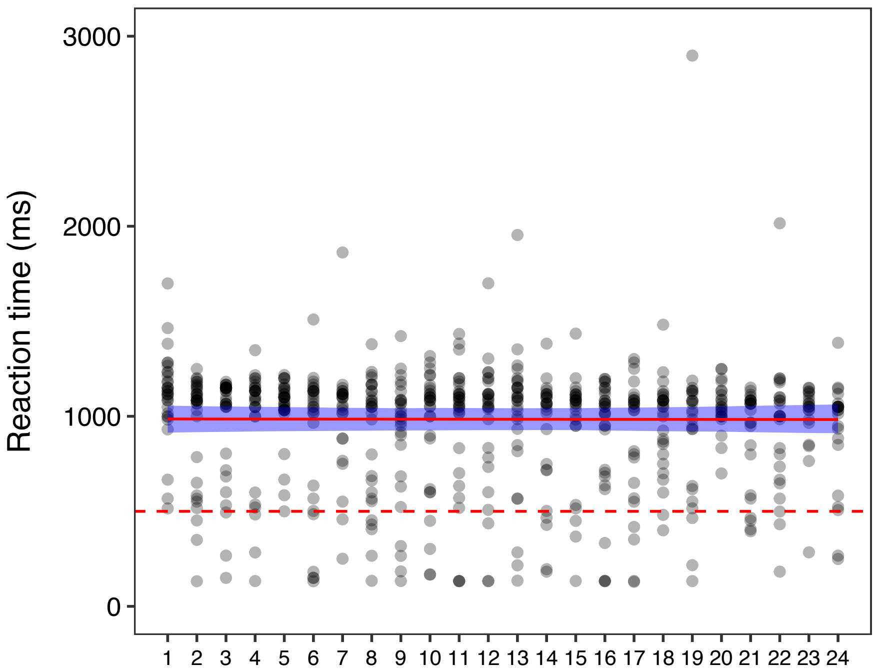

Figure 21. The timing of first looks at the boxes after decision onset throughout the Experiment 8. Points represent subjects' reaction times, the red line indicates the model-predicted reaction time for each trial, the blue-shaded region captures the 95% confidence interval. The red dashed line at 500 ms indicates the anticipation threshold.

increase in anticipatory looks across trials. The effect of session half was also positive, but not significant (β = 1.91, SE = 1.12, z = 1.70, p = 0.089), indicating a possible trend towards more anticipatory looks in the second half of the experiment. Importantly, the interaction between trial number and session half was significant (β = -0.18, SE = 0.09, z = -2.08, p = 0.037), suggesting that the relationship between trial progression and anticipatory looks differed across the two halves of the session. This suggests that the positive effect of trial progression on anticipatory looks was stronger in the first half and diminished in the second half. The random intercept for participants had a standard deviation of 0.88, indicating moderate variability across participants in their baseline likelihood of anticipatory looks, while the random slope for trial number had a low standard deviation of 0.08. Overall, results suggest that while the likelihood of anticipatory looks did not consistently increase across trials, there was evidence that this effect was moderated by session half, with trial progression having a stronger influence on anticipations during the first half of the experiment.

We examined the relationship between trial number and log-transformed fixation decision onset using a linear mixed model (Figure 22). In the analysis, trial number was included as a fixed effect, while participant identity was treated as a random effect, allowing both intercepts and slopes for trial number to vary by participant.

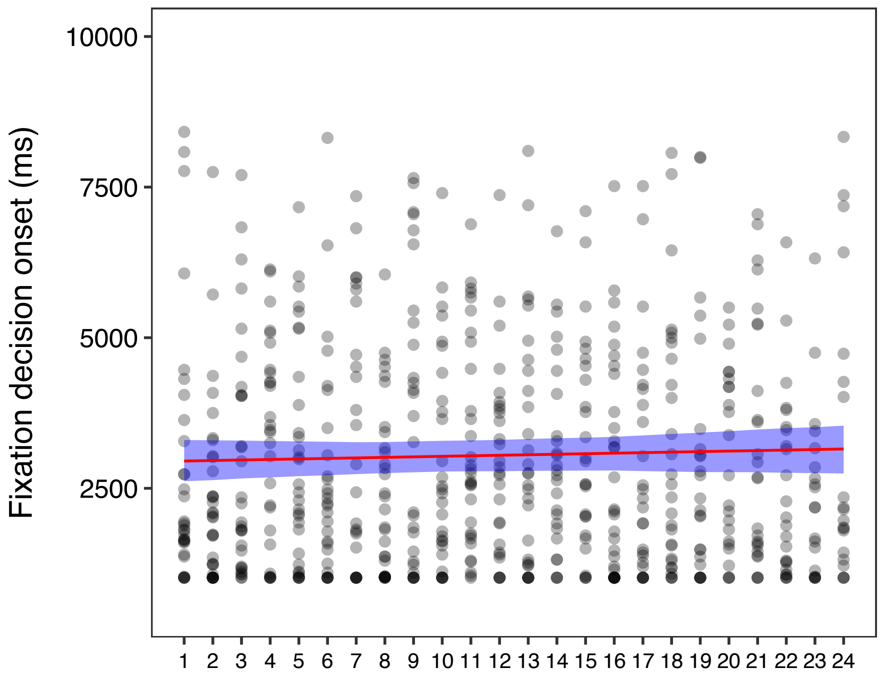

Figure 22. Fixation decision onsets throughout Experiment 8. Points represent subjects' decision onset, the red line indicates the model-predicted choice onset for each trial, the blue-shaded region captures the 95% confidence interval.

The effect of trial number was not significant (β = 0.003, SE = 0.004, t = 0.768, p = 0.448), suggesting that choice onsets did not increase (or decrease) linearly with session progression, differing from the findings of Experiment 7. This relative absence of a positive linear trend might be attributed to the session being twice as long as the training phase was in Experiment 7, or it could be because participants were more engaged with the task due to the non-uniform box probabilities introduced in the design.

To determine if participants exhibited a side bias in their fixation decisions during the trials, we used a GLMM to assess the binary outcome of choosing a particular side as the dependent variable. The model included random intercepts and slopes for trial number at the participant level, allowing both the baseline probability of choosing a side and the effect of trial number to vary across participants. The intercept representing the overall log-odds of selecting one side over the other was estimated at β = -0.229 (SE = 0.131, z = -1.752, p = 0.079), indicating no significant bias toward either side.

To assess choice strategies, we focused on analyzing the data from the second half of the experiment (trials 13-24). According to our model, this is when infants were expected to have learned the skewed probability distribution and their behavior would be more likely to reflect any of the decision-making strategies. To this end, we examined whether infants tended to choose the pair of boxes that aligned with information-related objectives (box pair AC, for objectives A and C) or full box shaking objectives (box pair AB, for strategies B and D). As these strategies predicted opposite choices in the second half of the session, we applied a GLMM to test whether the probability of choosing the button that shook the AC (rather than the AB) pair was significantly different from chance, accounting for individual differences and trial number as random factors.

The GLMM analysis revealed a significant effect for the intercept (β = 0.378, SE = 0.150, z = 2.518, p = 0.012), indicating that participants were significantly more likely to choose the AC option in the second half of the experiment (M_{AC} = 0.58, SD_{AC} = 0.20). This result supports the hypothesis that participants tended to adapt their strategies to favor choices that provided more information about trial outcomes, given the underlying distribution (Figure 23). Specifically, the 0.378 log-odds estimate indicates that the likelihood of choosing the informative button was about 10% above chance in the second, critical part of the experiment.

The random effects for participants showed relatively high variance ($\sigma^2 = 1.851$, SD = 1.360), suggesting considerable individual differences, while the random slope for trial number within participants had a very small variance ($\sigma^2 = 0.002$, SD = 0.045), suggesting minimal change in choice behavior over time within the second half of the session.

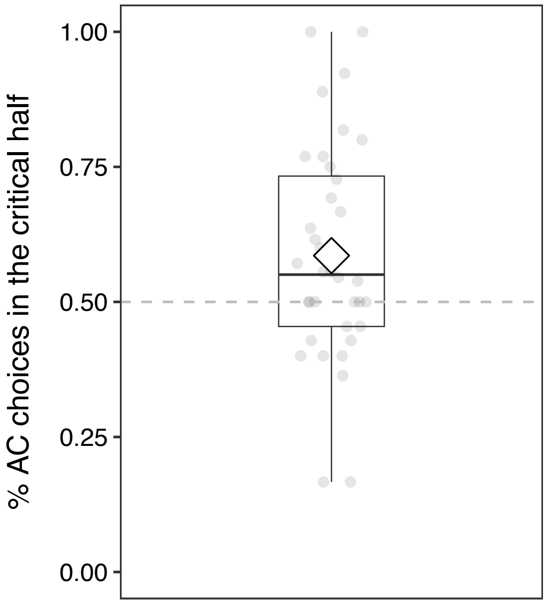

Figure 23. Boxplot of aggregate mean proportion of choosing the button shaking the AC versus AB box pair in Experiment 8 (N = 36). The white diamond indicates the mean, the bold horizontal line indicates the median, the box indicates the interquartile range, dots indicate individual aggregated proportions, the dashed line indicates chance performance.

Finally, we estimated the parameters and compared the fit of four models to infants' trial-by-trial choices, each model corresponding to one of the four objectives detailed previously. In what follows, the model implementing objective A (maximizing expected information gain) will be referred to as EXPINFO; the model implementing objective B (maximizing the probability of shaking the full box) will be referred to as EXPSHAKE; the model implementing objective C (selecting the option associated with maximum past information gain) will be referred to as PASTINFO; and the model implementing objective D (selecting the option associated with maximum past shaking sound rewards) will be referred to as PASTSHAKE.

The fitting process involved estimating the parameters τ_s (for all models) and η_s (for the PASTINFO and PASTSHAKE models) for each participant s, by maximizing the likelihood of the observed choices. The likelihood represents how probable the observed sequence of choices was, given a particular set of parameters. For each trial t, the likelihood of the observed choice $c_{s,t}$ by subject s was:

$$L_s(\theta_s) = \prod_{t=1}^{T_s} P(\, c_{s,t} \mid \theta_s \,)$$

where T_s is the total number of trials for subject s, and θ_s represents the model parameters (τ_s for EXPINFO and EXPSHAKE, and τ_s, η_s for PASTINFO and PASTSHAKE).

The likelihoods were derived from the softmax probabilities assigned to choices by the models. Log-likelihoods were used instead of raw likelihoods to handle numerical stability. The goal of the fitting process was to find the set of parameters θ_s^* that maximized the log-likelihood for each subject:

$$\theta_s^* = \arg\max_{\theta_s} \log L_s\,(\theta_s)$$

The optimization was performed using the L-BFGS-B algorithm in Python (using the SciPy library, Virtanen et al., 2020), with τ_s constrained between 0.01 and 1.0, and η_s bounded between 0.1 and 1 (to get realistic estimates and avoid fitting only noise in the data).

After obtaining the best-fitting parameters for each participant, the aggregated log-likelihood of each model was calculated as the sum of the log-likelihoods across all participants:

$$\log L = \sum_{s=1}^{S} \log L_s\,(\theta_s^*)$$

where S is the total number of participants. This aggregated log-likelihood provided an overall measure of how well a model fit the data across all participants.

To provide a measure that balances model fit and complexity, Bayesian Information Criterion (BIC) values were calculated. Lower BIC values indicate better models, considering both the goodness of fit and the simplicity of the model (here, the simplicity depended on the number of parameters, one or two).

To measure how much better the models fit the data compared to a null model (a model assuming all choices are equally likely), pseudo-R^2 was calculated as:

$$\text{Pseudo-R}^2 = 1 - \left(\frac{\log L}{\log L_{\text{null}}}\right)$$

where $\log L_{\text{null}}$ is the log-likelihood of the null model (random choice). Pseudo-R^2 values range from 0 to 1, with higher values indicating a better predictive accuracy.

	EXPINFO	EXPSHAKE	PASTINFO	PASTSHAKE
-LL	**437.11**	466.01	**424.09**	455.03
pseudo-R^2	**0.24**	0.19	**0.27**	0.21
# parameters	**1**	1	**2**	2
BIC	**880.99**	938.79	**861.70**	923.59
τ^*	**0.71** ± 0.37	0.80 ± 0.30	**0.75** ± 0.33	0.80 ± 0.36
η^*	**-**	-	**0.22** ± 0.24	0.22 ± 0.27

Table 3. Quality of behavioral fits from the 36 subjects in Experiment 8, for the four models. The best fitting models' metrics are highlighted with bold. -LL: negative log-likelihood. BIC: Bayesian Information Criterion. τ^*: estimated temperature parameter mean ± SD. η^*: estimated learning rate parameter mean ± SD.

The results of the model fitting and parameter estimation procedure are summarized at Table 3. The mean fitted τ (temperature) parameters ranged between 0.7 and 0.8 for all models, consistent with moderate randomness in choices, while the mean η (learning rate) parameter was estimated to be 0.22 for both backward-looking models, consistent with more gradual updates.

When comparing models, the PASTINFO model provided the best overall fit to the data, with the lowest negative log-likelihood ($-LL_{PASTINFO} = 424.09$) and the lowest BIC score ($BIC_{PASTINFO} = 861.70$), indicating that it not only captured the participants' choice behavior more accurately than the other models but also did it if accounting for its extra parameter. The EXPINFO model also performed well, with a slightly higher negative log-likelihood ($-LL_{EXPINFO} = 437.11$) and BIC score ($BIC_{EXPINFO} = 880.99$), suggesting that it was a competitive alternative to PASTINFO. In contrast, the EXPSHAKE and PASTSHAKE models had higher negative log-likelihoods ($-LL_{EXPSHAKE} = 466.02$, $-LL_{PASTHAKE} = 455.03$) and correspondingly higher BIC scores ($BIC_{EXPSHAKE} = 938.79$, $BIC_{PASTSHAKE} = 923.59$), indicating that these models were less predictive of infants' choices.

The pseudo-R^2 values further support these findings. The PASTINFO model achieved the highest pseudo-R^2 (pseudo-$R^2_{PASTINFO} = 0.27$), indicating that this model explains approximately 27% of the variance in the choice data beyond what could be expected by chance alone. The EXPINFO model followed closely (pseudo-$R^2_{EXPINFO} = 0.24$), while both the EXPSHAKE and PASTSHAKE models exhibited lower pseudo-R^2 (pseudo-$R^2_{EXPSHAKE} = 0.19$, pseudo-$R^2_{PASTSHAKE}$ =

0.21), consistent with their higher BIC scores and poorer log-likelihoods. The moderate pseudo-R^2 values exhibited by the two best models here are within the range often reported in studies targeting decision-making in young populations (e.g., Giron et al., 2022; Schulz, Wu, Ruggeri, & Meder, 2019).

In addition to these group-level findings, we conducted a participant-level analysis by extracting individual pseudo-R^2 values for each model and applying a linear mixed-effects model. This analysis, with participant identity as a random effect and model type as a fixed effect, revealed that the PASTINFO model, used as the reference, demonstrated superior predictive accuracy relative to both the EXPSHAKE and PASTSHAKE models. Specifically, the PASTSHAKE model exhibited a significantly lower pseudo-R^2, with an estimated difference of -0.051 (SE = 0.023, t(105) = -2.25, p = 0.026), and the EXPSHAKE model also showed a significantly lower pseudo-R^2, with an estimated difference of -0.070 (SE = 0.023, t(105) = -3.09, p = 0.003). These results are consistent with the lower group-level pseudo-R^2 values observed for the EXPSHAKE and PASTSHAKE models, supporting their relative inadequacy in participants' choices. In contrast, participants' pseudo-R^2 scores did not significantly differ under the EXPINFO (estimate = -0.022, SE = 0.023, t(105) = -0.98, p = 0.329) and PASTINFO models, indicating comparable performance between these two models. However, when comparing the relative evidence for these two candidate models, the Bayes Factor (using the models' BIC values) indicated strong evidence in favor of the PASTINFO model (BF = 15378.26).

As a form of posterior predictive check, using infants' empirically derived parameters, we simulated the behavior of 100 synthetic participants per real participant for each model. After generating the simulated data, we focused on the dynamics of the proportion of AC choices over time. To this end, we divided the overall 24 trials into 8 blocks, with each block containing 3 trials. For each synthetic participant, we calculated the mean proportion of AC choices within each block and plotted these simulated mean proportions against the observed mean proportions of AC choices from infants. The resulting figure (Figure 24) depicts the degree to which each model could reproduce the choice patterns observed in the experiment

The plot suggests that the EXPINFO and PASTINFO models were successful in capturing different aspects of infants' behavior on this task. Concentrating on blocks 4 to 8 (as in the first three blocks, both predictions and infants' choices fluctuated around chance), the sudden increase in AC choices around blocks 4 and 5 was better predicted by the EXPINFO model, while the decrease in AC choice proportion at block 8 (that is, at trials 16-18), followed by a relative increase in the last two blocks, aligned more closely with the behavior of the PASTINFO model. The drop of AC choices at block 8 under the PASTINFO model is likely due to its sensitivity to the consecutive A outcomes around this phase, resulting in the temporary devaluation of the AC option, as discussed previously.

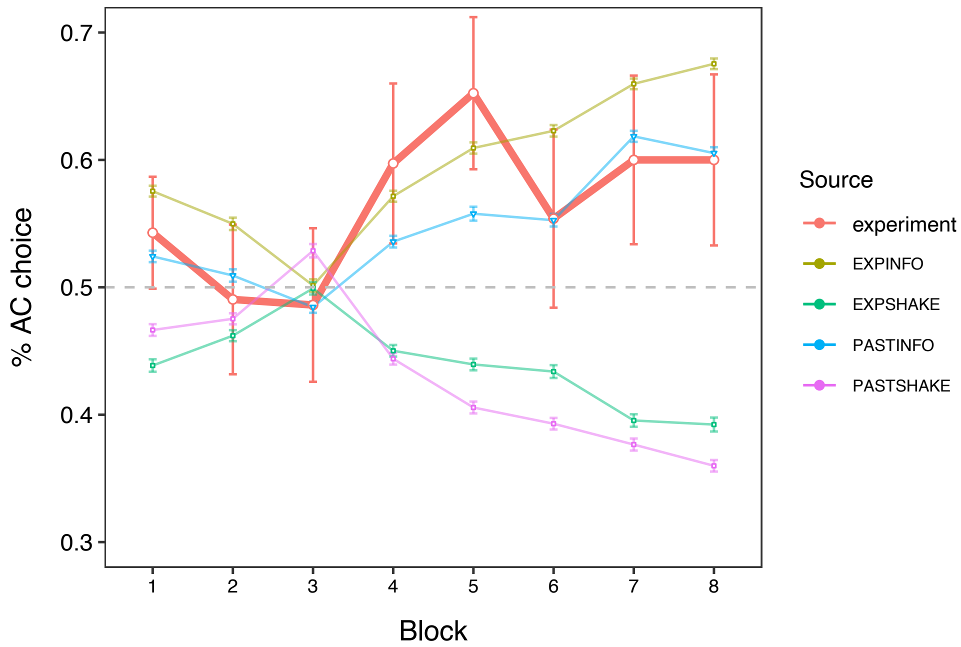

Figure 24. Comparison of mean proportion of AC choices across 8 blocks between real and simulated participants. The plot shows the mean proportion of AC choices made by real participants (labeled as "experiment") and those simulated by four different models: EXPINFO, EXPSHAKE, PASTINFO, and PASTSHAKE. Each block represents a set of 3 trials. Error bars indicate the standard error of the mean.

6.2.6 DISCUSSION

Experiment 8 aimed to investigate 12-month-olds' ability to actively use information sources in a modified gaze-contingent paradigm, building on the framework established in Experiment 7. In this version, infants were allowed to gradually learn and adapt to a skewed probability distribution rather than responding to a discrete shift in the probabilistic structure. The modeling approach used in this experiment enabled us to distinguish between information-directed strategies and other objectives, specifically, confirmatory and instrumental reward-driven (EXPSHAKE and PASTSHAKE) strategies. Additionally, it allowed us to predict distinct behavioral patterns under two different information-directed strategies: a forward-looking strategy focused on maximizing expected information gain (EXPINFO) and a backward-looking strategy based on past experiences and the success of previous actions (PASTINFO). Based on model simulations, we anticipated that these strategies would become apparent in the second half of the experimental sessions, once infants had presumably learned the underlying probability distribution guiding the trial outcomes.

The results demonstrated that infants, over the course of the trials, began to favor the informative box pair (AC) over the less informative pair (AB), especially in the second half of the experiment. This suggests that infants were not only sensitive to the underlying probability distribution but were also capable of adapting their behavior to options with higher information gain as the session progressed. When comparing different models, the PASTINFO model, which tracks past information gains, provided the best overall fit to the data. This model was closely followed by the EXPINFO model, which focused on maximizing future information gain. Exploratory analyses revealed that infants' behavior aligned with both the PASTINFO and EXPINFO models at different points during the experiment. Specifically, the initial increase in the selection of the informative box pair (AC) was better predicted by the EXPINFO model, while the later decline and subsequent recovery in this selection pattern aligned more closely with the PASTINFO model. This indicates that while infants may be influenced by forward-looking considerations, their choices are heavily shaped by reward history.

One might question whether infants' preference for the AC box pair could be attributed to the fact that the auditory outcome (the presence or absence of a shaking sound) from shaking these boxes was more variable and therefore more uncertain during the experiment. Research has shown that infants are drawn to stimuli with moderate rather than perfect predictability (Kidd, Piantadosi, & Aslin, 2012, 2014) and tend to engage in actions with moderate rather than perfect outcome contingency (Jacquey et al., 2020). Therefore, even though both choices provided the visual outcome of boxes shaking, a tendency to elicit more uncertain auditory outcomes may be considered a plausible alternative explanation of our findings. To explore this, we also modeled a strategy (detailed in the Appendix) that focuses on minimizing uncertainty about the auditory outcomes of shaking box pairs, independent of the final trial outcomes. Although this model provided a competitive fit compared to the top-performing models (PASTINFO and EXPINFO), its predicted behavior—characterized by a very gradual shift toward AC choices—failed to capture key aspects of infants' actual behavior, particularly the more abrupt changes observed during critical phases of the experiment (with the Bayes Factor analysis further favoring PASTINFO over this model). This suggests that while uncertainty reduction may influence infants' behavior, their choices in this paradigm are more accurately explained by strategies that integrate information gains.

Finally, it is important to emphasize that our predictions were derived from an *ideal* observer agent, which inherently simplifies the complexities of the task. All our models assumed that the agent makes decisions solely based on a single expected value, without accounting for learning about the experimental context itself. Infants may require time to familiarize themselves with the experimental setup, understanding the contingencies and the broader temporal structure of the task. This learning process could involve forming hypotheses about the relationship between their actions (e.g., fixating on buttons) and the subsequent outcomes (e.g., boxes shaking, opening, then the characters making their own characteristic movements), which the model assumes to be already established. For instance, participants might follow an "annealing schedule" (Gopnik, Griffiths, & Lucas, 2015; Kirkpatrick, Gelatt, & Vecchi, 1983): during the first few trials of the experiment, when they have not yet learned the button – box pair contingencies, they will likely make less optimal decisions with respect to the target objectives. Instead, they will be busy exploring the task scenario and the outcomes of fixating on different areas of the display. Future modeling could implement such a strategy by gradually decreasing the temperature and increasing the learning rate parameter over time.

6.3 **General discussion**

We designed Experiment 7 to investigate whether 12-month-old infants could evaluate and maximize expected information gain when choosing information sources in a gaze-contingent paradigm. However, findings failed to provide evidence for this ability, with the results being consistent with various alternative explanations. These included simple visual associations, auditory reward maximization (shaking boxes that contained rewards with highest probability, thus producing a shaking sound), and confirmatory hypothesis testing.

To address this, in Experiment 8, we introduced a different design where infants were not required to adapt to sudden changes in probabilistic structure. Instead, they were expected to gradually adjust to a skewed distribution where one option became increasingly more informative. The results provided strong evidence for information-maximizing strategies, as opposed to auditory reward maximization or confirmatory hypothesis testing. However, infants' behavior in this task could be explained by both a forward-looking strategy that maximized expected information gain and a backward-looking strategy based on past information gain, each accounting for different aspects of their choice patterns.

A significant limitation of our paradigm was its time-demanding nature. Infants were required to sit through a 7- to 8-minute experimental session with repetitive stimuli, often leading to early disengagement, as indicated by the high participant exclusion rates. This decrease in engagement likely affected their motivation to seek advanced information in later trials and to engage in the more computationally demanding, forward-looking evaluation of sources—especially since the information they gathered could not be directly used for any instrumental decision, but only to inform their expectations about trial outcomes. Future studies could improve our design by incorporating a more complex paradigm where infants could actually utilize the information they gather. Additionally, since our primary goal in Experiment 8 was to observe information maximization in a context where infants could gradually learn a skewed distribution, the experiment was not ideally suited to distinguish between forward- and backward-looking strategies of information gathering. More refined designs that carefully manipulate trial outcomes (and potentially a transient probability distribution) may be better equipped to address this distinction.

The conceptual distinction made between backward- and forward-looking strategies to maximize information can be directly related to the general dichotomy between model-free and model-based control of behavior. It is a longstanding idea that organisms may develop instrumental control systems of different complexity depending on what kind of organization of behavior outweighs its costs in their circumstances. Model-free (habitual) and model-based (goal-directed) systems are perhaps the most well-understood of these control mechanisms (Daw et al., 2005).

Model-free control can be linked to Thorndike's (1911) 'Law of Effect', and to operant conditioning, in general: those actions that led to reward in the past are more likely to get selected in the future. This retrospective system stores learned associations between environmental states and actions with predicted rewards, leading to decision-making that relies on selecting the action with the highest past reinforcement. However, since long-run utilities associated with each action in each state are learned through excessive experience, model-free learning, as a backward-looking strategy, is not robust to changes in the environment. Also, consistently choosing only those actions that led to rewards in the past could lead to substantial lost opportunities in the present (Dayan, 2008).

Instead of taking the substantial effort and risk of directly exposing themselves to the uncertain payoffs of their actions, organisms often facing novel situations are often better off with constructing, storing and updating a representation of their environment and its causal structure, and use it to guide decision-making. Such model-based control is characterized by the representation and prospective evaluation of potential action outcomes by utilizing an internal model of the environment (Dolan & Dayan, 2013). This internal model consists in a transition function, embodying the representation of probabilistic relationships between actions and their outcomes, and a reward function of probabilistic outcome-reward values. This system enables planning by creating decision trees of plausible future states and actions, allowing for the selection of actions with the highest long-term utility. This kind of forward-looking control is advantageous especially in transient, complex environments that tend to give rise to novel problems (e.g., Papineau, 2001, Dayan, 2008).

Distinguishing between model-based and model-free reinforcement learning in behavior is challenging, as both can produce similar outcomes in controlled environments. A common experimental approach to separate these influences is outcome devaluation (e.g., Adams & Dickinson, 1981). In such paradigms, subjects first learn to associate specific actions with rewards, and then one of the rewards is devalued in a different context from where the initial learning occurred. If subjects continue to perform an action despite the devalued reward, it suggests a model-free, habit-driven response, indicating reliance on past experiences rather than the current value of the reward. Conversely, if subjects adjust their behavior to avoid the devalued reward, it points to a model-based, goal-directed strategy, showing their ability to adapt actions based on updated information about potential outcomes. In fact, in Experiment 7, we aimed to adapt this devaluation scheme to the domain of information gain, by excluding one possible hiding location, suddenly rendering one option more informative than the other. However, this abrupt change may have also signaled a shift in the task structure, potentially triggering a positive testing strategy, which in our paradigm corresponded to choosing the less informative option (although this pattern in our data could also be explained by a simple response to the previous association between the remaining boxes and the button that previously activated them).

Although model-based and model-free components were traditionally conceptualized as being in competition with each other, recent research has challenged this idea, proposing architectures that are mixtures of model-free and model-based control, suggesting that these are best seen as prototypes, rather than independent modules (Collins & Cockburn, 2020). On one hand, transition models might be utilized to guide model-free credit assignment to actions, for instance by retrospectively applying a model to previous ambiguous choices (Moran et al., 2019). On the other hand, model-free values might be used instead of costly model-based evaluations during the planning process (Huys et al., 2015).

Studies from the recent years increasing suggest that control systems are not only evaluating instrumental rewards but also informational ones. A key focus of research has been on understanding how the brain assigns value to actions that lead to reward-relevant information. A common challenge in these studies is the natural correlation between gains in reward-predictive information and actual rewards, as more reliable cues typically result in more rewarding outcomes. To address this, Foley et al. (2017) developed a two-step decision task for macaques. In each trial, the monkeys first made a saccade to one of two cues, which then moved towards one of two targets. The monkeys were then required to make a second saccade towards one of the targets, receiving feedback in the form of a reward. The color of the cues provided information about cue validity, indicating the likelihood that the chosen cue's motion would accurately predict the rewarded target. By recording neuronal activity in the lateral intraparietal area (LIP) before the first saccades, the study found that monkeys preferred cues with higher validity and selected targets that aligned with cue predictions, thereby maximizing expected rewards. Furthermore, the firing rates of LIP neurons reflected the expected information gains related to the rewarded target, increasing as a function of cue validity. In a follow-up study, the researchers controlled for the confounding effects of information and reward gain by introducing a condition where pre-cues rendered the previously informative RF-cues non-informative, without altering their direction, timing, or reward expectations. LIP neurons selectively discriminated the validity of informative cues, ignoring the reward associations of non-informative items, suggesting that these parietal cells are modulated by expected information gain.

Parallel research has explored how expected rewards become associated with actions that provide evidence about the presence and amount of upcoming rewards, even when this evidence cannot be used for decision-making. For example, Bromberg-Martin and Hikosaka (2009, 2011) demonstrated that macaques developed a preference for cues leading to informative reward cues, even when the actual rewards were equal. Midbrain dopamine neurons showed increased activity between the initial decision and the onset of reward cues, reflecting a preference for gathering information. Later work by White et al. (2019) further refined these findings, showing that a neural network functionally similar to that encoding reward predictions motivates the gathering of reward-relevant information. This network's activity, originating in the anterior cingulate cortex and extending to the basal ganglia, was enhanced when monkeys directed their gaze toward informative cues and suppressed before they looked away, emphasizing the brain's valuation of information in the same currency as primary rewards.

What emerges from these results is indicative of a model-free control system from information gathering. According to Bromberg-Martin and Hikosaka's (2011) framework, actions that lead to informative sensory cues can be reinforced based on the information prediction errors they generate. Actions resulting in the unexpected denial of such cues would be penalized, whereas those followed by the predictable presentation of cues would maintain their current value. Building on this, White et al. (2019) proposed that information predictions linked to specific contexts could direct behavior towards acquiring information. In this model-free system, learned information predictions would indicate the expected information gain associated with particular actions within specific internal and environmental states (c.f. Carruthers & Williams, 2022). These internal states may include probabilistic representations defined by varying levels of uncertainty, guiding behavior towards maximizing information gathering.

Yet, it is likely that model-free, backward-looking control of information gathering is not enough in itself to explain the highly adaptive nature of human-level inquiry. Human-specific forms of information seeking, ranging from question asking and forging complex epistemic artefacts (e.g., telescope, radar, search engines) to setting up experiments, showcase the human capacity to exploit internal models to construct adaptive inquiries in the absence of prior experience with their informativity. Probably the most straightforward approach to test such prospective, model-based information-seeking capacities is to ask participants to generate questions. An early conceptual distinction made in this field is one between hypothesis-scanning and constraint-seeking questions (Mosher & Hornsby, 1966). While the former test hypotheses about variables directly (for example, when the variable of interest is what Sally ate for dinner, one could ask directly whether it was spaghetti), the latter reduce the space of hypotheses via testing features shared by many different hypotheses (e.g., asking whether the food contained meat) – which one of these is more informative in the current situation depends on the number of hypotheses and their current likelihoods. Optimal arbitration between these question types seem to develop relatively early: evidence indicates that 3- to 5-year-olds already show sensitivity to the prior probabilities of hypotheses, flexibly choosing questions that maximize information gain (for a review, see Jones, Swaboda, & Ruggeri, 2020).

However, as in the case of instrumental control systems, mechanisms responsible to evaluate and generate information gathering need not be purely model-free or model-based. For instance, people tend to reuse old questions from memory in novel situations, which likely helps constraining the vast space of possible candidate queries (Liquin, Rhodes, & Gureckis, 2022). Similarly, we might speculate that the pattern of choices found in Experiment 8 may be the result of an initial focus on expected information gain, followed by a backward-looking exploitation of the resulting expected values. This dual reliance on both forward-looking and backward-looking information-gathering strategies could represent an adaptive approach to navigating situations with recurring structures, as presented in our paradigm. Future research could explore integrating these strategies into a hybrid formal framework.

Chapter 7. **Conclusions and further directions**

The aim of the research project presented in this dissertation was to explore how infants reason about other agents' pursuit of information. In Chapter 1, I addressed the inductive challenge of inferring epistemic goals based on observed behavior. I suggested, in line with previous work on action understanding, that a set of assumptions about the underlying principles of information seeking might provide constraints to solve this problem, and I identified three principles that might already be represented by infants in their intuitive theory of goal-directed action.

The first principle, *instrumental relevance*, states that individuals often seek information that would aid them to optimally carry out their instrumental actions. Experiments 1 and 2, discussed in Chapter 2, were designed to investigate whether 14-month-old infants apply this principle when interpreting others' actions. In these experiments, infants watched an animated agent approach a goal object under different conditions, followed by the agent moving to a position that provided (or, in Experiment 2, *could* provide, according to the agent's false belief) a clear view of the object's location. This occurred either when the agent was uncertain about the object's location or when it already knew where the object was. The results showed that infants looked longer at the agent's action when it already knew the object's location, indicating that they found the information-seeking action more consistent with their expectations when the agent was uncertain, rather than certain, aligning with the idea that infants understand other agents' epistemic actions when put forward to aid instrumental goals.

The second principle I proposed was *drive toward novelty*, meaning that individuals are often drawn to seek information about novel events or objects in their environment. Experiment 3, presented in Chapter 3, adapted the paradigm used in the previous experiments to test whether infants operate with this principle in mind. Infants observed an animated agent moving to visually explore two objects, either when these objects were novel and ambiguous to the agent or when they had already been inspected. Results showed that infants looked comparatively more in the latter condition, suggesting that they found information gathering about these objects more reasonable when they were novel to the agent.

The third principle, *efficiency*, would suggest that individuals tend to gather information efficiently, balancing action costs with the value of the information obtained. Experiment 4, described in Chapter 4, tested infants' expectations regarding the efficiency of epistemic actions. Infants were shown an animated agent searching for a goal object by moving to one of two equally distant positions, each providing a different amount of information about the object's possible hiding locations. Infants looked longer when the agent chose a less informative position, indicating that they might expect agents to act efficiently when seeking information. This result provides some evidence for infants' expectations about the efficiency of epistemic actions, with the limitation that their observed pattern of looking time may also be consistent with an expectation to simply maximize the chance of looking at a target object in search scenarios, rather than with the general expectation to maximize expected information gain about a target variable.

Chapters 5 and 6 explored additional questions about infants' representations of epistemic actions. A key assumption in adult intuitive understanding of epistemic behavior is that agents seek information when they are uncertain about some variable. However, how infants represent such uncertainty is less clear. I proposed two possible representational schemes: a *modal scheme*, where uncertainty is represented as an attitude toward a set of possibilities, and an *open-ended scheme*, where uncertainty is represented as an attitude towards an open proposition with empty arguments. Experiments 5 and 6 in Chapter 5 aimed to provide evidence for the use of the modal representation of uncertainty in 14-month-old infants and 4-year-old children. The rationale for both experiments was that attributing disjunctive-inference-based knowledge requires the application of a modal representation of uncertainty, where different possibilities are explicitly considered. Thus, evidence supporting the ability to make such attributions would imply that infants and children are using the modal scheme to represent uncertainty.

In Experiment 5, infants were presented with a variation of the paradigm used in previous experiments. They observed an animated agent who was uncertain about the location of its goal object and could, through exclusion, potentially infer its location among three hiding places. We hypothesized that if infants were capable of attributing disjunctive-inference-based knowledge, they would look more at the scenario where the agent, despite being able to infer the object's location, still sought additional perceptual evidence, compared to when the agent sought evidence due to uncertainty. However, the infants did not show a differential looking pattern between these two conditions, indicating no clear evidence that they could attribute inference-based knowledge.

In Experiment 6, we targeted the ability of an older age group, specifically 4-year-olds, to make such attributions using a different paradigm. Children engaged with a tablet-based interactive game where they needed to seek help from one of two agents to feed animated animals hidden in a box. In the critical condition, one agent could infer the contents of the box using disjunctive reasoning and therefore assist in feeding the animal, while the other agent could not make such an inference. The results showed that children were more likely to seek help from the knowledgeable agent, suggesting they could attribute knowledge based on disjunctive reasoning. This provides evidence that 4-year-olds might use a modal representational scheme. However, for more robust conclusions, future studies should address the possibility that children simply preferred the "less ignorant" agent—the one who had knowledge about the truth or falsity of more possibilities (i.e., was aware of the emptiness of more potential box compartments) in the given scenario.

Finally, to apply the third proposed principle—efficiency—when attributing epistemic goals, infants must be capable of evaluating the potential information an observed action could yield for its actor regarding various candidate variables. In other words, they should be able to assess the *expected* information gain from an action. Experiments 7 and 8 explored 12-month-old infants' capacity to evaluate expected information gain from a first-person perspective, independent of third-person considerations. To this end, we developed a gaze-contingent paradigm, where, across a series of trials, infants could shake one of two partially overlapping pairs of boxes to gain advance information about animated characters' whereabouts between three potential hiding locations. The distribution of the characters' hiding locations was controlled across trials, allowing us to manipulate the expected information that shaking one pair of boxes over the other could provide. In Experiment 7, the probability distribution of the characters' hiding locations was uniform. However, in the final set of test trials, one hiding location was excluded at the beginning of each trial, making one pair of boxes suddenly more informative than the other. In Experiment 8, the probability distribution was skewed from the outset, and the informativeness of one pair of boxes over the other depended on the infants' representation of this distribution, which was modeled by an ideal observer model.

The results presented a complex picture of infants' performance in this paradigm. In Experiment 7, where the informativeness of one option over the other was introduced by a sudden change in context, infants unexpectedly tended to choose the less informative option above chance. This behavior could be interpreted as consistent with a confirmatory hypothesis-testing strategy, where the goal was to maximize the probability of shaking boxes that were full, or it might simply reflect a visual association between boxes and options. In Experiment 8, where the informativeness of one choice over the other gradually emerged as infants learned the skewed distribution, infants generally tended to choose the more informative option. However, their choices could be explained by both a strategy to prospectively maximize expected information gain and a tendency to choose options that had maximized information in the past.

In sum, what emerges from these findings is a preliminary understanding of epistemic behavior in infancy. It seems that infants can make sense of others' epistemic actions in simple situations, particularly when these actions provide uncertain agents with evidence about the location of a goal object or help identify novel objects. Moreover, in the former case, they seem to expect actions that could be more successful in locating the goal-object than other alternatives with equal cost.

However, we did not find conclusive evidence that infants conceptualize other agents' evaluation of epistemic goals in terms of expected information gain. Since the value of information is tied to how the probabilities of different possibilities shift following an observation, such a conceptualization would necessitate a modal representation of uncertainty, where an uncertain agent is represented as considering a set of possibilities. Our findings did not support the presence of this representational scheme in infancy; evidence for its use only emerged in 4-year-olds.

Second, we found mixed evidence regarding infants' ability to evaluate expected information gain from their own perspective when searching for information. While our results are consistent with the hypothesis that infants can evaluate expected information gain, our data can also be accounted for by a backward-looking strategy, where infants rely on the historical informativity of actions rather than prospective evaluation.

Moreover, it remains an open question for future research whether infants can interpret actions as epistemic in situations where the agent's inquiry involves variables beyond the location or identity of objects. Similarly, we focused on a single type of epistemic action mean: visual perception. However, humans can utilize a potentially infinite variety of actions for epistemic purposes, depending on their variable of interest. For example, hitting a wall with a hammer might initially appear to be a purely instrumental action, but apart from influencing the environment, the resulting environmental change might actually be used as a source of information for answering a set of questions — in this case, how fragile the wall is, or how thick it is. It is not yet clear when children begin to understand that individuals may alter their environment with epistemic goals in mind.

Finally, in all our studies, infants were first provided with clues about the observed agent's instrumental goal (approaching an object) and state of knowledge (being able to locate the object or not). Yet, evidence about these factors is often not immediately apparent and must be inferred based on the observed epistemic action. For instance, when observing someone flipping through the pages of a restaurant menu, one might infer that the person intends to eat there but is uncertain about the available options and which one to choose.

Relatedly, a comprehensive intuitive understanding of the causal factors behind epistemic actions should also support stable predictive inferences. For example, in our paradigm, one might expect infants to anticipate that the agent would position itself on the platform providing a view of its goal object, depending on the agent's level of uncertainty. However, since our paradigm was not designed to measure anticipatory fixations (with only one trial per test condition, leading to sparse and low-power gaze data), we currently lack evidence regarding infants' predictive expectations in this domain.

Huang, Hu, and Shao (2019) suggest that more advanced, productive inferences about epistemic actions may develop over a more extended period. In their study, children aged 5 to 7 years observed an agent's past actions and were asked to make either predictive inferences about the agent's future epistemic actions or backward inferences about the agent's prior knowledge. The paradigm involved three blocks hidden in three boxes, with two blocks sharing the same color (red) and one block having a different color (green). In the prediction task, children observed the agent opening one box, revealing either a block that shared its color with another block (red) or the uniquely colored block (green). They were then asked to predict whether the agent would open a second box to guess the content of the third box or make an immediate guess. Only children older than 6 years were able to make the correct prediction depending on whether the agent could infer the content of the third box. In the backward-inference task, children were shown that the agent opened the first box and then either chose to open or not open a second box. They were then asked which block was in the first box that would explain the agent's decision. Only 7-year-olds seemed to understand that finding the red block (that is, the one which was not unique) would explain why the agent opened the second box, and thus only this age group answered the task correctly. Although successful performance in these tasks required a combining reasoning about epistemic actions with counterfactual reasoning and disjunctive inference, these findings suggest that both forward- and backward-inferences regarding epistemic behavior may emerge later in development. Future studies could explore simpler paradigms for younger children to determine if these skills might show up earlier under less demanding conditions.

Beyond these limitations, several issues regarding the interpretation of epistemic behavior remain unaddressed in this dissertation. For example, although we can often identify information-seeking agents without prior knowledge of their mental states or goals, this ability is not straightforward. As previously argued, almost any action can provide information to an agent depending on the mental and environmental context. However, there may be certain sequences of movements that even infants can readily recognize as investigative, even without understanding the hypothesis space they constrain or the higher-order instrumental goal they serve. One such intuitive cue might be the gaze behavior of others, particularly if its representation is detailed enough. Currently, little is known about whether infants have specific expectations regarding agents that direct their gaze towards previously unseen locations (e.g., looking into boxes, inspecting the back of objects). Infants might be capable of distinguishing such stereotypical inquisitive gaze behaviors from shifts in attention triggered by external stimuli within the agent's prior field of view and may use this as a distinctive cue indicating information seeking.

A second pressing issue not explored in this work concerns whether and how children, as well as adults, differentiate between the epistemic and instrumental value of information. Many successful accounts of human naive psychology and action understanding posit that people reason about other agents' mental states and behavior by representing them as utility maximizers — agents aiming to maximize rewards while minimizing costs (Baker, Saxe, & Tenenbaum, 2009; Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017; Csibra et al., 1999; Csibra et al., 2003; Gergely et al., 1995; Gergely & Csibra, 2003; Jara-Ettinger, Tenenbaum, & Schulz, 2015; Jara-Ettinger, Floyd, Tenenbaum, & Schulz, 2017; Jara-Ettinger, Gweon, Tenenbaum, & Schulz, 2015; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Jara-Ettinger, 2019; Jern, Lucas, & Kemp, 2017; Liu & Spelke, 2017; Liu, Ullman, Tenenbaum, & Spelke, 2017; Lucas et al., 2014). These models predominantly focus on the assessment of instrumental rewards and corresponding costs. In contrast, there is currently a lack of clarity on how people think about informational rewards and how these are weighed against instrumental costs and rewards.

Information gain, the perhaps most common measure of information has been discussed in this dissertation at length. Under this notion, the value of information is proportional to the degree it reduces uncertainty about a well-defined set of hypotheses (Crupi et al., 2018; Lindley, 1956; Nelson, 2008; Shannon, 1948). However, integrating information gain with instrumental rewards in people's naive utility calculations is not straightforward, as uncertainty is a statistical notion, while instrumental rewards tend to be energy-based or monetary. Yet, this integration is crucial since the inferred rewards from potential actions must be comparable to effectively predict behavior and recognize intentions.

One possibility is that people primarily focus on the instrumental aspects of information – how it enhances the efficiency of pursuing instrumental goals. In other words, people may consider how others' expected utility may increase as a result of knowledge acquisition. This added instrumental bonus due to information can be formalized as expected utility gain. Consider a simple scenario where two mutually exclusive actions have exactly opposite utilities in two mutually exclusive, equally probable states, so that the average expected utilities are ultimately equal for all actions. Given this situation of indecision, where no action dominates, it might be rational to invest in epistemic behavior with a certain cost to gain certainty regarding relevant world states. More formally, adapting the analysis of Chater et al. (1998) and van Rooy (2003), the expected utility gain *EUG* from an epistemic action *EA* with the cost C and different possible *result*s, given the agent's initial knowledge K, can be expressed as:

$$EUG(EA) = \left[\sum_{result} P(result|K) \times EV(result)\right] - C(EA)$$

where the expected value of learning a *result* is the difference between the maximum expected utility of choosing with and without the information it provides (see Figure 25 for an illustration). This expected value can be expressed by the following equation, where $\mathcal{A}$ is the set of a actions and s are the possible world states that are considered in a given decision problem, and U is the utility of the outcome of action a in a possible world state s:

$$\begin{aligned} EV(result) &= UV(Learn\ the\ result,\ choose\ later) - UV(Choose\ now) \\ &= max_{a\in\mathcal{A}} EU(a,\ result) - max_{a\in\mathcal{A}} EU(a) \\ &= [max_{a\in\mathcal{A}} \textstyle\sum_s P(s|result) \times U(a,\ s)] - [max_{a\in\mathcal{A}} \textstyle\sum_s P(s) \times U(a,\ s)] \end{aligned}$$

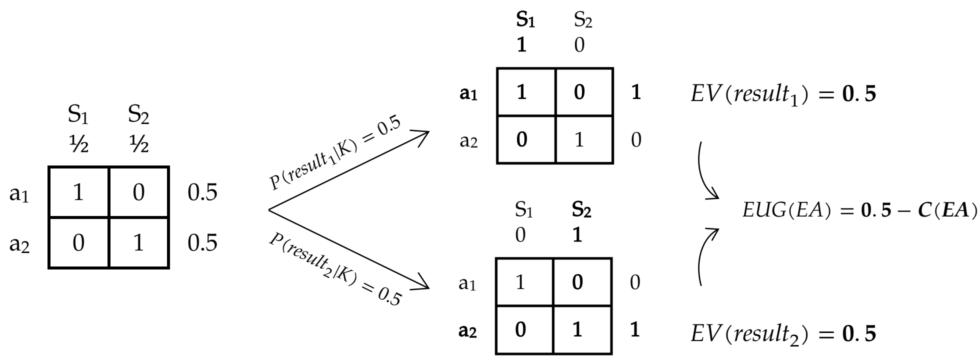

Figure 25. Example expected utility gain due to an epistemic action EA. The initial decision problem with two equally probable possible world states (S_1, S_2) and two possible actions (a_1, a_2) is updated by two equally probable results ($result_1$, $result_2$) via an epistemic action, changing the probabilities assigned to world states and thus also the maximum expected utilities by 0.5 in both cases. Weighed by the probabilities of the results, these expected values determine the expected benefit of the epistemic action, from which its costs can be subtracted.

If such a model is included in observers' naive utility calculus, several predictions should follow. Ceteris paribus, the value of a candidate epistemic action should be seen as higher when (i) the cost of the action is lower, (ii) the number of goal-relevant possible world states it disambiguates is larger, and (iii) the difference between the utilities of possible outcomes it helps to decide among is greater. These predictions, implying that information gathering is highly sensitive to the specific circumstances of the decision at hand, are quite intuitive.

However, this formulation faces challenges regarding its real-world validity. Firstly, it is uncertain whether people genuinely optimize their behavior based on expected utility gain (Markant & Gureckis, 2012). Secondly, both adults and, as indicated by the results of Experiment 3, infants seem to reason about curious agents seeking information for the sake of knowledge, without necessarily considering its practical uses.

Differentiating between instrumental and non-instrumental information gathering is notoriously challenging. This issue is particularly evident in research on animal curiosity. Unlike human subjects, who can articulate their goals when seeking information, animals cannot, which means researchers must ensure that their paradigms truly capture non-instrumental curiosity. For example, Wang, Sweis, and Hayden (2018) propose the following three criteria:

> "*First, operationally defined curiosity requires a willingness to sacrifice reward in order to obtain additional information (…). Second, the amount the subject is willing to pay must scale with the amount of available additional information (…). Third, that provided information must provide no obvious instrumental or strategic benefit, even subjectively.*" (Wang, Sweis, & Hayden, 2018, p. 8).

Evidently, the third criterion is the most difficult to establish, and arguably, it is the most challenging to figure out for observers of others' epistemic behavior as well.

To complicate issues further, human goals are typically hierarchically structured (Jackendoff, 2007; Velez-Ginorio et al., 2017). It is possible to distinguish very high-order, broadly defined goals, that may be called as desires (such as "wanting to eat") and lower-order, well-defined and detailed goal states that fulfill these desires (such as "eating a tuna sandwich at the deli on the corner"). When individuals are confronted with high-level goals without predetermined strategies for achieving them, it can often be challenging to predict how acquiring new information will impact their expected utilities. In such situations, people frequently seek to develop a generic model of the relevant factors that influence their decisions.

Consider the following example of participating in an escape room game, where the primary, high-level goal is to escape from a locked room. Upon entering the escape room, the first challenge is to understand the situation at hand. This initial phase involves uncertainty about the overall structure and components of the environment. To address this, first step may be to turn on the lights to get a clear view of the room and its contents. This action allows participants to construct a basic model of the situation, identifying the possible options available for achieving the goal of escaping. In this phase, participants may explore the room by asking questions such as: "*What are the objects present in the room?*", "*What are their features and causal properties?*", "*How do these objects relate to one another?*".

As the game progresses, participants might discover that operating a set of cranks on the wall occasionally yields pieces of a key needed to unlock the door. This introduces a new uncertainty pertaining to the parameters of the model they are constructing. In this case, the uncertainty revolves around the probability that each crank will produce a key piece. To resolve this, participants might begin exploring the cranks systematically, gathering data over time to refine their estimates of the likelihood that each crank will produce a key piece. This exploration reduces estimate uncertainty, allowing participants to make more informed decisions about which cranks to prioritize as time becomes increasingly limited.

Finally, suppose that through their exploration, participants learn that one particular crank has a 90% chance of yielding a key piece. At this point, the decision to use this crank appears rational. However, despite having a high probability of success, there is still a 10% chance of failure due to inherent variability in outcomes. Complicating matters further, participants may be constrained by time, with only 10 seconds left to complete the task. Here, participants might discover a way to peek into the pipes containing the key parts before using the crank. This action provides perfect information about the outcome, thereby eliminating outcome uncertainty and enhancing their chances of success.

This scenario highlights how individuals encounter and address different types of uncertainty at various levels of abstraction as they work toward implementing their high-level goals. It shows that even though participants' eventual goal was instrumentally defined (escaping the room), it required learning a generic model of their environment to be able to consider the specific action means of towards implementing it. Such learning is straightforward to formalize in terms of information gain, but problematic to account for in the expected utility gain framework. Finally, in the example provided, the higher-level goal—escaping from the room—was clearly defined from the outset. However, when observing others from a third-person perspective, without such prior information about higher-order goals, it is easy to misinterpret instrumental information gathering as non-instrumental.

But perhaps the view that recognizing a higher-level goal is necessary to interpret an action as instrumental might be fundamentally flawed. It is possible that people have a prior inductive bias to expect information gathering to be instrumental, justified by the basic human tendency of utility maximization. This expectation may arise not from an explicit recognition of a specific goal but from the understanding that most actions involving effort and resource expenditure are eventually must be serve some form of utility maximization. Even if information-seeking individuals sometimes have no specific instrumental goals in mind, the knowledge acquired through their curiosity could prove beneficial for future, unforeseen objectives.

Indeed, human decision-making processes are the product of evolutionary pressures, and, as Okasha (2017) points out, although biological rationality is logically independent of classical rationality, they are likely connected empirically. That is, an organism might have consistent preferences that do not necessarily maximize its fitness in certain environments, and conversely, behaviors that maximize fitness might not always align with the principles of classical rationality. However, these two forms of rationality are repeatedly correlated: for organisms with utility functions shaped by evolutionary pressures, the maximization of utilities often provides selective advantages over other strategies. This suggests that our subjective utility functions (our preferences and desires) are often not arbitrary but are instead shaped by the demands of survival and reproduction. As a result, even though curiosity might not have immediate practical benefits, it is likely advantageous in the long run, as it contributes to a broader knowledge base that can be leveraged in novel situations, ultimately enhancing fitness and survival.

The idea of considering epistemic behavior as primarily instrumental corresponds closely to the analysis of the human knowledge concept provided by Edward Craig in *Knowledge and the State of Nature* (1990). Craig suggests that the concept of knowledge may have evolved from practical needs within early human societies — specifically, the need to identify reliable informants who could provide accurate information when it mattered most. In this view, knowledge as a concept is rooted in social practices and survival strategies.

Relatedly, the notion of knowledge acquisition may also be seen as predominantly instrumental in our naive psychology.

Yet, even if we implicitly assume that most information gathering is driven by some instrumental purpose (even when we cannot always clearly pinpoint this goal), there remains a folk concept of non-instrumental curiosity that we find deeply intuitive. Future research should investigate how humans, both children and adults, reconcile these notions and how this is reflected in their understanding of epistemic behavior.

References

Abbeduto, L., & Rosenberg, S. (1985). Children's knowledge of the presuppositions of know and other cognitive verbs. *Journal of Child Language, 12*, 621–641.

Aboody, R., Zhou, C., & Jara-Ettinger, J. (2021). In pursuit of knowledge: Preschoolers expect agents to weigh information gain and information cost when deciding whether to explore. *Child Development, 92*(5), 1919–1931.

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B, 33*(2b), 109–121.

Adriaans, P., & van Benthem, J. (2008). Information is what information does. In *P. Adriaans & J. van Benthem (Eds.), Philosophy of Information (pp. 3–26)*. Amsterdam: North-Holland.

Aguirre, M., Brun, M., Couderc, A., Reboul, A., Senez, P., & Mascaro, O. (2022). Knowledge in Sight: Toddlers Plan Efficient Epistemic Actions by Anticipating Learning Gains. *Cognitive Science, 46*(2), e13103.

Altmann, E. C., Bazhydai, M., & Westermann, G. (2021, June). 'Until it bores me': Learning Progress Maximization as the Reward Mechanism to solve the Exploration-Exploitation Dilemma in Infants. In *Development in Motion Conference 2021*.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407.

Aslin, R. N. (2017). Statistical learning: a powerful mechanism that operates by mere exposure. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(1–2), e1373.

Baer, C., & Friedman, O. (2018). Fitting the message to the listener: Children selectively mention general and specific facts. *Child Development, 89*(2), 461–475.

Baillargeon, R. (1987). Object permanence in 3 1/2-and 4 1/2-month-old infants. *Developmental Psychology, 23*(5), 655–664.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110–118.

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), 598.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. Cognition, 113(3), 329–349.

Baldwin, D., Loucks, J., & Sabbagh, M. (2008). Pragmatics of Human Action. In T. F. Shipley & J. M. Zacks (Eds.), *Understanding Events: From Perception to Action (pp. 96–129)*. Oxford: Oxford University Press.

Bar-Hillel, Y., & Carnap, R. (1952). An Outline of a Theory of Semantic Information. *Technical Report No. 247, Research Laboratory of Electronics*.

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior, 34*(3), 164–175.

Barclay, P., & Willer, R. (2007). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences, 247*, 749–753.

Barone, P., Corradi, G., & Gomila, A. (2019). Infants' performance in spontaneous-response false belief tasks: A review and meta-analysis. *Infant Behavior and Development, 57*, 101350.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models Using LME4. *Journal of Statistical Software, 67*(1).

Baumard, N., André, J.-B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences, 36*(1), 59–122.

Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77*(2), 413–426.

Begus, K., & Bonawitz, E. (2020). The rhythm of learning: Theta oscillations as an index of active learning in infancy. *Developmental Cognitive Neuroscience, 45*, 100810.

Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function. *Developmental Science, 15*(5), 611–617.

Begus, K., Gliga, T., & Southgate, V. (2014). Infants learn what they want to learn: Responding to infant pointing leads to superior learning. *PloS One, 9*(10), e108817.

Begus, K., Gliga, T., & Southgate, V. (2016). Infants' preferences for native speakers are associated with an expectation of information. *Proceedings of the National Academy of Sciences, 113*(44), 12397–12402.

Bennett, D., Bode, S., Brydevall, M., Warren, H., & Murawski, C. (2016). Intrinsic valuation of information in decision making under uncertainty. *PLoS Computational Biology, 12*(7).

Berlyne, D. (1960). *Conflict, Arousal and Curiosity*. New York: McGraw-Hill.

Bickerton, D., & Szathmáry, E. (2011). Confrontational scavenging as a possible source for language and cooperation. *BMC Evolutionary Biology, 11*, 1–7.

Bion, R. A. H., Borovsky, A., & Fernald, A. (2012). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, 126, 39–53.

Bohus, K. A., Cesana-Arlotti, N., Martín-Salguero, A., & Bonatti, L. L. (2023). The scope and role of deduction in infant cognition. *Current Biology, 33*(18), 4014-4020.

Bonawitz, E. B., Bass, I., & Lapidow, E. S. (2018). Choosing to Learn: Evidence evaluation for active learning and teaching in early childhood. In P. Ganea & M. Saylor (Eds.), *Active Learning from Infancy to Childhood: Social Motivation, Cognition, and Linguistic Mechanisms (pp. 213–231)*. Springer International.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The Cultural Niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences, 108*(2), 10918–10925.

Bramley, N. R., Gerstenberg, T., Tenenbaum, J. B., & Gureckis, T. M. (2018). Intuitive experimentation in the physical world. *Cognitive Psychology, 105*, 9–38.

Braunlich, K., & Love, B. C. (2022). Bidirectional influences of information sampling and concept learning. *Psychological Review, 129*(2), 213–234.

Bridgers, S., Jara-Ettinger, J., & Gweon, H. (2020). Young children consider the expected utility of others' learning to decide what to teach. *Nature Human Behaviour, 4*(2), 144–152.

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron, 63*(1), 119–126.

Bromberg-Martin, E. S., & Hikosaka, O. (2011). Lateral habenula neurons signal errors in the prediction of reward information. *Nature Neuroscience, 14* (9), 1209–1216.

Bromberg-Martin, E. S., & Monosov, I. E. (2020). Neural circuitry of information seeking. *Current Opinion in Behavioral Sciences, 35*, 62–70.

Brosseau-Liard, P. E., & Poulin-Dubois, D. (2014). Sensitivity to confidence cues increases during the second year of life. *Infancy, 19*(5), 461–475.

Bruner, J. S. (1973). Organization of early skilled action. *Child Development, 44*(1), 1–11.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*(2), 337–342.

Buttelmann, F., & Kovács, Á. M. (2019). 14-Month-olds anticipate others' actions based on their belief about an object's identity. *Infancy, 24*(5), 738-751.

Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology, 131*, 94–103.

Carruthers, P. (2018). Basic questions. *Mind & Language, 33*(2), 130–147.

Carruthers, P. (2019). Questions in development. *In L. Butler, S. Ronfard, & K. Corriveau (Eds.), The Questioning Child: Insights from Psychology and Education (pp. 6–28)*. Cambridge: Cambridge University Press.

Carruthers, P. (2024). The contents and causes of curiosity. *British Journal for the Philosophy of Science*.

Carruthers, P., & Williams, D. M. (2022). Model-free metacognition. *Cognition, 225*, 105117.

Casillas, M., & Frank, M. C. (2017). The development of children's ability to track and predict turn structure in conversation. *Journal of Memory and Language, 92*, 234–253.

Caza, J. S., Atance, C. M., & Bernstein, D. M. (2016). Older (but not younger) preschoolers understand that knowledge differs between people and across time. *British Journal of Developmental Psychology, 34*(3), 313-324.

Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science, 359*(6381), 1263-1266.

Cesana-Arlotti, N., Varga, B., & Téglás, E. (2022). The pupillometry of the possible: an investigation of infants' representation of alternative possibilities. *Philosophical Transactions of the Royal Society B: Biological Sciences, 377*(1866), 20210343.

Chater, N., Crocker, M. J., & Pickering, M. J. (1998). The Rational Analysis of Inquiry: The case of parsing. In M. Oaksford & N. Chater (Eds.), *Rational Models of Cognition*. Oxford: Oxford University Press.

Ciardelli, I., Groenendijk, J., & Roelofsen, F. (2019). *Inquisitive Semantics*. Oxford: Oxford University Press.

Cimpian, A., & Scott, R. M. (2012). Children expect generic knowledge to be widely shared. *Cognition, 123*(3), 419-433.

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

Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology, 79*, 102–133.

Collins, A. G., & Cockburn, J. (2020). Beyond dichotomies in reinforcement learning. *Nature Reviews Neuroscience, 21*(10), 576–586.

Craig, E. (1990). *Knowledge and the State of Nature: An essay in conceptual synthesis*. New York: Oxford University Press.

Crivello, C., & Poulin-Dubois, D. (2018). Infants' false belief understanding: A non-replication of the helping task. *Cognitive Development, 46*, 51–57.

Croom, S., Zhou, H., & Firestone, C. (2023). Seeing and understanding epistemic actions. *Proceedings of the National Academy of Sciences, 120*(49), e2303162120.

Crupi, V., Nelson, J. D., Meder, B., Cevolani, G., & Tentori, K. (2018). Generalized information theory meets human cognition: Introducing a unified framework to model uncertainty and information search. *Cognitive Science, 42*(5), 1410–1456.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*(2), 705–717.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1567), 1149–1157.

Csibra, G., & Volein, Á. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology, 26*(1), 1–11.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27*(1), 111–133.

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of "pure reason" in infancy. *Cognition, 72*(3), 237–267.

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology, 52*(4), 521.

Cubit, L. S., Canale, R., Handsman, R., Kidd, C., & Bennetto, L. (2021). Visual attention preference for intermediate predictability in young children. *Child Development, 92*(2), 691–703.

Davidson, D. (1963). Actions, Reasons, and Causes. *The Journal of Philosophy, 60*(23), 685–700.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*(12), 1704–1711.

Dayan, P. (2008). The role of value systems in decision making. In C. Engel & W. Singer (Eds.), *Better than conscious? Implications for performance and institutional analysis* (pp. 51–70). Cambridge, MA: MIT Press.

Del Bianco, T., Falck-Ytter, T., Thorup, E., & Gredebäck, G. (2019). The developmental origins of gaze-following in human infants. *Infancy, 24*(3), 433–454.

Denison, S., & Xu, F. (2010a). Integrating physical constraints in statistical inference by 11-month-old infants. *Cognitive Science, 34*(6), 885-908.

Denison, S., & Xu, F. (2010b). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*(5), 798-803.

Dennett, D. C. (1971). Intentional Systems. *The Journal of Philosophy, 68*(4), 87–106.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron, 80*(2), 312–325.

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. *Psychological Review, 127*(3), 455.

Eliaz, K., & Schotter, A. (2010). Paying for confidence: An experimental study of the demand for noninstrumental information. *Games and Economic Behavior, 70*(2), 304–324.

Emery, N. J., & Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature, 414*(6862), 443–446.

Feiman, R., Mody, S., & Carey, S. (2022). The development of reasoning by exclusion in infancy. *Cognitive Psychology, 135*, 101473.

Foley, N. C., Kelly, S. P., Mhatre, H., Lopes, M., & Gottlieb, J. (2017). Parietal neurons encode expected gains in instrumental information. *Proceedings of the National Academy of Sciences, 114*(16), E3315–E3323.

Fontenelle, S. A., Alexander Kahrs, B., Ashley Neal, S., Taylor Newton, A., & Lockman, J. J. (2007). Infant manual exploration of composite substrates. *Journal of Experimental Child Psychology, 98*(3), 153–167.

Foushee, R., Srinivasan, M., & Xu, F. (2023). Active learning in language development. *Current Directions in Psychological Science, 32*(3), 250–257.

Frankenhuis, W. E., & Amir, D. (2022). What is the expected human childhood? Insights from evolutionary anthropology. *Development and Psychopathology, 34*(2), 473–497.

Frankland, S. M., Webb, T. W., & Cohen, J. D. (2021). No Coincidence, George: Capacity Limits are the Curse of Compositionality. *PsyArXiv*. https://doi.org/10.31234/osf.io/cjuxb

Friedman, J. (2013). Question-directed attitudes. *Philosophical Perspectives, 27*(1), 145–174.

Geffen, S., & Mintz, T. H. (2015). Can you believe it? 12-month-olds use word order to distinguish between declaratives and polar interrogatives. *Language Learning and Development, 11*(3), 270–284.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: the naïve theory of rational action. *Trends in Cognitive Sciences, 7*(7), 287–292.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*(2), 165–193.

Gervain, J., Bonatti, L., & Filippin, L. (2009). *The PsyCode Software Manual*. http://psy.ns.sissa.it/

Ghilardi, T., Poli, F., Meyer, M., Colizoli, O., & Hunnius, S. (2024). Early roots of information-seeking: Infants predict and generalize the value of information. *eLife*, 13.

Giron, A. P., Ciranka, S., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2023). Developmental changes in exploration resemble stochastic optimization. *Nature Human Behaviour, 7*(11), 1955–1967.

Godfrey-Smith, P. (1996). *Complexity and the Function of Mind in Nature*. Cambridge: Cambridge University Press.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language, 21*, 125–155.

Good, I. J., & Card, W. I. (1971). The diagnostic process with special reference to errors. *Methods of Information in Medicine, 10*, 176–188.

Goodhue, D., Hacquard, V., & Lidz, J. (2023). 18-month-olds understand the links between declaratives and assertions, and interrogatives and questions. In *Proceedings of the Boston University Conference on Language Development (Vol. 47)*.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences, 20*(11), 818–829.

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 375*(1803), 1–10.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin, 138*(6), 1085.

Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2015). When younger learners can be better (or at least more open-minded) than older ones. *Current Directions in Psychological Science, 24*(2), 87–92.

Gottlieb, J., & Oudeyer, P. Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience, 19*(12), 758–770.

Goupil, L., Romand-Monnier, M., & Kouider, S. (2016). Infants ask for help when they know they don't know. *Proceedings of the National Academy of Sciences, 113*(13), 3492–3496.

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498.

Greenberg, R., & Mettke-Hofmann, C. (2001). Ecological aspects of neophobia and neophilia in birds. In V. Nolan Jr. & C. F. Thompson (Eds.*), Current Ornithology (Vol. 16, pp. 119–178)*. Kluwer Academic/Plenum Publishers.

Gweon, H., & Schulz, L. (2019). From exploration to instruction: Children learn from exploration and tailor their demonstrations to observers' goals and competence. *Child Development, 90*(1), e148–e164.

Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (2014). Sins of omission: Children selectively explore when teachers are under-informative. *Cognition, 132*(3), 335–341.

Gweon, H., Shafto, P., & Schulz, L. (2018). Development of children's sensitivity to overinformativeness in learning and teaching. *Developmental Psychology, 54*(11), 2113–2125.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*, B23–B34.

Hamblin, C. L. (1958). Questions. *Australasian Journal of Philosophy*, *36*(3), 159–168.

Hare, B., Call, J., Agnetta, B., & Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771–785.

Hayhoe, M. M., & Rothkopf, C. A. (2011). Vision in the natural world. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(2), 158–166.

Hebb, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychological Review, 62*, 243–254.

Hoppe, D., & Rothkopf, C. A. (2016). Learning rational temporal eye movement strategies. *Proceedings of the National Academy of Sciences, 113*(29), 8332–8337.

Hoppe, D., & Rothkopf, C. A. (2019). Multi-step planning of eye movements in visual search. *Scientific Reports, 9*(1), 144.

Huang, Z., Hu, Q., & Shao, Y. (2019). Understanding others' knowledge certainty from inference and information-seeking behaviors in children. *Developmental Psychology, 55*(7), 1372.

Hunt, J. M. (1965) Intrinsic motivation and its role in psychological development. *Nebraska Symposium on Motivation, 13*, 189–282.

Hunt, L. T., Daw, N. D., Kaanders, P., MacIver, M. A., Mugan, U., Procyk, E., ... & Kolling, N. (2021). Formalizing planning and information search in naturalistic decision-making. *Nature Neuroscience, 24*(8), 1051–1064.

Huys, Q. J., Lally, N., Faulkner, P., Eshel, N., Seifritz, E., Gershman, S. J., ... & Roiser, J. P. (2015). Interplay of approximate planning strategies. *Proceedings of the National Academy of Sciences*, 112(10), 3098–3103.

Jackendoff, R. (2007). *Language, Consciousness, Culture: Essays on Mental Structure*. Cambridge: MIT Press.

Jacquey, L., Fagard, J., Esseily, R., & O'Regan, J. K. (2020). Detection of sensorimotor contingencies in infants before the age of 1 year: A comprehensive review. *Developmental Psychology, 56*(7), 1233–1251.

Jara-Ettinger, J. (2019). Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences, 29*, 105–110.

Jara-Ettinger, J., Floyd, S., Tenenbaum, J. B., & Schulz, L. E. (2017). Children understand that agents maximize expected utilities. *Journal of Experimental Psychology: General, 146*(11), 1574–1585.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The Naïve Utility Calculus: Computational Principles Underlying Commonsense Psychology. *Trends in Cognitive Sciences, 20*(8), 589–604.

Jara-Ettinger, J., Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2015). Children's understanding of the costs and rewards underlying rational action. *Cognition, 140*, 14–23.

Jara-Ettinger, J., Tenenbaum, J. B., & Schulz, L. E. (2015). Not so innocent: Toddlers' inferences about costs and culpability. *Psychological Science, 26*(5), 633–640.

Jern, A., Lucas, C. G., & Kemp, C. (2017). People learn other people's preferences through inverse decision-making. *Cognition, 168*, 46–64.

Johnson, C. N., & Maratsos, M. P. (1977). Early comprehension of mental verbs: Think and know. *Child Development, 48*(4), 1743–1747.

Jones, A., Swaboda, N., & Ruggeri, A. (2020). Developmental Changes in Question-Asking. In L. Butler, S. Ronfard, & K. Corriveau (Eds.), *The Questioning Child: Insights from Psychology and Education (pp. 118–143)*. Cambridge: Cambridge University Press.

Kaltefleiter, L. J., Schuwerk, T., Wiesmann, C. G., Kristen-Antonow, S., Jarvers, I., & Sodian, B. (2022). Evidence for goal- and mixed evidence for false belief-based action prediction in 2-to 4-year-old children: A large-scale longitudinal anticipatory looking replication study. *Developmental Science, 25*(4), e13224.

Kampis, D., & Kovács, Á. M. (2021). Seeing the world from others' perspective: 14-month-olds show altercentric modulation effects by others' beliefs. *Open Mind, 5*, 189–207.

Kampis, D., Kármán, P., Csibra, G., Southgate, V., & Hernik, M. (2021). A two-lab direct replication attempt of Southgate, Senju and Csibra (2007). *Royal Society Open Science, 8*(8), 210190.

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. *Proceedings of the Royal Society B: Biological Sciences, 282*(1819), 20151683.

Kaplan, F., & Oudeyer, P. Y. (2007). In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience, 1*, 17.

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy, 1*(1), 3–44.

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron, 88*(3), 449–460.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS One, 7*(5), e36399.

Kidd, C., Piantadosi, S.T., & Aslin, R.N. (2014). The Goldilocks effect in infant auditory attention. *Child Development, 85*, 1795–1804.

Kim, S., Sodian, B., & Proust, J. (2020). 12-and 24-month-old infants' search behavior under informational uncertainty. *Frontiers in Psychology, 11*, 566.

Király, I., Oláh, K., Csibra, G., & Kovács, Á. M. (2018). Retrospective attribution of false beliefs in 3-year-old children. *Proceedings of the National Academy of Sciences, 115*(45), 11477–11482.

Kirkpatrick, S., Gelatt Jr, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, 220*(4598), 671–680.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18*(4), 513–549.

Kominsky, J. F. (2019) PyHab: Open-source real time infant gaze coding and stimulus presentation software. *Infant Behavior & Development, 54*, 114–119.

Kovács, Á. M., Tauzin, T., Téglás, E., Gergely, G., & Csibra, G. (2014). Pointing as epistemic request: 12-month-olds point to receive new information. *Infancy, 19*(6), 543–557.

Kovács, Á. M., Téglás, E., & Csibra, G. (2021). Can infants adopt underspecified contents into attributed beliefs? Representational prerequisites of theory of mind. *Cognition, 213*, 104640.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*(6012), 1830–1834.

Kulke, L., Reiß, M., Krist, H., & Rakoczy, H. (2018). How robust are anticipatory looking measures of Theory of Mind? Replication attempts across the life span. *Cognitive Development, 46*, 97–111.

Kulke, L., von Duhn, B., Schneider, D., & Rakoczy, H. (2018). Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychological Science, 29*(6), 888–900.

Lancy, D. F. (2018). *Anthropological Perspectives on Children as Helpers, Workers, Artisans, and Laborers*. New York: Palgrave Macmillan.

Lapidow, E., & Walker, C. M. (2020a). Informative experimentation in intuitive science: Children select and learn from their own causal interventions. *Cognition, 201*(4), 104315.

Lapidow, E., & Walker, C. M. (2020b). The search for invariance: Repeated positive testing serves the goals of causal learning. In J. B. Childers (Ed.), *Language and Concept Acquisition from Infancy Through Childhood: Learning from Multiple Exemplars (pp. 197–219)*. Springer International Publishing.

Lawson, C. A., & Rakison, D. H. (2013). Expectations about single event probabilities in the first year of life: The influence of perceptual and statistical information. *Infancy, 18*(6), 961–982.

Leahy, B., & Carey, S. E. (2020). The acquisition of modal concepts. *Trends in Cognitive Sciences, 24*(1), 65–78.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences, 43*, e1, 1–60.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics, 27*(4), 986–1005.

Liquin, E. G., Callaway, F., & Lombrozo, T. (2020). Quantifying Curiosity: A Formal Approach to Dissociating Causes of Curiosity. In S. Denison., M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 309–315). Cognitive Science Society.

Liquin, E. G., Callaway, F., & Lombrozo, T. (2021). Developmental change in what elicits curiosity. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd Annual Conference of the Cognitive Science Society*. (pp. 1360–1366). Cognitive Science Society.

Liquin, E. G., Rhodes, M, & Gureckis, T. M. (2024). Reuse and remixing in question asking across development. In L. K. Samuelson, S. Frank, M. Toneva, A. Mackey & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society* (pp. 1137–1144). Cognitive Science Society.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition, 108*(3), 732–739.

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). Twelve- and 18-month-olds point to provide information for others. *Journal of Cognition and Development, 7*(2), 173–187.

Liu, S., & Spelke, E. S. (2017). Six-month-old infants expect agents to minimize the cost of their actions. Cognition, 160, 35–42.

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science, 358*(6366), 1038–1041.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75–98.

Lövdén, M., Garzón, B., & Lindenberger, U. (2020). Human skill learning: Expansion, exploration, selection, and refinement. *Current Opinion in Behavioral Sciences, 36*, 163–168.

Lucas, C. G., Griffiths, T. L., Xu, F., Fawcett, C., Gopnik, A., Kushnir, T., ... & Hu, J. (2014). The child as econometrician: A rational model of preference understanding in children. *PloS one, 9*(3), e92160.

Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition, 105*(3), 489–512.

Luo, Y., & Johnson, S. C. (2009). Recognizing the role of perception in action at 6 months. *Developmental Science, 12*(1), 142–149.

Luo, Y., Baillargeon, R., Brueckner, L., & Munakata, Y. (2003). Reasoning about a hidden object after a delay: Evidence for robust representations in 5-month-old infants. *Cognition, 88*(3), B23–B32.

Macnamara, J., Baker, E., & Olson, C. L. (1976). Four-year-olds' understanding of "pretend", "forget", and "know": Evidence for propositional operations. Child Development, 62–70.

Margolis, E., & Laurence, S. (2011). Learning matters: The role of learning in concept acquisition. *Mind & Language, 26*(5), 507–539.

Markant, D., & Gureckis, T. (2012). Does the utility of information influence sampling behavior?. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 2434–2439). Cognitive Science Society.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*(2), 121–157.

McCall, R. B. (1974). Exploratory manipulation and play in the human infant. *Monographs of the Society for Research in Child Development, 39*(2), 1–88.

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., ... & Gonzalez, C. (2015). Unpacking the exploration-exploitation tradeoff: A synthesis of human and animal literatures. *Decision, 2*(3), 191–215.

Mettke-Hofmann, C. (2014). Cognitive ecology: ecological factors, life-styles, and cognition. *WIREs Cognitive Science, 5*(3), 345–360.

Miller, G. A. (1983). Informavores. In F. Machlup & U. Mansfield (Eds.), *The Study of Information: Interdisciplinary Messages (pp. 111–113)*. New York: Wiley.

Miller, S. A., Hardin, C. A., & Montgomery, D. E. (2003). Young children's understanding of the conditions for knowledge acquisition. *Journal of Cognition and Development, 4*(3), 325–356.

Millikan, R. G. (1989). In defense of proper functions. *Philosophy of Science, 56*(2), 288–302.

Mitchell, P., & Robinson, E. J. (1992). Children's understanding of the evidential connotation of 'know'in relation to overestimation of their own knowledge. Journal of Child Language, 19(1), 167–182.

Moore, C., & Davidge, J. (1989). The development of mental terms: Pragmatics or semantics? Journal of Child Language, 16(03), 633–641.

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development, 60*(1), 167–171.

Moran, R., Keramati, M., Dayan, P., & Dolan, R. J. (2019). Retrospective model-based inference guides model-free credit assignment. *Nature Communications, 10*(1), 1–14.

Mosher, F., & Hornsby, J. (1966). On asking questions. In J. Bruner, R. Oliver, & P. Greenfield (Eds.). *Studies in Cognitive Qrowth*. New York: Wiley.

Nelson, J. D. (2008). Towards a rational theory of human information acquisition. In M. Oaksford & N. Chater (Eds*.), The Probabilistic Mind: Prospects for Rational Models of Cognition (pp. 143–163)*. Oxford: Oxford University Press.

Nelson, J. D., Meder, B., & Jones, M. (2018). Towards a theory of heuristic and optimal planning for sequential information search. *PsyArXiv*. https://doi.org/10.31234/osf.io/bxdf4

Okasha, S. (2017). Biology and the Theory of Rationality. In D. Livingstone Smith (Ed.), *How Biology Shapes Philosophy: New Foundations for Naturalism* (pp. 141–160). Cambridge: Cambridge University Press.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*(5719), 255–258.

Papineau, D. (2001). The Evolution of Means–End Reasoning. In D. Walsh (Ed.), *Naturalism, Evolution and Mind* (pp. 145–79). Oxford: Oxford University Press.

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology, 43*(5), 1216–1226.

Pavlov, I. P. (1927). *Conditioned Reflexes*. Oxford: Clarendon Press.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods.*

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*, 195–203.

Perner, J. (2012). MiniMeta: In search of minimal criteria for metacognition. In M. J. Beran, J. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of Metacognition (pp. 94–116)*. Oxford: Oxford University Press.

Piaget, J. (1952). *The Origins of Intelligence in Children*. New York: International Universities Press. (Original work published 1937)

Pietraszewski, D., Wertz, A. E., Bryant, G. A., & Wynn, K. (2017). Three-month-old human infants use vocal cues of body size. *Proceedings of the Royal Society B: Biological Sciences, 284*(1856), 20170656.

Pillow, B. (1999). Children's understanding of inferential knowledge. *Journal of Genetic Psychology, 160*, 419–428.

Pillow, B. H. (2002). Children's and adults' evaluation of the certainty of deductive inferences, inductive inferences, and guesses. *Child Development, 73*(3), 779–792.

Pillow, B. H., & Anderson, K. L. (2006). Children's awareness of their own certainty and understanding of deduction and guessing. *British Journal of Developmental Psychology, 24*(4), 823–849.

Pillow, B. H., Hill, V., Boyce, A., & Stein, C. (2000). Understanding inference as a source of knowledge: Children's ability to evaluate the certainty of deduction, perception, and guessing. *Developmental Psychology, 36*, 169–179.

Poli, F., Ghilardi, T., Mars, R. B., Hinne, M., & Hunnius, S. (2023). Eight-month-old infants meta-learn by downweighting irrelevant evidence. *Open Mind, 7*, 141-155.

Poli, F., Meyer, M., Mars, R. B., & Hunnius, S. (2022). Contributions of expected learning progress and perceptual novelty to curiosity-driven exploration. *Cognition, 225*, 105119.

Poli, F., O'Reilly, J. X., Mars, R. B., & Hunnius, S. (2024). Curiosity and the dynamics of optimal exploration. *Trends in Cognitive Sciences, 28*(5), 441-453.

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science Advances, 6*(39), eabb5053.

Quillien, T. (2023). Rational information search in welfare-tradeoff cognition. *Cognition, 231*, 105317.

Rai, R., & Mitchell, P. (2006). Children's ability to impute inferentially based knowledge. *Child Development, 77*(4), 1081–1093.

Raz, G., & Saxe, R. (2020). Learning in infancy is active, endogenously motivated, and depends on the prefrontal cortices. *Annual Review of Developmental Psychology, 2*, 247–268.

Raz, G., & Saxe, R. (2020). Learning in infancy is active, endogenously motivated, and depends on the prefrontal cortices. *Annual Review of Developmental Psychology, 2*, 247-268.

Redshaw, J., & Suddendorf, T. (2016). Children's and apes' preparatory responses to two mutually exclusive possibilities. *Current Biology, 26*(13), 1758–1762.

Rochat, P. (1989). Object manipulation and exploration in 2- and 5-month-old infants. *Developmental Psychology, 25*(6), 871–884.

Rovee, C. K., & Rovee, D. T. (1969). Conjugate reinforcement of infant exploratory behavior. *Journal of Experimental Child Psychology, 8*(1), 33–39.

Ruggeri, A. (2022). An introduction to ecological active learning. *Current Directions in Psychological Science, 31*(6), 471–479.

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology, 69*, 181–203.

Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General, 141*(3), 433.

Schulz, A. W. (2011). Simulation, simplicity, and selection: An evolutionary perspective on high-level mindreading. *Philosophical Studies, 152*, 271–285.

Schulz, E., & Gershman, S. J. (2019). The algorithmic architecture of exploration in the human brain. *Current Opinion in Neurobiology, 55*, 7–14.

Schulz, E., Bhui, R., Love, B. C., Brier, B., Todd, M. T., & Gershman, S. J. (2019). Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences, 116*(28), 13903–13908.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for rewards like a child means less generalization and more directed exploration. *Psychological science, 30*(11), 1561–1572.

Schulz, L. (2012). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences, 16*(7), 382–389.

Schuwerk, T., Kampis, D., Baillargeon, R., Biro, S., Bohn, M., Byers-Heinlein, K., Dörrenberg, S., Fisher, C., Franchin, L., Fulcher, T., Garbisch, I., Geraci, A., Grosse Wiesmann, C. G., Hamlin, K., Haun, D. B. M., Hepach, R., Hunnius, S., Hyde, D. C., Karman, P., … Rakoczy, H. (2021). Action anticipation based on an agent's epistemic state in toddlers and adults. *PsyArXiv*. https://doi.org/10.31234/osf.io/x4jbm

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development, 80*(4), 1172–1196.

Scott, R. M., Richman, J. C., & Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning. *Cognitive Psychology, 82*, 32–56.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal, 27*(3), 379–423.

Sodian, B., & Wimmer, H. (1987). Children's understanding of inference as a source of knowledge. *Child Development, 58*, 424–433.

Sokolov, E. (1966). Orienting reflex as information regulator. In A. Leontiev, A. Luria, & S. Sminov (Eds.), *Psychological Research in the USSR (pp. 334–360)*. Moscow: Progress Publishers.

Sommer-Trembo, C., Santos, M. E., Clark, B., Werner, M., Fages, A., Matschiner, M., ... & Salzburger, W. (2024). The genetics of niche-specific behavioral tendencies in an adaptive radiation of cichlid fishes. *Science, 384*(6694), 470–475.

Song, H. J., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology, 44*(6), 1789.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology, 45*(6), 1794–1798.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*(7), 587–592.

Sperber, D. (2000). Metarepresentations in an evolutionary perspective. In Sperber, D. (Ed.), *Metarepresentations: A Multidisciplinary Perspective (pp. 117–137)*. New York, NY: Oxford University Press.

Sperber, D., & Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences, 8*(1), 40-46.

Sperber, D., & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind & Language, 17*(1&2), 3–23.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359–393.

Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin & Review, 17*(3), 412–416.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science, 348*(6230), 91–94.

Sterelny, K. (2003). *Thought in a Hostile World: The Evolution of Human Cognition*. Oxford: Blackwell.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27*(3), 453–489.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science, 18*(7), 580–586.

Tauzin, T., & Gergely, G. (2018). Communicative mind-reading in preverbal infants. *Scientific Reports, 8*(1), 1–9.

Tebbich, S., Fessl, B., & Blomqvist, D. (2009). Exploration and ecology in Darwin's finches. *Evolutionary Ecology, 23*(5), 591–605.

Téglás, E., & Bonatti, L. L. (2016). Infants anticipate probabilistic but not deterministic outcomes. *Cognition, 157*, 227-236.

Téglás, E., Girotto, V., Gonzalez, M., & Bonatti, L. L. (2007). Intuitions of probabilities shape expectations about the future at 12 months and beyond. *Proceedings of the National Academy of Sciences, 104*(48), 19156-19159.

Téglás, E., Ibanez-Lillo, A., Costa, A., & Bonatti, L. L. (2015). Numerical representations and intuitions of probabilities at 12 months. *Developmental Science, 18*(2), 183-193.

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332*(6033), 1054-1059.

Ten, A., Kaushik, P., Oudeyer, P. Y., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications, 12*(1), 5972.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331*(6022), 1279–1285.

Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental Psychology, 47*(4), 1065–1077.

Thoermer, C., Sodian, B., Vuori, M., Perst, H., & Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *British Journal of Developmental Psychology, 30*(1), 172–187.

Thorndike, E. (1911). *Animal Intelligence: Experimental Studies*. New York: MacMillan.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*(5), 675–691.

Tooby, J., & Cosmides, L. (2020). Natural Selection and the Nature of Communication. In K. Floyd & R. Weber (Eds*.), The Handbook of Communication Science and Biology (pp. 21–49)*. New York: Routledge.

Török, G., Pomiechowska, B., Csibra, G., & Sebanz, N. (2019). Rationality in joint action: Maximizing coefficiency in coordination. *Psychological Science, 30*(6), 930-941.

Turan-Küçük, E. N., & Kibbe, M. M. (2024). Three-year-olds' ability to plan for mutually exclusive future possibilities is limited primarily by their representations of possible plans, not possible events. *Cognition, 244*, 105712.

Tversky, A., & Edwards, W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology, 71*(5), 680.

Twomey, K. E., & Westermann, G. (2018). Curiosity-based learning in infants: A neurocomputational approach. *Developmental Science, 21*(4), e12629.

Ullman, T. D., & Tenenbaum, J. B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual Review of Developmental Psychology, 2*, 533–558.

Ullman, T. D., Goodman, N. D., & Tenenbaum, J. B. (2012). Theory learning as stochastic search in the language of thought. *Cognitive Development, 27*(4), 455-480.

van Rooij, I., Blokpoel, M., Kwisthout, J., & Wareham, T. (2019). *Cognition and Intractability: A guide to classical and parameterized complexity analysis*. Cambridge: Cambridge University Press.

van Rooy, R. (2003). Questioning to Resolve Decision Problems. *Linguistics and Philosophy, 26*(6), 727–763.

van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.

Velez-Ginorio, J., Siegel, M. H., Tenenbaum, J. B., & Jara-Ettinger, J. (2017). Interpreting actions by attributing compositional desires. In G. Gunzelmann, A. Howes, T. Tenbrink, and E. Davelaar (Eds). *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. Cognitive Science Society.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... & Van Mulbregt, P. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods, 17*(3), 261–272.

Vouloumanos, A., Onishi, K. H., & Pogue, A. (2012). Twelve-month-old infants recognize that speech can communicate unobservable intentions. *Proceedings of the National Academy of Sciences, 109*(32), 12933–12937.

Wagner, L., & Carey, S. (2005). 12-month-old infants represent probable endings of motion events. *Infancy, 7*(1), 73–83.

Wang, M. Z., Sweis, B., & Hayden, B. Y. (2018). A testable definition of curiosity. *IEEE. CDS Newsletter*.

Wang, Q., Bolhuis, J., Rothkopf, C. A., Kolling, T., Knopf, M., & Triesch, J. (2012). Infants in control: rapid anticipation of action outcomes in a gaze-contingent paradigm. *PloS one, 7*(2), e30884.

Wellman, H. (2011). Reinvigorating explanations for the study of early cognitive development. *Child Development Perspectives, 5*, 33–38.

White, J. K., Bromberg-Martin, E. S., Heilbronner, S. R., Zhang, K., Pai, J., Haber, S. N., & Monosov, I. E. (2019). A neural network for information seeking. *Nature Communications, 10*(1), 1–19.

Winkler, H., & Leisler, B. (1999). Exploration and curiosity in birds: functions and mechanisms. In *Proceedings of the 22nd International Ornithological Congress (pp. 915-932).*

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1–34.

Wu, S., Blanchard, T., Meschke, E., Aslin, R. N., Hayden, B. Y., & Kidd, C. (2022). Macaques preferentially attend to intermediately surprising information. *Biology Letters, 18*(7), 20220144.

Wu, Y., Merrick, M., & Gweon, H. (2024). Expecting the unexpected: infants use others' surprise to revise their own expectations. *Open Mind, 8*, 67–83.

Yang, S. C. H., Wolpert, D. M., & Lengyel, M. (2016). Theoretical perspectives on active sensing. *Current Opinion in Behavioral Sciences, 11*, 100–108.

Appendix: Supplementary modeling results for Experiment 8

In addition to the objectives described in relation to Experiment 8 in Chapter 6, we introduced yet another strategy for the Ideal Shaker agent, which we call UNCERTAINSHAKE. This objective introduces a different approach to decision-making where the agent aims to minimize its uncertainty about the auditory outcome of shaking a specific pair of boxes (whether or not a sound is heard), without considering the final trial outcomes (the location of the hidden character) and assuming that only the full box elicits a sound when shaken. The UNCERTAINSHAKE strategy leads the agent to choose the option whose auditory outcome it is most uncertain about, based on its history of previous outcomes.

For this objective, the agent maintains a separate Beta distribution for each possible action (i.e., each pair of boxes). The Beta distribution, parameterized by α and β, represents the agent's belief about the likelihood of hearing a sound when shaking a specific pair of boxes. Here, α represents the number of successful auditory outcomes (i.e., instances where a sound was heard), while β represents the number of unsuccessful outcomes (no sound heard). Initially, the agent's belief is modeled with a uniform prior, with both α and β set to 1. This uniform prior reflects an initial assumption that both outcomes (sound or no sound) are equally likely. As the agent progresses through trials, it updates the Beta distribution parameters based on the observed outcomes using the following update rule:

$$\alpha_{ij} \leftarrow \alpha_{ij} + 1 \quad \text{if sound is heard (success)}$$

$$\beta_{ij} \leftarrow \beta_{ij} + 1 \quad \text{if no sound is heard (failure)}$$

The UNCERTAINSHAKE strategy leverages the entropy of the binary outcome (sound or no sound) of each option to guide behavior. The entropy is highest ($H = 1$) when the probability of hearing a sound, p, is closest to 0.5, reflecting maximum uncertainty about the outcome. For each action (i.e., each pair of boxes), the entropy is calculated as:

$$H(\text{outcome}) = -p \log_2(p) - (1-p)\log_2(1-p)$$

where $p = \frac{\alpha}{\alpha+\beta}$ represents the agent's current belief about the probability of hearing a sound for that pair of boxes.

The agent aims to reduce this uncertainty by selecting the action (pair of boxes) with the highest entropy. Therefore, the expected value for the UNCERTAINSHAKE objective is the given action's entropy value. In each trial, the agent selects its action (pair of boxes) using the same softmax policy as it did under all the other objectives.

The results the model fitting and parameter estimation are summarized at Table S1. The mean fitted τ (temperature) parameter was 0.75, similar in range to the previous models. When comparing to previous models, UNCERTAINSHAKE was comparable to the other two best fitting candidates, PASTINFO and EXPINFO, across all metrics. However, when comparing the relative evidence for pastinfo versus UNCERTAINSHAKE, the Bayes Factor (using the models' BIC values) indicated strong evidence in favor of the PASTINFO model (BF = 23.31).

	UNCERTAINSHAKE
-LL	430.62
pseudo-R^2	0.25
# parameters	1
BIC	868.00
τ^*	0.75 ± 0.34

Table S1. Quality of behavioral fits from the 36 subjects in Experiment 8, for the UNCERTAINSHAKE model. -LL: negative log-likelihood. BIC: Bayesian Information Criterion. τ^*: estimated temperature parameter mean ± SD.

To qualitatively evaluate the model's predictions, we simulated the behavior of 100 synthetic participants for each real participant under the UNCERTAINSHAKE objective, using empirically derived parameters from infants. Figure S1 displays the predicted dynamics of the proportion of AC choices over time, based on the best-fitting parameters. The figure shows that, for most participants, the model predicts a very gradual shift towards choosing the AC option as the experiment progresses. The bottom panel of the figure reveals the underlying cause: the difference between the learned probabilities of sound outcomes increases slowly, with the most significant differences appearing in the final six trials. This gradual change is likely due to frequent trials with box A outcomes, where shaking both pairs of boxes results in a shaking sound. Despite the competitive model fit metrics, the predicted choice pattern of UNCERTAINSHAKE fails to capture the more abrupt increase in AC choices observed in infants' behavior during the middle and early second part of the session. Additionally, it does not reflect the temporary, sharp decrease in AC choices seen during consecutive A outcome trials around trials 15 to 17.

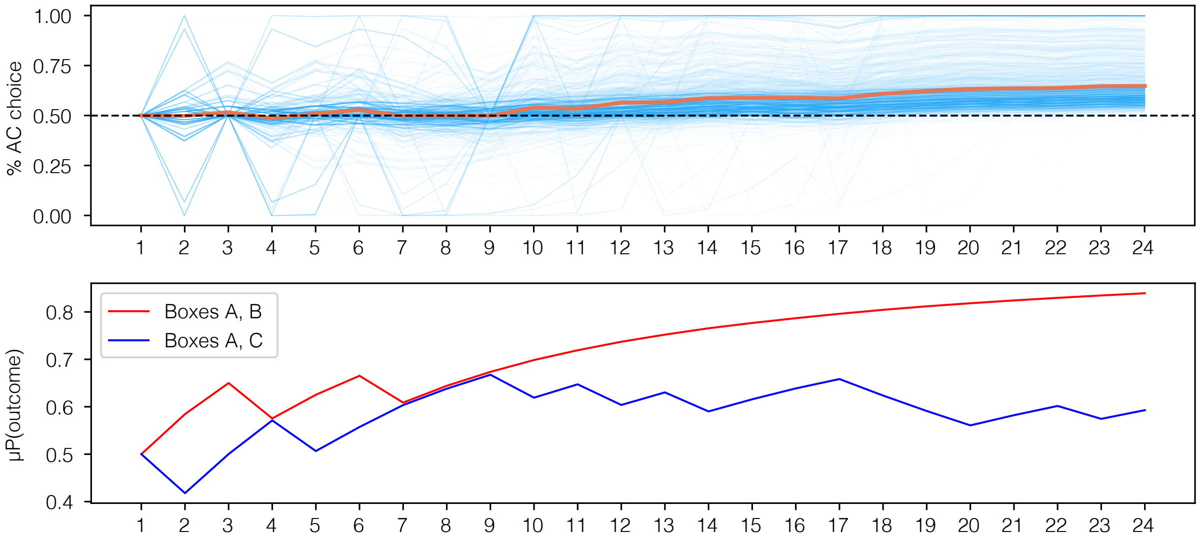

Figure S1. Predicted dynamics of AC choices and underlying learned probabilities in the UNCERTAINSHAKE model. The top panel depicts the simulated predicted proportion of AC choices across 100 synthetic participants per real participant, simulated under the UNCERTAINSHAKE objective, with the empirically derived parameters from infant participants. Blue lines represent the difference in softmax probabilities for choosing the AC option across trials for each participant, while the red line represent the mean of individual softmax probabilities. The model predicts a gradual increase in the likelihood of selecting the AC option as the experiment progresses, reflecting the slow differentiation in learned probabilities. The bottom panel shows the evolution of the estimated probabilities for sound outcomes associated with box A and box C over the course of the experiment. The mean trajectories indicate that the difference in learned probabilities tends to grow slowly, with the most significant divergence occurring in the final six trials. Despite the model's fit to the overall data, it fails to capture the more abrupt shifts in AC choices observed in infants.