Technische Universität München
TUM School of Social Sciences and Technology

Prerequisites for Self-Guided Learning: How Children Adapt to Task Structures, Goals, and Their Abilities

Daniil L. Serko, M.Sc.

Vollständiger Abdruck der von der TUM School of Social Sciences and Technology der Technischen Universität München zur Erlangung eines Doktors der Philosophie (Dr. phil.) genehmigten Dissertation.

Vorsitz: Prof. Dr. Andreas Obersteiner

Prüfende der Dissertation:

1. Prof. Dr. Dr. Azzurra Ruggeri
2. Assoc. Prof. Dr. Patricia Kanngiesser

Die Dissertation wurde am 10.04.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Social Sciences and Technology am 30.09.2024 angenommen.

Abstract

Children's capacity to self-direct their learning has been the subject of extensive debate among psychologists, educators, and philosophers. Despite extensive research, there appears to be little consensus among researchers on whether and from what age children can make adaptive learning decisions central to self-guided learning. My thesis presents behavioral experiments conducted with children between the ages of 2 and 8 that explore three questions at the core of this controversy: 1) Do children adapt their learning to the *task structures* they encounter? 2) Do they adapt their learning to achieve different *goals*? and 3) Can they adapt their learning decisions to their *abilities*?

In the first chapter, I introduce the theoretical framework and research findings that contextualize my empirical work. In the experiments, presented in the following chapters, I systematically expand the breadth of the research questions introduced above. In Chapter 2, I investigate children's exploration when facing causal relations that vary across contexts. The results indicate that, with age, children's exploration becomes increasingly attuned to the causal *structures* they face. In Chapter 3, I explore children's goal-directed information search. The results demonstrate that even 2-year-olds can successfully adapt their information searches in response to the *task structure* and their *goals*. In Chapter 4, I investigate children's abilities to make adaptive practice decisions when facing uncertainty about a future goal. Here, the results show that with scaffolding, even 4-year-olds prepare themselves for a future test situation, underscoring their ability to integrate information about *task structures*, *goals*, and their *abilities*.

Taken together, the experiments presented in this thesis suggest that, when tested with age-appropriate paradigms, even preschoolers are capable of directing their learning paths, but they might rely on scaffolding to succeed. Thus, this thesis contributes to resolving the longstanding debate about children's self-guided learning by suggesting that previous divergent findings regarding the developmental trajectory of self-guided learning abilities may stem from methodological choices, and underscore the importance of age-appropriate measures to assess children's cognitive capacities. More broadly, this thesis enhances cognitive developmental research by demonstrating that young children are capable of making adaptive learning decisions in diverse contexts, thereby having significant implications for refining educational practices to better support young learners.

Zusammenfassung

Die Fähigkeit von Kindern, ihr Lernen selbst zu steuern, ist Gegenstand umfangreicher Debatten unter Psychologinnen, Pädagoginnen und Philosophinnen. Trotz zahlreicher Forschungsarbeiten scheint es bisweilen keinen Konsens darüber zu geben, ob und ab welchem Alter Kinder adaptive Lernentscheidungen treffen können, die für selbstbestimmtes Lernen zentral sind. In meiner Dissertation stelle ich Verhaltensexperimente vor, die mit Kindern im Alter von 2 bis 8 Jahren durchgeführt wurden und die drei Kernfragen dieser Kontroverse untersuchen: 1) Passen Kinder ihr Lernen an unterschiedliche *Eigenschaften von Aufgaben* an? 2) Passen sie ihr Lernen an, um unterschiedliche *Ziele* zu erreichen? und 3) Können sie ihre Lernentscheidungen an *ihre Fähigkeiten* anpassen?

Im ersten Kapitel stelle ich den theoretischen Rahmen und die Forschungsergebnisse vor, die meine empirische Arbeit kontextualisieren. In den Experimenten, die in den folgenden Kapiteln vorgestellt werden, beleuchte ich die oben eingeführten Forschungsfragen aus unterschiedlichen Perspektiven. In Kapitel 2 untersuche ich das Explorationsverhalten von Kindern, wenn sie mit kausalen Zusammenhängen konfrontiert werden, die in verschiedenen Kontexten variieren. Die Ergebnisse deuten darauf hin, dass sich das Explorationverhalten von Kindern mit zunehmendem Alter immer mehr an die Aufgabeneigenschaften anpasst. In Kapitel 3 untersuche ich die zielgerichtete Informationssuche von Kindern. Die Ergebnisse zeigen, dass bereits 2-Jährige ihre Informationssuche erfolgreich an Aufgabeneigenschaften und Ziele anpassen können. In Kapitel 4 untersuche ich die Fähigkeit von Kindern, adaptive Übungsentscheidungen zu treffen. Hier zeigen die Ergebnisse, dass sich sogar 4-Jährige auf eine zukünftige Testsituation vorbereiten, wenn sie mit altersgerechten Studienmaterialen und Designs präsentiert werden. Die Ergebnisse unterstreichen die Fähigkeit von Kindern, Informationen über Eigenschaften von Aufgaben, Ziele und ihre eigenen Fähigkeiten zusammenzubringen und somit adaptive Lernentscheidungen zu treffen.

Zusammengefasst deuten die in dieser Arbeit vorgestellten Experimente darauf hin, dass selbst Vorschulkinder, in der Lage sind, selbstständige Lernentscheidungen zu treffen. Hierfür sind sie jedoch auf Unterstützung in Form von altersgerechten Studiendesigns und Materialien angewiesen. Meine Arbeit leistet somit einen Beitrag zur Debatte über die Fähigkeiten von Kindern, ihr Lernen selbst zu bestimmen. Insbesondere unterstreichen meine Ergebnisse die Bedeutung altersgerechter Materialien und Designs zur Erfassung der kognitiven Fähigkeiten von Kindern und legt nahe, dass frühere abweichende Forschungsergebnisse über selbstbestimmtes Lernen von Kindern möglicherweise auf methodische Entscheidungen in den Studiendurchführungen zurückzuführen sind. Im weiteren Sinne trägt diese Arbeit dazu bei, besser zu verstehen, wie junge Kinder in unterschiedlichen Kontexten adaptive und selbstbestimmte Lernentscheidungen treffen. Entsprechend haben die hier vorgestellten Ergebnisse eine erhebliche Relevanz für die Verbesserung pädagogischer Praxis, welche darauf abzielt junge Lernende zu unterstützen.

Acknowledgements

Every so often, when reading a paper, formulating my ideas, or designing a new study, I am struck by a realization that leaves me incredibly joyful, thinking: *"Wow, this is actually my job; I am a researcher, and I even get paid to do this!"* This would not have been possible without the support and help of some amazing people whom I would like to thank on the following pages.

First and foremost, I want to thank my parents, Leonard and Tatiana. Their decision to immigrate from Kyiv to Munich in May 1992, instead of staying in Ukraine or moving to Israel, has profoundly shaped my life. Had they decided otherwise and not endured the rigors of migration, my life would be profoundly different and I would likely be far from the privilege of spending my days unraveling the workings of the human mind.[1]

Being joyful about my work is largely due to my supervisor, Azzurra Ruggeri, who is everything at once: a brilliant researcher and an amazing mentor. Thank you, Azzu, for trusting me with projects, for your advice on conducting research, and for your support with everything from writing papers and analyzing data to giving talks at conferences. Beyond your role as a supervisor, you have cultivated an extraordinary lab atmosphere, and I am honored to be part of it.

Speaking of the iSearch lab, I want to thank all my colleagues for the enriching conversations, exchanges of ideas, and helpful advice I continually receive from you. While each lab member contributes to making iSearch an incredible place, I particularly want to thank my former lab members: Andreas Domberg, for his mentorship early in my PhD, and Nora Swaboda, for our productive collaboration. I also want to express my gratitude to my current lab members: Yilin Li, for sharing her expertise in infant research with me; Oana Stanciu, for inspiring and supporting me in writing and coding; and Hande Melis Altunay, for her enormous contribution to kick-starting the iSearch lab in Munich. A big thank you goes to Georgina Török as well; working with you on a common project is exciting, and I am grateful for your insightful feedback on my work.

I want to thank Julia Leonard, whose lab I had the pleasure of visiting as a guest researcher in 2022. Julia, your vast knowledge of children's cognition and your invaluable advice throughout our collaborative project (presented in Chapter 4 of this dissertation) have been tremendously helpful. I also extend my gratitude to all members of the Leonard Learning Lab I had the pleasure to meet—Mika, Reut, Melissa, Brandon, and Flora. You made my time there truly memorable.

I owe a big thank you to Ny Vasil for being an exceptional scientist and for inspiring me to think, write, and research the complexities of human thought. My gratitude also extends to Samuel Ronfard for sharing his expertise and knowledge in study design and academic writing, and for always being approachable and ready to assist with our joint project on intellectual humility.

[1] Translation to Russian: Прежде всего, я хочу выразить благодарность своим родителям, Леонарду и Татьяне. Ваше решение переехать из Киева в Мюнхен в мае 1992 года, вместо того чтобы остаться на Украине или эмигрировать в Израиль, радикально изменило мою жизнь. Благодаря вашему выбору и усилиям по адаптации в новой стране, моя жизнь приобрела совершенно иной оборот, и иначе я вряд ли смог бы посвятить себя изучению детской психологии.

I might have never entered academia without the support from my former supervisor, Patricia Kanngießer. Patricia, thank you for providing me with a candid overview of academic life and for supporting my decision to pursue an academic career. Similarly, I owe a debt of gratitude to Katja Liebal. Katja, your tremendously helpful advice on taking my first steps toward becoming a researcher motivated me to write my first bold email to my supervisor, Azzurra, asking if she would be willing to accept me as a PhD student—and it turns out she said yes!

Speaking of feedback and advice, Rosie Aboody definitely deserves a mention. It is always an absolute pleasure to exchange ideas with you, and my work would be three steps behind without your excellent feedback on my writing and research ideas. With regards to feedback, I want to thank Anna Thoma. Anna, your comments on my writing have been incredibly helpful for improving clarity and truly expressing what I intend to say. Thank you for that! I also want to express my gratitude to Jan Engelmann for his helpful advice on navigating academia as a young researcher.

A special thanks goes to my brother Konstantin and his family. Konny, you motivated me to think deeply about things and here I am submitting my dissertation. I want to thank my nieces Valerie, and Amelie for sharing their views on the world with me and for being amazing human beings. Valerie thank you for your fruitful advice on designing the study presented in Chapter 2.

I am fortunate to call some of the most amazing people in the world my friends, all of whom have accompanied and supported me in different ways throughout my life. What would all intellectual endeavors be worth if I could not share my thoughts with you, ask for your advice, or simply enjoy life with you?

I want to begin by thanking Alex for always always always helping me with everything and for your ability to elucidate even the most complex (Hegelian) thoughts to me. At this point, I need to thank Maria and Romino for generously hosting me every so often in Munich. Max H.—Братан!—thank you for a range of experiences, from intellectual debates to unforgettable parties, and for mastering both the talk and the walk. Luise, thank you for the small and thoughtful presents every now and then, and for knowing how to ask the right questions at the right moments, thereby helping me to understand the complexities and simplicities of life. Thank you, thank you, thank you, Max D., for the lavish dinners, surprises, and above all, your immense patience in assisting me with programming my projects. Without you, I would have long lost my sanity while trying to program experiments, and I would probably be storing my ideas by typing on an old typewriter instead of having a digital Zettelkasten. Premil, I want to thank you for reminding me that rationality and enlightenment, devoid of love and vulnerability, risk becoming their antithesis.

None of the thoughts I have would be the way they are without my former—and continuing—philosophy teacher, Bernhard. Thank you for posing the right questions, providing me with the right books, and patiently explaining complex thoughts to me. You contributed significantly to my becoming a curious thinker rather than a cynical one. Tantschik! Thank you for helping me discover aspects of my own biography and history by sharing your insights into philosophy, politics, martial arts, and the Soviet Union with me. Zozan, I need to thank you for uniquely combining knowledge and love. Till, I want to thank you for sharing your knowledge and providing feedback on my ideas. Thanks to both of you for the memorable dinners, deep conversations, and the ability to enjoy the immediate joys of life without losing sight of the bigger picture. Nico, I feel like I can always reach out to you, and you would always support me, be it with academia or life's various challenges. Thank you for that, brother! And Margit, thank you for sharing your life experiences and perspectives with me. In reflecting on support, I need to extend my gratitude to Katha. Thank you for your constant help and for always being ready to share with me, be it money, lunch, or your most recent musical discoveries. Hagen! Hagen! Hagen! Thank you! Beyond just being an amazing person, you inspired me to approach questions scientifically. Your work with primates and young children has been an incredible source of inspiration for me. Chris, thank you for demonstrating that a perfect harmony of fine taste in literature, food, architecture, and intellect can exist in one person. I am grateful for our friendship that has lasted for about half of my life.

Speaking of old friendships, I want to thank Sören for being a role model of an amazing father and for holding my hand when I need it. Timmy, even though you are in the “Strudel” right now, I know you will come back stronger than ever. Special thanks to your daughter Afra for her contributions to piloting the study presented in Chapter 2. Janu, you have been part of my journey since I was 13 years old—that’s more than 20 years! Thank you for being my Zotero, Obsidian, and LaTeX buddy and for your support with these tools. I am so happy you are embarking on your PhD journey soon! At this point, I also need to thank you, Matteo; besides being Janu’s amazing husband, you make me feel like part of the family whenever I visit. Thede and Anastasia, thank you for entrusting me with the godfather role for the unique Kiril and for offering advice and help whenever needed. Theresa, thank you for demystifying the conundrums of Holzkamp’s complicated words with me and for exemplifying that balancing parenthood and other life pursuits is no contradiction. Lukas, thank you for pushing me to think critically, sharing your knowledge about the human body and parenthood, and for being my friend for over 20 years. Jerome, I owe you for making my arrival in Berlin ten years ago so welcoming; and a special thanks to your son, Marek, for his contributions to the pilot study presented in Chapter 4. Laura, I want to extend my thanks for the many spontaneous walks and meetings in what was once my new neighborhood, and for sharing your insights into politics. Robert, thank you for being the best interim flatmate Munich has to offer. Pola, thank you for always creating an atmosphere that makes me feel I can share whatever I need. Denes, thank you for being an outstanding training partner and an amazingly helpful human being.

And finally, a heartfelt thank you to Hannah—for everything. I did not even know that this exists.

Contents

List of Figures

List of Tables

Chapter 1

General Introduction

Can children direct their own learning? This question—central to psychology, education, and philosophy—remains unresolved, yet, the debates revolving around it have profound implications for young learners. Consider, for instance, the widespread conception that young children lack the necessary cognitive abilities to autonomously navigate their own learning. Proponents of this view support a more guided approach in which children are explicitly directed and instructed by teachers. Supporters of an alternative view emphasize that from early on children are equipped with the necessary tools to make learning decisions by themselves. These proponents advocate for a more self-guided approach in which children can individually decide what and when they want to learn. Although this decades—if not century—old debate calls for an empirical answer, research has not yet converged on a clear agreement as to whether and from what age children can make adaptive learning decisions across contexts central to engage in self-guided learning.

My thesis aims to contribute to filling this research gap by investigating children's learning behavior concerning *three* key dimensions essential for successful self-guided learning. In particular, I investigate children's ability to make adaptive learning decisions in light of different *task structures*, *goals*, and to their *abilities*. By examining children's learning decisions with regards to these specific dimensions, rather than seeking to answer the question of children's self-guided learning in a binary way (i.e., whether children can self-guide their learning or not), my research offers a more nuanced understanding of the processes fundamental for their self-guided learning. This approach will enhance our understanding of the cognitive processes at the core of self-guided learning and can therefore help us gain more clarity about what we can expect children to master by themselves and where children may benefit from help from others and, therefore, contribute to solving the old debate between proponents of guided vs. self-guided learning.

Having presented the aim of this thesis in the preceding paragraph, and before delving into the details of what I mean by adapting one's learning to *task structures*, *goals*, and an individual's *abilities*, I want to clarify a central term for this thesis: *adaptation*. Adaptations can manifest in various forms, and assessing them in absolute terms—that is, evaluating adaptive behaviors against a singular, fixed standard—may be impossible or misleading (Gigerenzer & Selten, 2001; Todd & Gigerenzer, 2012). Instead, Todd and Gigerenzer (2012) propose to view adaptation as a match between an individual's cognitive abilities and the specific requirements that are imposed onto them by their environment. Behavior that is adaptive in some situations (e.g., thinking through a subject in depth—devoting hours, weeks, or even years to a single topic), may be non-adaptive in other situations (e.g., when we have to make quick decisions under pressure). Importantly, this view underscores that even when cognitive resources are imperfect or constrained, individuals can still successfully adapt to meet the challenges presented by their surroundings. The target population of this thesis—young children—makes such an emphasis particularly crucial because their cognitive capacities are still developing. For a long time, young children's cognition was seen as inferior and non-adaptive as compared to adults (Piaget, 1930). However, recent approaches have suggested that it may be more fruitful to perceive children's behavior as context specific adaptations rather than interpreting their behaviors in absolute terms, or comparing them with adult performance as a benchmark (Ruggeri, 2022).

Understanding adaptation as a match between an individual's abilities and the specific requirements imposed by their environment (Todd & Gigerenzer, 2012) suggests approaching adaptation from more than one perspective. Therefore, I suggest to investigate three dimensions (although there may be more) fundamental for self-guided learning: *task structures*, *goals*, and an individual's *abilities*. To understand this better, consider, for instance, a child entering school education. To thrive in this new environment they must adapt their behaviors to those three broad dimensions (task structures, goals, and their abilities). First, they must understand the *task structures*—or bundles of characteristics relevant for specific tasks—they face (see, e.g., Ruggeri & Lombrozo, 2015). For a child just entering school, this might involve understanding the routine of attending lectures on designated days and recognizing that every school day follows a specific schedule with designated classes and recess times. It also includes understanding that some subjects are more challenging than others and that the likelihood of certain events, such as changes in the timetable, visits from guest speakers, or outdoor activities, may vary depending on the context or day. Second, but equally important for the student's success, adaptation to specific *goals* is crucial. For example, the immediate goal of joining a group of friends during recess, or longer-term goals like learning to read. Each goal will require distinct actions and adjustments from the child. Finally, beyond adaptation to task structures and goals, the student's success will depend on a third dimension: the student's abilities. Our hypothetical student will not adapt to task structures and goals in the abstract space, but rather from the concrete standpoint of their own self. In other words, to navigate their studies efficiently, the student must direct their adaptation skills from the world surrounding them to their own self by reflecting on their own abilities, strengths, and weaknesses and then drawing conclusions about their practice decisions based on these reflections. This can involve understanding that they may need more practice with reading than with math, or recognizing when they are uncertain about something. Adaptation to these three dimensions (task structures, goals, and their own abilities) will equip the child to make adaptive and self-guided decisions about how they wish to navigate their own learning. Although I treat these dimensions as distinct, it would be overly simplified to view them as entirely separated. As the above example suggests a child who adapts to a goal does so equipped with individual abilities that will influence their adaptation. Similarly, every goal a person pursues is embedded in specific structures that may facilitate or complicate reaching it. With this in mind, my dissertation seeks to provide answers to the following three research questions:

1. Do children adapt their learning to the *task structures* they encounter?
2. Do children adapt their learning to achieve different *goals*?
3. Can children adapt their learning to their *abilities*?

I will approach these questions by presenting three empirical studies that address the three respective areas of adaptation (task structures, goals, individual abilities) to different extends seeking to provide a cohesive picture of children's abilities to direct their own learning. In particular, I begin by examining children's adaptation to *task structures* in Chapter 2. In a free exploration game, where children are facing no explicit goals, I study how they adapt their exploration to causal relations that can vary across different contexts. In Chapter 3, I then add the dimension of *goals* to the investigation. Here, I study how children assess the informativeness of different information when facing the explicit goal of finding a hidden present. In Chapter 4, I then continue to integrate all of the three dimensions by studying children's capacity to prepare themselves for the future through deliberate practice based on an assessment of the *task structures*, *goals*, and their *abilities*. Thereby, these studies provide empirical evidence about whether and when we can expect children to be able to direct their own learning and when they might profit from scaffolding. Before I start presenting my own empirical work, I will review the relevant literature on children's adaption to task structures, goals, and their abilities that will theoretically embed my empirical studies.

1.1 Adaptation to Task Structures

In this chapter, I examine the existing empirical evidence concerning the question: *How do children adapt their behavior to different task structures?* Similar to the example of a child entering school provided above, I focus on various facets of adaptation to task structures, including how children respond to the different likelihood of events to occur, evaluate the informativeness of available information and potential questions, and how they estimate the difficulty of the tasks they face.

The available literature presents a complex picture: Infants show remarkable abilities to adapt their actions in response to environmental cues, suggesting an early emerging capacity for adaptive learning (Begus & Southgate, 2012; Goupil et al., 2016; Kovács et al., 2014; Leonard et al., 2017; Perez & Feigenson, 2020; Sim & Xu, 2017; Stahl & Feigenson, 2015; Téglás et al., 2007; Xu & Garcia, 2008). Yet, as children enter preschool age, the available evidence becomes less clear. Some studies show that these early present abilities do not linearly progress as children mature (Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018; Ciranka & van den Bos, 2021; Schulze & Hertwig, 2021), while other studies emphasize the sophisticated nature of their adaptations during these years (Leonard et al., 2020; Magid et al., 2018; Ruggeri et al., 2019; Sumner et al., 2019; Swaboda et al., 2022; Wang & Bonawitz, 2022).

In the early stages of developmental psychology, infants' behavior was mostly perceived as unintentional and non-directed (see, e.g., Piaget, 1952). This view contrasts strongly with more recent evidence that has highlighted how even infants adapt their behaviors to specific features of their environment including object properties, probabilities of events, and the knowledge of others (Begus & Southgate, 2012; Goupil et al., 2016; Kovács et al., 2014; Perez & Feigenson, 2020; Sim & Xu, 2017; Stahl & Feigenson, 2015; Téglás et al., 2007; Xu & Garcia, 2008). Much of this research is based on the idea that from a young age, children possess expectations about their environment and notice when these expectations are violated (Gopnik et al., 1999; L. Schulz, 2015; Spelke et al., 1992). For example, at the age of 11 months, infants' exploration varies significantly when they encounter surprising and unexpected events that violate their prior expectations (Sim & Xu, 2017; Stahl & Feigenson, 2015; for an overview, see Schulz, 2015). Stahl and Feigenson (2015) investigated infants' exploration behaviors after witnessing events that violated physical principles. For instance, they saw how an object slid over a gap that, under laws of gravity, would have forced the object to fall down. Infants who saw such improbable events were more likely to engage with the target object compared to a novel distractor object. Importantly, the infants' explorations aimed at testing the expectations they held about the objects involved in the scenes. They were more likely to drop objects that seemed to float over a gap and to knock objects against the floor that appeared to have passed through a wall. A different study, using a similar design, but providing children with a reasonable explanation for the observed phenomena (e.g., children saw a hole in the wall through which the object could have passed) found that exploration behavior was significantly reduced (Perez & Feigenson, 2020). Together these studies demonstrate how young infants are not simply passive observers, but that their active exploration is guided by expectations they hold about their environment.

Building on the understanding that infants adapt their exploration actions based on their expectations, subsequent research utilizing infants' looking time as an indicator of their increased interest and attention has expanded these findings. This work highlights their rudimentary awareness of probabilities when observing different events and indicates that they form future-directed expectations based on what they observe (Téglás et al., 2007; Xu & Garcia, 2008). For example, one seminal study (Xu & Garcia, 2008) showed 8-month-old infants a concealed container from which an experimenter sampled red and white balls. The experimenters then uncovered the container and revealed the total distribution of red and white balls inside of it. Infants looked significantly longer at the container when the drawn sample was unlikely given its content (i.e., the container contained mostly red balls, whereas the sample consisted of mostly white balls) compared to when the drawn sample and the content of the container aligned (i.e., the container contained mostly red balls and the sample consisted of mostly red balls; Xu & Garcia, 2008). On a similar note, Téglás et al. (2007) presented 12-month-old infants with short movies in which four objects moved around a container with an opening on its bottom. Importantly, three of the objects were identical in color and shape, whereas one object looked different from the other three. When the scene was briefly occluded and one object fell out of the container, children looked longer at the scene when this object was the rare object and not one of the three identical objects. These findings highlight not only that infants' attention is attuned to events defying their expectations, but also their adaptive responses, underlining their rudimentary yet early emerging awareness of the likelihood of events occurring.

Another example of infants' behavioral adaptation to the task structures is their use of pointing gestures (Begus & Southgate, 2012; Goupil et al., 2016; Kovács et al., 2014). Research in this area highlights how infants adapt their information-search behavior to the knowledgeability of their prospective partners and selectively request information from potentially knowledgeable partners. For example, 12-month-olds use more pointing gestures when an experimenter uses new words (e.g., "this is a dax") to label familiar items with unfamiliar attributes (e.g., a cat wearing shoes) than when the experimenter labeled familiar objects using familiar words (e.g., "this is a kitty"; Kovács et al., 2014). Similarly, 16-month-old infants are more likely to point at an unfamiliar object when they interact with someone who has previously labeled known objects correctly, compared to when interacting with someone who has labeled familiar objects incorrectly (Begus & Southgate, 2012), suggesting that infants compare other people's state of knowledge with their own and use their pointing to selectively request epistemic information.

However, despite the early presence of some of the reported competencies of young infants, research with preschool- and school-aged children paints a less clear picture of their abilities to adapt to task structures. In fact, a substantial body of work indicates that infants' early abilities do not directly translate into more sophisticated adaptations as they mature (Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018; Ciranka & van den Bos, 2021; Gopnik et al., 2017; Schulze & Hertwig, 2021). The work I will review next, suggests that young children often struggle to efficiently navigate their explorative efforts (Giron et al., 2023; Meder et al., 2021; E. Schulz & Gershman, 2019), that they often fail to integrate probabilistic information into their decisions (Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018), and that they have difficulties asking and selecting informative questions (Herwig, 1982; Ronfard et al., 2018; Ruggeri & Feufel, 2015; Ruggeri et al., 2016; Ruggeri et al., 2017).

For example, research on exploration has highlighted how children's behavioral adaptations change significantly as they mature (Ciranka & van den Bos, 2021; Gopnik et al., 2017). Studies comparing adults' and children's exploration using tablet-based grid games, where participants can click on a grid to discover rewards, found notable differences in behavior between these age groups (Giron et al., 2023; Meder et al., 2021; E. Schulz & Gershman, 2019). In these games, rewards are spatially correlated, with high-reward options adjoining other high-reward options, and low-reward options adjoining other low-reward options. Adults tend to discover high-reward options and quickly generalize, focusing on exploiting these options without much random exploration (Wu et al., 2020; Wu et al., 2018). In contrast, children engage in uncertainty-directed and random exploration, with the amount of directed exploration increasing with age and becoming more focused on specific goals (Bonawitz et al., 2014; Giron et al., 2023; Meder et al., 2021; E. Schulz & Gershman, 2019). This observation—where young children's exploration transitions from random to increasingly goal-directed and outcome-focused as they mature—has been described as a *cooling off process* (Ciranka & van den Bos, 2021; Giron et al., 2023; Gopnik et al., 2017; Meder et al., 2021).

Similarly, research has indicated that children often struggle to integrate probabilistic information into their decisions (Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018). For instance, in a study by Betsch et al. (2018), 5- to 10-year-old children and adult participants were presented with an information board game in which they had to find a hidden treasure in one of several potential locations. Participants first learned that three animals could predict the location of a treasure with different accuracies over several rounds. The children learned that two animals had lower probabilities of correctly predicting the location, while one animal was more likely to correctly predict the location. The researchers then measured whether children based their decisions on the animal that was more likely to make correct suggestions. The results indicated that preschoolers, unlike school-aged children and adults, failed to integrate probabilistic information into their information searches. This information-search ability improved with age, such that elementary school-age children eventually used the available information comparably to adults (Betsch et al., 2018). Related to these findings, research has found that children tend to be overly optimistic about the probabilities of future outcomes when they desire these to occur, an observation termed "wishful thinking". I will revisit this concept in Chapter 4, discussing the findings I obtained (see, e.g., Bernard et al., 2016; Wente et al., 2019).

Research on question asking offers another example of children's still-developing abilities to adapt their learning to task structures, finding that children's ability to select and ask informative questions undergoes substantial improvements (for an overview, see De Simone & Ruggeri, 2022; Ronfard et al., 2018). Although research on children's selective trust highlights that they track the knowledgeability and trustworthiness of potential information partners from preschool age (for an overview, see Harris et al., 2018), it takes them until the end of preschool to *also ask informative questions*. For example, Mills et al. (2010) showed that 3-year-olds failed to generate questions that would allow them to disambiguate the location of a hidden novel toy and at the same time choose whom to ask where the toy is hidden; 4-year-olds were able to select one of two informants who was more likely to know the location of the toy, but they still struggled to ask them informative questions. Only at the age of 5 did children begin to select the right informant *and* ask informative questions to find the target toy (Fitneva et al., 2013; Mills & Landrum, 2016; Mills et al., 2011).

Similarly, research on the efficiency of children's question asking suggests that young children often fail to select the most informative question (Herwig, 1982; Ruggeri & Feufel, 2015; Ruggeri et al., 2016; Ruggeri et al., 2017, for an overview, see Ronfard et al. 2018). Such research often utilizes the 20-questions game, in which children must disambiguate a target item from multiple possible options by asking as few yes/no questions. The efficiency of participants' questions depends on the hypothesis space participants are facing. Suppose that half of the items in the game are birds and the rest are other animals: then a good question would be "Can the animal fly?", because this single question can reduce the hypothesis space by 50%. Conversely, asking "Is it the red parrot?" targets only one specific animal and does not provide as much clarity given the above situation. However, referring back to our definition of adaptation, whether a question is adaptive also depends on the information structure one faces (i.e., the likelihood of the options). If all options are equally likely, as in the above example, then it makes sense to ask more general questions that help reduce the hypothesis space. However, if some options are more likely than others (e.g., if there are multiple red birds) then it can be more advisable to use more specific questions (e.g., “Is it one of the red birds?”; Ruggeri et al., 2017). Research has indicated that preschool children struggle to select the most informative question types in 20-questions games and while by age 5 they can choose which of two questions is more informative (Ruggeri et al., 2017), it takes them until age 10 to begin consistently employing effective question-asking strategies (Herwig, 1982; Ruggeri & Feufel, 2015; Ruggeri et al., 2016). However, by age 6, children begin to ask questions that allow them to rule out large parts of the hypothesis space (Ruggeri & Feufel, 2015) and by age 7 they adapt their questions to the feedback they receive (Ruggeri & Lombrozo, 2015).

So far, the reviewed literature suggests that although infants possess some remarkable abilities to adapt their behaviors to task structures (Begus & Southgate, 2012; Goupil et al., 2016; Kovács et al., 2014; Perez & Feigenson, 2020; Stahl & Feigenson, 2015; Téglás et al., 2007; Xu & Garcia, 2008), older children’s skills in exploration (Bonawitz et al., 2014; Giron et al., 2023; Meder et al., 2021; E. Schulz & Gershman, 2019), probability-based decision making (Bernard et al., 2016; Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018; Wente et al., 2019), and question-asking (Herwig, 1982; Ruggeri & Feufel, 2015; Ruggeri & Lombrozo, 2015), crucial for navigating various task structures, continue to develop into their school years. However, a contrasting line of research, highlights the early emerging abilities of young children to adapt their behaviors to task structures (Giron et al., 2023; Gottlieb et al., 2013; Leonard et al., 2020; Leonard et al., 2017; Lucca et al., 2020; Magid et al., 2018; Oudeyer et al., 2007; Ruggeri et al., 2019; Sumner et al., 2019; Swaboda et al., 2022; Wang & Bonawitz, 2022). These studies point out how young children’s exploration does not stem solely from a preference for uncertainty (e.g., in the form of novel or surprising information), because these factors may sometimes lead to overly complex tasks (L. Schulz, 2015). Instead, their exploration takes into account whether situations allow for optimal learning (Gottlieb et al., 2013; Oudeyer et al., 2007; Poli et al., 2020), thus, further underlining the importance of understanding adaptation as a match between the environment and an individual’s specific abilities (Todd & Gigerenzer, 2012). Moreover, research highlights how, when presented with child-appropriate paradigms, even young children can integrate relevant information into their information search and question asking behavior (Lindow, 2021; Ruggeri et al., 2023; Schulze & Hertwig, 2021, 2022; Swaboda et al., 2022) and how they take task relevant features such as their difficulty or potential rewards into account when deciding which tasks to pursue and how long to persist on them (Leonard et al., 2020; Leonard et al., 2017; Lucca et al., 2020; Magid et al., 2018; Wang & Bonawitz, 2022). In the following section, I will delve into this contrasting line of research that underscores young children's abilities to adapt their behaviors to various task structures.

For example, work on children's exploration has suggested that it may be too simplistic to interpret young children's behavior in grid games as merely undirected and random (Giron et al., 2023; Meder et al., 2021). In particular, a recent study that compared the learning choices of participants aged 5 to 55 with those of simulated learning algorithms found that children's exploration evolves with age to become less random and more goal-directed, as well as increasingly efficient across several factors relevant for learning (Giron et al., 2023). According to that work, children's exploration resembles stochastic optimizations of their learning processes. These include adaptations in their reward generalization, uncertainty-directed exploration, and learning-focused decision making. Research targeting younger children has shown that even 4-year-olds do not base all of their exploration on random sampling; instead, they are more likely to explore those options that have the highest uncertainty and aim to make generalized predictions about potential outcomes (Meder et al., 2021). It has also been suggested that young children's tendency to engage in undirected exploration more than adults could indeed put them in an advantageous position because it allows to discover new options, rewards, or changes in the environment that adults often overlook (Lucas et al., 2014; Mata et al., 2013; Sumner et al., 2019). These findings challenge some of the work reporting children's shortcoming's in their adaptive behaviors.

Work on children's information-search and question asking further highlights children's sophisticated adaptation to the available information structures when the experimental designs account for age-appropriate tasks (Lindow, 2021; Ruggeri et al., 2023; Schulze & Hertwig, 2021, 2022; Swaboda et al., 2022) and the potential influence of background knowledge (Ruggeri & Feufel, 2015). For instance, information board games are often complex, which can negatively affect children's performance. In a simplified version of an information-search task developed by Lindow (2021) to which we will return more specifically in Chapter 3, children as young as 5 showed remarkable search efficiency. When having to disambiguate the location of a target hidden in one of four boxes by selecting among four different information cards, each revealing a feature of the target box, the majority of 5-year-olds selected the informative card that would allow them to find the hidden target and stopped their information search after revealing the relevant information. Similarly, research on children's ecological active learning further highlights the importance of age-appropriate research designs and materials to uncover young children's early present adaptation behaviors (Ruggeri et al., 2019; Swaboda et al., 2022). This research also demonstrates that previous studies relying on children's verbal answers (Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018; Herwig, 1982; Ruggeri & Feufel, 2015; Ruggeri et al., 2016) may have underestimated their cognitive capacities and highlights that researchers should utilize designs that facilitate capturing these early competencies if they seek to capture young children's adaptations.

Moreover, children's ability to adapt their learning to different task structures becomes evident when we look at research investigating how they estimate a task's difficulty and allocate their efforts based on such estimates. Research highlights how these abilities are present in early ages and becomes more sophisticated as children mature (Bridgers et al., 2020; Leonard et al., 2020; Leonard et al., 2017; Lucca et al., 2020; Magid et al., 2018; Wang & Bonawitz, 2022). Specifically, studies have shown that young children use cues related to physical features or rewards of a task (Bridgers et al., 2020; Wang & Bonawitz, 2022) and social learning (Leonard et al., 2020; Leonard et al., 2017) to infer their future and others' potential performance in these tasks. For example, one study (Bridgers et al., 2020) found that children infer the task difficulty of a toy from the number of buttons it has when deciding what to teach another person (see also Magid et al., 2018). Similarly, a recent study (Wang & Bonawitz, 2022) shows that from the age of 4, children begin to evaluate tasks based on difficulty and reward levels, preferring simpler or more rewarding tasks. Research on effort allocation indicates that both infants and preschoolers use social cues to decide on task persistence (Leonard et al., 2020; Leonard et al., 2017; Lucca et al., 2020). Infants consider factors such as an adult's performance and persistence in tasks (Leonard et al., 2017; Lucca et al., 2020), while preschoolers assess adults' invested effort, success, and verbal indications of task difficulty (e.g., "this will be hard!") to calibrate their efforts (Leonard et al., 2020). These findings suggest that even from an early age, children not only estimate the difficulty of tasks by evaluating physical and social cues but also use these evaluations to strategically allocate their efforts.

Together, the body of work reviewed here presents an inconclusive picture and highlights the need for further research. On the one hand, infant research broadly agrees that from a young age children adapt their behaviors in reaction to violations of their expectations about object properties (Perez & Feigenson, 2020; L. Schulz, 2015; Stahl & Feigenson, 2015), take into account the probabilities and the likelihood of observed samples (Téglás et al., 2007; Xu & Garcia, 2008), and consider the knowledge of their potential interaction partners when actively gathering information (Begus & Southgate, 2012; Goupil et al., 2016; Kovács et al., 2014). On the other hand, research with preschool- and school-aged children paints a more contradictory picture. Some research suggests that the cognitive abilities relevant for children's adaptation to different task structures may be still developing (Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018; Giron et al., 2023; Herwig, 1982; Meder et al., 2021; Ruggeri & Feufel, 2015; Ruggeri et al., 2016; E. Schulz & Gershman, 2019). In particular, research suggests that young children explore in a random fashion rather than focusing on high-reward outcomes (Bonawitz et al., 2014; Giron et al., 2023; Meder et al., 2021; E. Schulz & Gershman, 2019), that they often fail to base their information-search decisions on probabilistic information (Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018), and that they struggle to ask (Aguiar et al., 2012; Fitneva et al., 2013) and select the most efficient questions (Herwig, 1982; Ruggeri & Feufel, 2015; Ruggeri et al., 2016). However, a growing body of recent work has highlighted that some of the behaviors reported among young children can be quite adaptive in some environments, even to the extent that children outperform adults (Lucas et al., 2014; Sumner et al., 2019). Work on ecological active learning has further demonstrated that even preschool-aged children do adapt their information-search abilities to the different task structures (Ruggeri et al., 2019; Swaboda et al., 2022) provided these tasks are presented to them in a child-friendly way (for an overview, see Ruggeri, 2022). Moreover, research suggests that even young children evaluate the difficulty of tasks based on physical and social information and that they use these information to guide their decisions (Bridgers et al., 2020; Leonard et al., 2020; Leonard et al., 2017; Lucca et al., 2020; Wang & Bonawitz, 2022).

1.2 Adaptation to Goals

Beyond the adaptation of behavior to different task structures, individuals who want to self-direct their learning must adapt their behaviors to different goals. On an abstract level, goals can be conceptualized as the cognitive representations of desired states or outcomes that require the individual to execute specific actions in order to attain them (see Austin & Vancouver, 1996, as cited in Brandstaetter & Bernecker, 2022). As previously illustrated, a child joining elementary school may face immediate goals such as joining a group of friends during recess, or long-term goals such as learning to read. Each goal will require distinct actions and behavioral adjustments from the child. I now turn to research on infants' early present abilities to pursue immediate goals (Rovee-Collier et al., 1978; Rovee-Collier, 1999; Rovee-Collier et al., 2001). I will then present evidence about how young children adapt their behavior to achieve present (Klossek et al., 2008; Klossek et al., 2011; Paulus & Sodian, 2015) and future directed goals (Redshaw & Suddendorf, 2013; Suddendorf & Busby, 2005; Suddendorf & Moore, 2011; Suddendorf et al., 2011), and how they engage in deliberate practice in the here-and-now to achieve goals in the future (Brinums et al., 2018; Davis et al., 2016).

Research with infants underlines that from the first weeks of their lives, they are sensitive to whether their actions trigger outcomes (Kalnins & Bruner, 1973). They also exhibit actions that cause desired outcomes even after a temporal gap between learning how a specific behavior triggers an outcome and repeated exposure to these triggers (Rovee-Collier et al., 1978; Rovee-Collier, 1999; Rovee-Collier et al., 2001). For example, two-month-old infants will perform specific leg motions to activate a mobile hanging above them, even one day after initial exposure to the toy (Rovee-Collier, 1999). However, these studies leave unclear whether infants' actions are explicitly goal-directed, or whether they follow a learned association between stimulus and response (i.e., foot movement and mobile movement). To disentangle these options, researchers use a paradigm adapted from animal behavior studies (Adams & Dickinson, 1981; Dickinson, 1985); such studies allow participants to learn an association between a response and an outcome and then see whether this behavior changes or persists after devaluating the outcome by repeated exposure without the initial behavioral trigger. More specifically, Klossek et al. (2008) presented 16- to 38-month-old children with a screen showing two butterfly icons, each of which produced a different video when clicked. After this training phase, one of the videos was devaluated, that is, children saw the video associated with one of the icons several times without having to click the icon (this non-requested presentation of the video should have made it less interesting compared to the other video—and children should, therefore, be more interested in watching the other video). The experimenters then measured whether participants showed a preference for clicking one icon over the other. Assuming that children's interest in the repeatedly presented video had decreased, they would be expected to exhibit goal-directed behavior toward selectively clicking on the non-devaluated video. Indeed, as predicted, studies using such a paradigm indicate that starting from the age of 18 months, children pursue goals by selectively clicking on the non-devaluated option (Klossek & Dickinson, 2012; Klossek et al., 2011; see also Paulus & Sodian, 2015, for a different approach indicating similar results).

The literature reviewed above suggests that from an early age, children can adapt their behavior to goals in the present. However, the ability to select specific actions in the here-and-now to achieve a future goal presents a more elaborate aspect of goal-directed adaptation. This involves a number of sophisticated cognitive abilities, namely the integration of memory (i.e., remembering a specific task), maintaining this information in the working memory, and planning ahead using episodic foresight (i.e., projecting oneself to a future situation). To investigate this ability in children, research has implemented the *two-room paradigm* (see Suddendorf & Busby, 2005), in which children are led into one room where they encounter a task that they cannot solve (e.g., a toy that needs a specifically shaped key to work, or a puzzle with a missing piece). Next, the children are moved to another room and, after a little break during which they play different games, they are presented with several objects. Among these, one is crucial to solving the task in the initial room. Several studies—varying the target objects and temporal gap between the presentation of the target toy and the potential games—have demonstrated that starting from the age of 4, children can remember a problem in one room and select the right objects that allow them to solve it in another room, even with a temporal gap of 15 minutes between the two tasks during which children play distractor games (Redshaw & Suddendorf, 2013; Suddendorf & Busby, 2005; Suddendorf & Moore, 2011; Suddendorf et al., 2011).

Building upon the two-room paradigm, some studies have investigated children's *deliberate practice* for the future (Brinums et al., 2023; Brinums et al., 2018; Davis et al., 2016; Suddendorf et al., 2016). Deliberate practice can be defined as the repetition of specific actions to improve one's skills to achieve a future goal (Suddendorf et al., 2016). It builds on the ability to remember a goal and to select actions in the present that enable its achievement in the future. Deliberate practice proves indispensable in almost all activities, including simple skills such as throwing, but also more sophisticated abilities such as playing an instrument or writing a text. Although much research has focused on deliberate practice in adult athletes as a means to achieve mastery (Côté et al., 2007; Ericsson, 2008; Ericsson et al., 2009), only a few studies have investigated whether and how young children engage in deliberate practice (Brinums et al., 2023; Brinums et al., 2018; Davis et al., 2016). The existing investigations share a similar approach: children are told that they will eventually face a test and can then select between different actions allowing them to prepare for this future test. For instance, Davis et al. (2016) presented 3- to 5-year-old children with four alternative variations of the same motor-ability task (e.g., four hot-wire games in different shapes) and told children that they would win a sticker if they were able to complete one specific of the tasks later. The children then played a distractor game and entered the second room, where they were presented with four identical games as in room one. Before returning to the first room, children could practice with one of the games for one minute and were asked to explain why they chose this specific game. The results of this study indicated that starting from the age of 4, children engage in deliberate practice (Davis et al., 2016). However, it takes until the age of 5 to gain an explicit understanding that it takes practice to improve in a specific skill. Using a similar approach, Brinums et al. (2018) found that 6- and 7-year-olds are able to explicitly state why practice benefits their skill development—specifically practicing the target game they will later be tested on—and are more likely to engage with the target game first, as compared to 4- and 5-year-olds, who did not mention the importance of practice and did not practice the target game more often than the distractor games. Interestingly, a recent study has emphasized how 8-year-old children's ability to engage in successful deliberate practice correlated with their ability to envision themselves in the future (Brinums et al., 2023). This work highlighted how children's deliberate practice ability improved when they were prompted to imagine how they will feel in the future after having successfully completed a task.

Taken together, the evidence reviewed so far opens up a versatile picture of children's goal adaptation. Even young children can adapt their own behavior to achieve immediate goals (Klossek & Dickinson, 2012; Klossek et al., 2011; Paulus & Sodian, 2015) and from the age of 4 children can adapt their behavior to solve spatially and temporally separated tasks (Redshaw & Suddendorf, 2013; Suddendorf & Moore, 2011). However, the available evidence about deliberate practice is inconclusive (Brinums et al., 2023; Brinums et al., 2018; Davis et al., 2016). While Davis et al. (2016) reports that already at the age of 4 children engage in deliberate practice, children in the study by Brinums et al. (2018) engaged in deliberate practice only from age 6. This available evidence on children's underscores the need for further empirical studies to clarify when and under which circumstances young children can adapt their practice to pursue future goals.

1.3 Adaptation to One's Abilities

Early on in this work, I introduced a definition of *adaptation* as a match between an individual's cognitive abilities and the specific requirements that are imposed onto them by their environment (Todd & Gigerenzer, 2012). Throughout the introduction of this thesis, I have reviewed literature on how infants and young children adapt their behaviors when facing various task structures and goals. After having looked at one part of the equation (that is, adaptation to the environment), we now turn our attention to the other part of the equation: children's abilities to adapt their behavior to their *abilities*. Each person adapts to task structures and goals from their unique and individual standpoint in the world. Given that all people differ, we must take into consideration that successful adaption necessitates adaptation to the specifics of one self—or, more precisely, an individual's specific abilities, characteristics, and skills. A tall person who wants to learn a new sport will have to adapt their movements differently compared to a shorter person. Likewise, compared to a novice, someone with experience in a topic will make different learning decisions, thus adapting to their individual *abilities*. Consequently, there can be no one-size-fits-all approach to learning from the standpoint of the individual learner; on the contrary, each person must individually adapt their learning to *themselves*. Adapting to one's abilities provides significant advantages by preventing individuals from spending too much time on tasks that are too difficult or even impossible for them to achieve. Moreover, it allows them to make informed learning decisions based on an assessment of what they already know, such as whether they can still improve in the task. How much they can navigate their learning under no one's leadership but their own will depend greatly on the degree to which they can reliably assess what they already know, and what they can or cannot do. Therefore, in the following paragraphs we will be delving into literature that investigates the question, *How do children reason about their abilities, and can they then base their learning decisions on this reasoning?*

The ability to reason about ones own knowledge, abilities, and thinking is often referred to as *metacognition* (Flavell, 1978, 1979). Early work on metacognition (Flavell, 1978, 1979; Piaget & Inhelder, 1964) suggested that it is not before the age of 5, that children show first adaptations to their own abilities. For instance, when 5-year-olds are tasked with memorizing a list of objects, they begin to employ a rehearsal strategy that aids in their memorization if this strategy is explicitly introduced to them. However, without this introduction, they fail to adopt the strategy in 90% of the cases (for an overview, see Whitebread & Neale, 2020). The early work on metacognition, has motivated a large body of literature with more recent studies expanding the understanding of young children's abilities to engage in metacognitive processes (e.g., Destan et al., 2014; Gascoine et al., 2017; Geurten & Bastin, 2019; Metcalfe & Finn, 2013). Studies have shown that infants and young children have some early awareness of their knowledge, or lack thereof, and adapt their behaviors accordingly (Goupil et al., 2016; Lyons & Ghetti, 2010; Moore et al., 1989; Moore et al., 1994; Shatz et al., 1983). For instance, 20-month-old children selectively seek assistance from their caregivers when they are uncertain about where a toy is hidden (Goupil et al., 2016). Research further suggests that by the end of age of 2, children start using words indicating their knowledge (e.g., *I know, I think, I do not know;* Moore et al., 1994). Although this may be a simple reflection of pragmatic use of these words in their environment (Lyons & Ghetti, 2010), research has indicated that by the age of 2.5, children indeed intentionally use these terms to indicate what they know and do not know (Moore et al., 1989; Moore et al., 1994; Shatz et al., 1983). To study this, Marazita and Merriman (2004) presented children with real and fake words, or real and novel objects, and asked whether they knew or did not know those words or objects. They found that 2.5-year-old children accurately indicated their knowledge and ignorance of words and objects. Similarly, in various studies, children were questioned about their knowledge of the contents of closed box (Pratt & Bryant, 1990; Rohwer et al., 2012; Wimmer et al., 1988). Starting from 2 years, children who previously peeked inside the box indicated knowing its content, whereas children who had not peeked inside the box were significantly less likely to indicate knowledge of the box's contents (Pratt & Bryant, 1990; Rohwer et al., 2012; Wimmer et al., 1988). Between 3 and 5 years of age, these introspective abilities improve. Children recognize when they are uncertain, and are more likely to seek help on those trials of perceptual identification tasks for which they indicated more uncertainty (Coughlin et al., 2015).

Although awareness of one's knowledge and the ability to verbalize it can serve as initial steps in adapting to one's abilities, it is less clear whether children adjust their learning and information-search behaviors based on evaluations of their knowledge and areas of ignorance. Furthermore, do children base these adaptations on a fine-grained assessment of their strengths and weaknesses, moving beyond merely identifying what they know to estimating *how well* they know something? The evidence regarding these more nuanced questions is mixed. On the one hand, research has indicated that children spontaneously begin tracking their previous performance by age 5 (Niebaum & Munakata, 2020), but also that they are often overly confident about their performance (Finn & Metcalfe, 2014; Metcalfe & Finn, 2013). Even if they can metacognitively assess their knowledge, they often fail to make efficient learning decisions based on these assessments until they reach fifth grade (Metcalfe & Finn, 2013). Research examining preschool- and school-aged children's metacognition has shown strong developmental changes in the effectiveness of their training strategies (Finn & Metcalfe, 2014). More specifically, third-grade children do not dedicate additional time to repeat items they failed to remember on an interim memory check before a memory test, despite reporting lower confidence in their memory of these items (Bisanz et al., 1978) and although preschoolers can indicate which items they know and do not know, they fail to focus on these items when given the opportunity to practice before a memory test (Metcalfe & Finn, 2013). Similarly, 4- and 5-year-olds do not study items they previously failed to answer (Flavell et al., 1970). Whereas children between 5 and 7 years assess the mathematical items that are respectively easy and difficult for them to solve and then choose to be tested on the items that are easier for them, 3- to 4-year-olds do not pick items based on their previous performance (Baer & Odic, 2019).

Successfully integrating metacognitive estimates into one's behavior requires the integration of processes from executive functioning, working memory, and inhibition control (Lyons & Zelazo, 2011; Roebers, 2017; Spiess et al., 2016). Such processes enable individuals to integrate information into current decision-making, impacting what and how they choose to learn (Mata et al., 2015).

Particularly for young learners whose cognitive competencies are still developing such integration might be challenging (Blackwell & Munakata, 2014; Chatham et al., 2009; Chevalier et al., 2015; Kail, 2007; Legare et al., 2013; Mata et al., 2015; Roebers, 2017). Studies that control for these potential confounds and provide scaffolds for children (e.g., in the form of memory aids of their prior performance) have found that even preschoolers can adapt their behavior based on their previous performance (Hembacher & Ghetti, 2014; Leonard et al., 2023; Siegel et al., 2021). For instance, 4- to 5-year-old children express greater confidence in items they previously guessed correctly compared to those they got wrong (Hembacher & Ghetti, 2014) and they adjust their exploration efforts in response to the outcomes of *their own* previous actions (Siegel et al., 2021), while 4- to 6-year-olds use their prior performance to decide how to allocate their efforts (Leonard et al., 2023). In the study by Leonard et al. (2023), children are more likely to persist on a task in which they had previously been improving, as opposed to when their performance reached a plateau, even when performance rewards were consistent across both conditions. This suggests that children of these ages can adapt their efforts to their own actions and abilities.

Together, the reviewed body of work underlines the need for more empirical research on how children adapt their behaviors to their individual set of abilities, strengths, and weaknesses. Although there appears to be evidence that even young children possess some awareness of their knowledge and ignorance (Goupil et al., 2016; Lyons & Ghetti, 2011), and that they can adapt their verbal utterances accordingly (Moore et al., 1989; Moore et al., 1994; Shatz et al., 1983), the findings for older children are mixed. A substantial body of research highlights children's tendency to overestimate their abilities and knowledge (Destan et al., 2014; Finn & Metcalfe, 2014; Metcalfe & Finn, 2013; Salles et al., 2016) noting that even where children manage to reliably assess their skill level, they fail to put these estimates into practice (Bisanz et al., 1978; Metcalfe & Finn, 2013). However, a growing body of research paints a more optimistic picture and emphasizes that even young children can reliably estimate their abilities (that is, their state of knowledge and their prior performance in a task; Lyons & Ghetti, 2011) and that these estimates can guide their actions (Leonard et al., 2023; Siegel et al., 2021). The work presented raises questions about how effectively children can direct their learning based on an assessment of their abilities, strengths, and weaknesses.

1.4 Outline of the Thesis

In this thesis, I investigate children's adaptive learning across three dimensions central for self-guided learning. Specifically, the crucial dimensions I explore are how children adapt their learning to the *task structures*, the *goals* they face, and their *abilities*. I empirically approach children's adaptive learning in across these dimensions through three experimental studies, each focusing either on one specific dimension or integrating multiple dimensions in which adaptation can manifest. Two of these studies (see Chapter 2 and 3) have been published in the Proceedings of the Annual Conference of the Cognitive Science Society, one study (see Chapter 4) is currently under review in Child Development.

Chapter 2 investigates how 5- to 7-year-old children adapt their explorative efforts to *task structures*. More specifically, how they adapt their exploration to causal relationships (i.e., monsters that produce energy by shaking) that vary in their causal strength across backgrounds (i.e., different planets that the monsters can travel to). To study this, I create a virtual environment where children made a number of forced-choice decisions and could engage in free exploration by sending specific monsters (i.e., causes) to different planets (i.e., backgrounds), thereby learning about their underlying causal strength and stability. Results indicate that children's forced-choice decisions are guided by a sensitivity for the stability of the target monsters and that with increasing age, children are more likely to specifically explore and re-explore some monster and planet combination than others.

Chapter 3 then adds the dimension of *goals* to the investigation. Here I focus on 2- to 4-year-old's ability to adapt their information-search decisions to the information structure of a game in which they have to disambiguate the location of a hidden present by choosing between relevant and irrelevant information cards. By making a paradigm initially presented by Lindow (2021) more accessible for young children by reducing task demands, I find that at the age of 2 years, children can already select the information card that allows them to infer the location of the present. This effect holds even when children are presented with a new set of materials, demonstrating their adaptation across contexts.

Chapter 4 investigates 4- to 8-year-old children's and adults' abilities to integrate all three dimensions of adaptation (*task structures*, *goals*, and their *abilities*). More specifically, I study the capacity to prepare for future test scenarios through deliberate practice based on an assessment of one's prior performance. To investigate this, I present participants with two games: one in which it was fairly easy to succeed, and another more difficult game. After familiarizing themselves with these games, the participants are informed that they will later face a test in either the easy, the difficult, or a randomly chosen game. I discover that already at the age of 4, children consider the future goals and adjust their practice choices based on varying task characteristics, such as game difficulty or the likelihood of being tested in one game compared to another and their own performance in these games. Chapter 5 provides a comprehensive discussion of the key empirical contributions, outlining potential directions for future research and implications for research and educational practice.

1.5 Methodology

Samples

The studies presented in this dissertation focused on children within the age range of 2 to 8 years. Participants for the experiments described in Chapter 2 and the first experiment in Chapter 4 were recruited and tested online due to restrictions from the COVID-19 pandemic. Participants for the study in Chapter 3 and the second experiment in Chapter 4 were recruited and tested either at the Berlin Zoo or in the lab of the Max Planck Institute for Human Development in Berlin. All participants were fluent in German. Before participation, written informed consent was obtained from the legal guardians of all children. Children provided verbal consent and were rewarded with small gifts, such as stickers, for participating in the studies. The Ethics Committee of the Max Planck Institute for Human Development in Berlin approved all studies.

General Methodological Approach

In this thesis, adaptation—understood as a match between the requirements of the environment and an individual's specific abilities (Todd & Gigerenzer, 2012)—has been described as a central concept. Beyond interpreting children's behavior in experimental studies using this understanding of adaptation, I propose that when choosing the methods to conduct scientific research, scientists should ask whether the chosen methods are *adaptive* for the research question they aim to answer. Following this logic, there is no single correct approach to conducting research with children. On the contrary, different research questions and different age groups demand *adaptive* choices of methods. With this in mind, and given that I targeted a broad age range—starting from children as young as 2 years to children aged 8 years and adults—I decided to use a *behavioral* approach to target the research I will present henceforth. This approach allows me to study the decisions children make in specific situations, even if in some cases they may still be too young to verbally indicate their decisions or explain why they decided to show these specific behaviors (Köymen & Tomasello, 2020). To approach the different research questions I ask in the respective studies, I developed experimental games that allowed me to investigate children's adaptations across the three different dimensions that I deem essential for successful self-guided learning: task structures, goals, and their abilities. All materials were designed to be fun and engaging for children while still maintaining strict requisites of scientific standards. All data analyses were carried out using the R statistical software. All studies, except for the study presented in Chapter 3, were pre-registered on OSF. Any deviations from the pre-registrations are explicitly reported in the respective studies.

1.6 References

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*, *33*(2b), 109–121. https://doi.org/10/gk6c49

Aguiar, N. R., Stoess, C. J., & Taylor, M. (2012). The development of children's ability to fill the gaps in their knowledge by consulting experts. *Child Development*, *83*(4), 1368–1381. https://doi.org/10/f379gz

Baer, C., & Odic, D. (2019). Certainty in numerical judgments develops independently of the approximate number system. *Cognitive Development*, *52*. https://doi.org/10/gf88gk

Begus, K., & Southgate, V. (2012). Infant pointing serves an interrogative function. *Developmental Science*, *15*(5), 611–617. https://doi.org/10/f38vgj

Bernard, S., Clement, F., & Mercier, H. (2016). Wishful thinking in preschoolers. *Journal of Experimental Child Psychology*, *141*, 267–274. https://doi.org/10/ghz9t9

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology*, *115*(1), 108–125. https://doi.org/10/f4vxqn

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica*, *153*, 74–86. https://doi.org/10/f6vd8w

Betsch, T., Lehmann, A., Jekel, M., Lindow, S., & Glöckner, A. (2018). Children's application of decision strategies in a compensatory environment. *Judgment and Decision Making*, *13*(6), 514–528. https://doi.org/10/grqb3t

Bisanz, G. L., Vesonder, G. T., & Voss, J. F. (1978). Knowledge of one's own responding and the relation of such knowledge to learning. *Journal of Experimental Child Psychology*, *25*(1), 116–128. https://doi.org/10/dvvrfv

Blackwell, K. A., & Munakata, Y. (2014). Costs and benefits linked to developments in cognitive control. *Developmental Science*, *17*(2), 203–211. https://doi.org/10/f5t7rn

Bonawitz, E., Denison, S., Griffiths, T. L., & Gopnik, A. (2014). Probabilistic models, learning algorithms, and response variability: Sampling in cognitive development. *Trends in Cognitive Sciences*, *18*(10), 497–500. https://doi.org/10/f6kqqs

Brandstaetter, V., & Bernecker, K. (2022). Persistence and disengagement in personal goal pursuit. *Annual Review of Psychology*, *73*, 271–299. https://doi.org/10/gm39k5

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. https://doi.org/10/gq92d5

Brinums, M., Franco, C., Kang, J., Suddendorf, T., & Imuta, K. (2023). Driven by emotion: Anticipated feelings motivate children's deliberate practice. *Cognitive Development*, *66*, 101340. https://doi.org/10/gtphnj

Brinums, M., Imuta, K., & Suddendorf, T. (2018). Practicing for the future: Deliberate practice in early childhood. *Child Development*, *89*(6), 2051–2058. https://doi.org/10/gfj2wh

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences*, *106*(14), 5529–5533. https://doi.org/10/b5x9gz

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience*, *27*(6), 1125–1136. https://doi.org/10/gc5ss6

Ciranka, S., & van den Bos, W. (2021). Adolescent risk-taking in the context of exploration and social influence. *Developmental Review*, *61*, 100979. https://doi.org/10/gnrth3

Côté, J., Baker, J., & Abernethy, B. (2007). Practice and play in the development of sport expertise. In *Handbook of sport psychology* (pp. 184–202). John Wiley & Sons, Ltd.

Coughlin, C., Hembacher, E., Lyons, K. E., & Ghetti, S. (2015). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science*, *18*(6), 957–971. https://doi.org/10/f7vdbp

Davis, J. T. M., Cullen, E., & Suddendorf, T. (2016). Understanding deliberate practice in preschool-aged children. *Quarterly Journal of Experimental Psychology*, *69*(2), 361–380. https://doi.org/10/gh7x9t

De Simone, C., & Ruggeri, A. (2022). Searching for information, from infancy to adolescence. In I. C. Dezza, E. Schulz, & C. M. Wu (Eds.), *The drive for knowledge: The science of human information-seeking.* (pp. 77–100). Cambridge University Press.

Destan, N., Hembacher, E., Ghetti, S., & Roebers, C. M. (2014). Early metacognitive abilities: The interplay of monitoring and control processes in 5- to 7-year-old children. *Journal of Experimental Child Psychology*, *126*, 213–228. https://doi.org/10/gjvrdn

Dickinson, A. (1985). Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *308*(1135), 67–78. https://doi.org/10/bxmmqh

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine*, *15*(11), 988–994. https://doi.org/10/d2v2n7

Ericsson, K. A., Nandagopal, K., & Roring, R. W. (2009). Toward a science of exceptional achievement. *Annals of the New York Academy of Sciences*, *1172*(1), 199–217. https://doi.org/10/cxdpk8

Finn, B., & Metcalfe, J. (2014). Overconfidence in children's multi-trial judgments of learning. *Learning and Instruction*, *32*, 1–9. https://doi.org/10/gtphnq

Fitneva, S. A., Lam, N. H. L., & Dunfield, K. A. (2013). The development of children's information gathering: To look or to ask? *Developmental Psychology*, *49*(3), 533–542. https://doi.org/10/f4p6mt

Flavell, J. H. (1978). Metacognitive development. In J. M. Scandura & C. J. Brainerd (Eds.), *Structural/process models of complex human behavior* (pp. 213–246). Springer Science & Business Media.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906–911. https://doi.org/10/cg33qs

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology*, *1*(4), 324–340. https://doi.org/10/dt7tpf

Gascoine, L., Higgins, S., & Wall, K. (2017). The assessment of metacognition in children aged 4–16 years: A systematic review. *Review of Education*, *5*(1), 3–57. https://doi.org/10/ghhfc8

Geurten, M., & Bastin, C. (2019). Behaviors speak louder than explicit reports: Implicit metacognition in 2.5-year-old children. *Developmental Science*, *22*(2). https://doi.org/10/gd9tzr

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality: The adaptive toolbox*. The MIT Press https://www.doi.org/10.7551/mitpress/1654.001.0001.

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*, 1–13. https://doi.org/10/gs93v3

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: Minds, brains, and how children learn*. William Morrow & Co.

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., Aboody, R., Fung, H., & Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences*, *114*(30), 7892–7899. https://doi.org/10/gbq863

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593. https://doi.org/10/f5hpnt

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. https://doi.org/10/f8gq87

Harris, P. L., Koenig, M. A., Corriveau, K. H., & Jaswal, V. K. (2018). Cognitive foundations of learning from testimony. *Annual Review of Psychology*, *69*, 251–273. https://doi.org/10/gdrbzv

Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science*, *25*(9), 1768–1776. https://doi.org/10/f6ks7j

Herwig, J. E. (1982). Effects of age, stimuli, and category recognition factors in children's inquiry behavior. *Journal of Experimental Child Psychology*, *33*(2), 196–206. https://doi.org/10/dfd86w

Kail, R. V. (2007). Longitudinal evidence that increases in processing speed and working memory enhance children's reasoning. *Psychological Science*, *18*(4), 312–313. https://doi.org/10/bfq748

Kalnins, I. V., & Bruner, J. S. (1973). The coordination of visual observation and instrumental behavior in early infancy. *Perception*, *2*(3), 307–314. https://doi.org/10/bkmxtz

Klossek, U., & Dickinson, A. (2012). Rational action selection in 11/2-to 3-year-olds following an extended training experience. *Journal of Experimental Child Psychology*, *111*(2), 197–211. https://doi.org/10/dr5rpv

Klossek, U., Russell, J., & Dickinson, A. (2008). The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *Journal of Experimental Psychology: General*, *137*(1), 39. https://doi.org/10/dnbf3b

Klossek, U., Yu, S., & Dickinson, A. (2011). Choice and goal-directed behavior in preschool children. *Learning & Behavior*, *39*(4), 350–357. https://doi.org/10/ddrwmw

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. https://doi.org/10/gtphn2

Köymen, B., & Tomasello, M. (2020). The early ontogeny of reason giving. *Child Development Perspectives*, *14*(4), 215–220. https://doi.org/10/gtphn3

Legare, C. H., Mills, C. M., Souza, A. L., Plummer, L. E., & Yasskin, R. (2013). The use of questions as problem-solving strategies during early childhood. *Journal of Experimental Child Psychology*, *114*(1), 63–76. https://doi.org/10/f4jghf

Leonard, J. A., Cordrey, S. R., Liu, H. Z., & Mackey, A. P. (2023). Young children calibrate effort based on the trajectory of their performance. *Developmental Psychology*, *59*(3), 609–619. https://doi.org/10/gtphn7

Leonard, J. A., Garcia, A., & Schulz, L. E. (2020). How adults' actions, outcomes, and testimony affect preschoolers' persistence. *Child Development*, *91*(4), 1254–1271. https://doi.org/10.1111/cdev.13305

Leonard, J. A., Lee, Y., & Schulz, L. E. (2017). Infants make more attempts to achieve a goal when they see adults persist. *Science*, *357*(6357), 1290–1294. https://doi.org/10/gb2cwh

Lindow, S. (2021). On searching and finding: The development of information search abilities. *Cognitive Development*, *58*, 101011. https://doi.org/10/grcfkx

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, *131*(2), 284–299. https://doi.org/10/f5x276

Lucca, K., Horton, R., & Sommerville, J. A. (2020). Infants rationally decide when and how to deploy effort. *Nature Human Behaviour*, *4*(4), 372–379. https://doi.org/10/gq6jbr

Lyons, K. E., & Ghetti, S. (2010). Metacognitive development in early childhood: New questions about old assumptions. In A. Efklides & P. Misailidi (Eds.), *Trends and prospects in metacognition research* (pp. 259–278). Springer US.

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*(6), 1778–1787. https://doi.org/10/bkn4tg

Lyons, K. E., & Zelazo, P. D. (2011). Chapter 10 - monitoring, metacognition, and executive function: Elucidating the role of self-reflection in the development of self-regulation. In J. B. Benson (Ed.), *Advances in child development and behavior* (pp. 379–412). JAI.

Magid, R. W., DePascale, M., & Schulz, L. E. (2018). Four-and 5-year-olds infer differences in relative ability and appropriately allocate roles to achieve cooperative, competitive, and prosocial goals. *Open Mind*, *2*(2), 72–85. https://doi.org/10/grqb4h

Marazita, J. M., & Merriman, W. E. (2004). Young children's judgment of whether they know names for objects: The metalinguistic ability it reflects and the processes it involves. *Journal of Memory and Language*, *51*(3), 458–472. https://doi.org/10/d282zn

Mata, R., Josef, A., & Lemaire, P. (2015). Chapter 6 - adaptive decision making and aging. In *Aging and decision making: Empirical and applied perspectives* (pp. 105–126).

Mata, R., Wilke, A., & Czienskowski, U. (2013). Foraging across the life span: Is there a reduction in exploration with aging? *Frontiers in Neuroscience*, *7*, 53. https://doi.org/10/gfkpgh

Meder, B., Wu, C. M., Schulz, E., & Ruggeri, A. (2021). Development of directed and random exploration in children. *Developmental Science*, *24*(4). https://doi.org/10/gq6jcr

Metcalfe, J., & Finn, B. (2013). Metacognition and control of study choice in children. *Metacognition and Learning*, *8*(1), 19–46. https://doi.org/10/f4qkjg

Mills, C. M., & Landrum, A. R. (2016). Learning who knows what: Children adjust their inquiry to gather information from others. *Frontiers in Psychology*, *7*, 951. https://doi.org/10/f8sw97

Mills, C. M., Legare, C. H., Bills, M., & Mejias, C. (2010). Preschoolers use questions as a tool to acquire knowledge from different sources. *Journal of Cognition and Development*, *11*(4), 533–560. https://doi.org/10/cmfzmx

Mills, C. M., Legare, C. H., Grant, M. G., & Landrum, A. R. (2011). Determining who to question, what to ask, and how much information to ask for: The development of inquiry in young children. *Journal of Experimental Child Psychology*, *110*(4), 539–560. https://doi.org/10/dhmcff

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development*, *60*(1), 167–171. https://doi.org/10/bmgcz3

Moore, C., Furrow, D., Chiasson, L., & Patriquin, M. (1994). Developmental relationships between production and comprehension of mental terms. *First Language*, *14*(42-43), 001–17. https://doi.org/10/gtphpd

Niebaum, J., & Munakata, Y. (2020). Deciding what to do: Developments in children's spontaneous monitoring of cognitive demands. *Child Development Perspectives*, *14*(4), 202–207. https://doi.org/10/ghmsp8

Oudeyer, P.-Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, *11*(2), 265–286. https://doi.org/10/fck77s

Paulus, M., & Sodian, B. (2015). Which way to take? Infants select an efficient path to their goal. *Journal of Experimental Child Psychology*, *137*, 111–124. https://doi.org/10/gtphph

Perez, J., & Feigenson, L. (2020). Violations of expectation trigger infants to search for explanations. https://doi.org/https://www.doi.org/10.31234/osf.io/eahjd

Piaget, J. (1930). *The child's conception of causality* (Vol. 5). Routledge.

Piaget, J. (1952). *The origins of intelligence in children.* International Universities Press https://www.doi.org/10.1037/11494-000.

Piaget, J., & Inhelder, B. (1964). *The early growth of logic in the child: Classification and seriation.* New York.

Poli, F., Serino, G., Mars, R. B., & Hunnius, S. (2020). Infants tailor their attention to maximize learning. *Science Advances*, *6*(39), eabb5053. https://doi.org/10/ghdx9g

Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development*, *61*(4), 973–982. https://doi.org/10/bt9n9h

Redshaw, J., & Suddendorf, T. (2013). Foresight beyond the very next event: Four-year-olds can link past and deferred future episodes. *Frontiers in Psychology*, *4*, 404. https://doi.org/10/gbfpwz

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review*, *45*, 31–51. https://doi.org/10/gbqdb7

Rohwer, M., Kloo, D., & Perner, J. (2012). Escape from metaignorance: How children develop an understanding of their own lack of knowledge. *Child Development*, *83*(6), 1869–1883. https://doi.org/10/f3ssk9

Ronfard, S., Zambrana, I. M., Hermansen, T. K., & Kelemen, D. (2018). Question-asking in childhood: A review of the literature and a framework for understanding its development. *Developmental Review*, *49*, 101–120. https://doi.org/10/gd9fvz

Rovee-Collier, C., Morrongiello, B. A., Aron, M., & Kupersmidt, J. (1978). Topographical response differentiation and reversal in 3-month-old infants. *Infant Behavior and Development*, *1*, 323–333. https://doi.org/10/bt4pbg

Rovee-Collier, C. (1999). The development of infant memory. *Current Directions in Psychological Science*, *8*(3), 80–85. https://doi.org/10/dvprrb

Rovee-Collier, C., Hayne, H., & Colombo, M. (2001). *The development of implicit and explicit memory.* John Benjamins Publishing Company.

Ruggeri, A. (2022). An introduction to ecological active learning. *Current Directions in Psychological Science*. https://doi.org/10/grbccz

Ruggeri, A., & Feufel, M. (2015). How basic-level objects facilitate question-asking in a categorization task. *Frontiers in Psychology*, *6*, 918. https://doi.org/10/f7jmf7

Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, *143*, 203–216. https://doi.org/10/f7n6j2

Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2016). Sources of developmental change in the efficiency of information search. *Developmental Psychology*, *52*(12), 2159. https://doi.org/10/f9gbx8

Ruggeri, A., Sim, Z. L., & Xu, F. (2017). "Why is Toma late to school again?" Preschoolers identify the most informative questions. *Developmental Psychology*, *53*(9), 1620. https://doi.org/10/gcjpjg

Ruggeri, A., Stanciu, O., Pelz, M., Gopnik, A., & Schulz, E. (2023). Preschoolers search longer when there is more information to be gained. *Developmental Science*. https://doi.org/10/gsbgvt

Ruggeri, A., Swaboda, N., Sim, Z. L., & Gopnik, A. (2019). Shake it baby, but only when needed: Preschoolers adapt their exploratory strategies to the information structure of the task. *Cognition*, *193*, 104013. https://doi.org/10/grqb4j

Salles, A., Ais, J., Semelman, M., Sigman, M., & Calero, C. I. (2016). The metacognitive abilities of children and adults. *Cognitive Development*, *40*, 101–110. https://doi.org/10/f9czkm

Schulz, E., & Gershman, S. J. (2019). The algorithmic architecture of exploration in the human brain. *Current Opinion in Neurobiology*, *55*, 7–14. https://doi.org/10/gft7gh

Schulz, L. (2015). Infants explore the unexpected. *Science*, *348*(6230), 42–43. https://doi.org/10/gtphpq

Schulze, C., & Hertwig, R. (2021). A description–experience gap in statistical intuitions: Of smart babies, risk-savvy chimps, intuitive statisticians, and stupid grown-ups. *Cognition*, *210*, 104580. https://doi.org/10/gh6zr6

Schulze, C., & Hertwig, R. (2022). Experiencing statistical information improves children's and adults' inferences. *Psychonomic Bulletin & Review*, *29*(6), 2302–2313. https://doi.org/10/gtchdq

Shatz, M., Wellman, H. M., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental state. *Cognition*, *14*(3), 301–321. https://doi.org/10/dv47rt

Siegel, M. H., Magid, R. W., Pelz, M., Tenenbaum, J. B., & Schulz, L. E. (2021). Children's exploratory play tracks the discriminability of hypotheses. *Nature Communications*, *12*(1), 3598. https://doi.org/10/grqb37

Sim, Z. L., & Xu, F. (2017). Infants preferentially approach and explore the unexpected. *British Journal of Developmental Psychology*, *35*(4), 596–608. https://doi.org/10/gq6jb6

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*(4), 605–632. https://doi.org/10/frctvd

Spiess, M. A., Meier, B., & Roebers, C. M. (2016). Development and longitudinal relationships between children's executive functions, prospective memory, and metacognition. *Cognitive Development*, *38*, 99–113. https://doi.org/10/gjpxwd

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*(6230), 91–94. https://doi.org/10/f66td4

Suddendorf, T., Brinums, M., & Imuta, K. (2016). Shaping one's future self: The development of deliberate practice. In *Seeing the future: Theoretical perspectives on future-oriented mental time travel* (pp. 343–366). Oxford University Press.

Suddendorf, T., & Busby, J. (2005). Making decisions with the future in mind: Developmental and comparative identification of mental time travel. *Learning and Motivation*, *36*(2), 110–125. https://doi.org/10/bkj78h

Suddendorf, T., & Moore, C. (2011). Introduction to the special issue: The development of episodic foresight. *Cognitive Development*, *26*(4), 295–298. https://doi.org/10/c94fqn

Suddendorf, T., Nielsen, M., & von Gehlen, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science*, *14*(1), 26–33. https://doi.org/10/chqpbw

Sumner, E., Li, A. X., Perfors, A., Hayes, B., Navarro, D., & Sarnecka, B. W. (2019). The exploration advantage: Children's instinct to explore allows them to find information that adults miss. https://doi.org/10/grqb4n

Swaboda, N., Meder, B., & Ruggeri, A. (2022). Finding the (most efficient) way out of a maze is easier than asking (good) questions. *Developmental Psychology*, *58*(9), 1730–1746. https://doi.org/10/grn9jp

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. https://doi.org/10/fpk2mt

Todd, P. M., & Gigerenzer, G. (2012). *Ecological rationality: Intelligence in the world*. Oxford University Press, USA.

Wang, J., & Bonawitz, E. (2022). Children's sensitivity to difficulty and reward probability when deciding to take on a task. *Journal of Cognition and Development*. https://doi.org/10/grmj56

Wente, A., Goddu, M., Garcia, T., Posner, E., Flecha, M., & Gopnik, A. (2019). Young children are wishful thinkers: The development of wishful thinking in 3- to 10-year-old children. *Child Development*, *91*. https://doi.org/10/gf6294

Whitebread, D., & Neale, D. (2020). Metacognition in early child development. *Translational Issues in Psychological Science*, *6*(1), 8–14. https://doi.org/10/gtphpx

Wimmer, H., Hogrefe, G.-J., & Perner, J. (1988). Children's understanding of informational access as source of knowledge. *Child Development*, *59*(2), 386–396. https://doi.org/10/c649tg

Wu, C. M., Schulz, E., Garvert, M. M., Meder, B., & Schuck, N. W. (2020). Similarities and differences in spatial and non-spatial cognitive maps. *PLoS computational biology*, *16*(9). https://doi.org/10/gmchzt

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour*, *2*(12), 915–924. https://doi.org/10/gfjwtq

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences*, *105*(13), 5012–5015. https://doi.org/10/d2mmgv

Chapter 2

The Role of Causal Stability in Children's Active Exploration

Serko, D., Vasil, N., & Ruggeri, A. (2023).

Proceedings of the 45th Annual Conference of the Cognitive Science Society.

This paper can be accessed at:

https://escholarship.org/uc/item/8h38n375

Abstract

Previous research documented adults' preference for stable causal relationships that do not vary in strength across backgrounds (Vasilyeva et al., 2018). In this study, we investigate the role of causal stability in guiding children's exploration behavior. We developed a computerized version of an active information-search paradigm to study how children dynamically explore different agents and backgrounds to learn more about their causal stability. Five- to seven-year-old children (n = 60) were presented with stable and unstable causes (i.e., causes with fixed or variable causal efficacy across backgrounds). We assessed children's causal attributions of outcomes and their exploratory behavior as they tried out previously observed and novel causes across previously observed and novel backgrounds. We find that children in this age range acknowledge causal instability in their causal attributions, and they become increasingly adept at tracking causal efficacy across multiple factors simultaneously (causes and backgrounds), but this does not translate into a blanket preference for exploring stable or unstable causes. We suggest a possibility that causal (in)stability guides exploration in more subtle and indirect ways and discuss the implications of our findings for the development of active exploration.

Introduction

Imagine you and your neighbor both decide to grow strawberries. Each of you plants the seeds in your own backyard and waits for the outcome. To your dismay, nothing happens in your garden — while your neighbor is harvesting bushels of juicy berries. You are puzzled: both you and the neighbor did the same thing, i.e., intervened on the same cause. Why the different outcomes? It turns out the chances that a cause will in fact produce the desired outcome—sweet and juicy strawberries—depend on several background variables, which might significantly increase or decrease the likelihood that the cause will produce an effect. For example, planting strawberry seeds is more likely to result in strawberries if the soil is acidic. If your backyard has alkaline soil, the same cause may not be as effective—so you end up with no berries.

This example illustrates the notion of causal *instability* across *backgrounds*. Philosophers have defined *stable* causal relationships as those that hold with similar strength across different backgrounds (where *backgrounds* can refer to any variable other than the cause and effect, Woodward, 2006, 2010). Causal stability can be defined in terms of variability in causal *strength*. Various measures of causal strength exist, with important differences among them (Cartwright, 1989, 2009; Cheng, 2000; Cheng et al., 2013; Cheng & Lu, 2017; Liljeholm & Cheng, 2007), but it can be glossed as a generalized measure of efficacy of a cause in generating an outcome, controlling for other factors. Suppose you place 8 seeds in your garden's soil and 8 in store-bought pre-fertilized soil. The number of strawberries you eventually obtain will reveal the causal strength in these specific backgrounds: more strawberries indicate higher causal strength. Assessing overall causal strength is important for selecting the most effective interventions from a range of possibilities (e.g., the best-producing variety of strawberries, or the best soil) (Meder et al., 2014). But beyond that, assessing how much causal strength *varies* across backgrounds, its stability, can offer better guidance for predictions within and across backgrounds (Blanchard et al., 2018; Liljeholm & Cheng, 2007).

Empirical work with adults

Prior work has argued that it is particularly advantageous to keep track of stable causal relationships to generalize knowledge to new situations and contexts and demonstrated that adults prefer stable causal relationships over unstable ones in generalization and intervention (Blanchard et al., 2018; Lombrozo, 2010; Lombrozo & Carey, 2006; Woodward, 2006, 2010). For example, researchers presented participants with a hypothetical scenario of a supplement that is supposed to increase bone density and manipulated whether participants received information that the supplement had a stable effect on bone density (non-moderated group) or whether the effect varied depending on whether or not participants carried a specific gene (moderated group). Adults in the non-moderated group were more likely to agree with causal generalizations that the supplement increases bone density, even though the moderated and non-moderated causes had equal causal strength on average. Participants were also more likely to intervene on the stable cause (i.e., decide to take the pill to increase their own bone density, under the conditions of uncertainty about the background, Vasilyeva et al., 2018). This research shows that adults prefer stable relationships over unstable ones when making causal generalizations across contexts and when deciding whether to intervene on a cause to produce an outcome.

However, most causal relationships are not stable, as they can be influenced—at least to some degree—by other variables. For example, even the relatively stable causal relationship of water boiling at 100 degrees Celsius is impacted by the altitude at which the water is set to boil. Basing one's predictions on the assumption of causal stability alone might therefore be misleading (Cheng, 2000). Adults seem to monitor which relationships are unstable, and use this information to make rich inferences. For example, when participants learned that a pill's side effect (headaches) varied across treatment groups, they inferred that the causal relationship between pill and headaches might interact with another non-observable background factor, suggesting that people can infer the influence of additional background variables when they encounter *unstable* causal relationships (Liljeholm & Cheng, 2007).

Empirical work with children

For children, learning about and understanding causal relationships is particularly crucial, given that they are navigating the world with less data and experience than adults. Research indicates that young children are motivated causal learners: they spontaneously intervene on novel causal systems to infer the underlying causal structure, and form predictions about outcomes of their interventions

Previous work suggests that children show some sensitivity to causal stability, and can use this information to guide their interventions (Cheng et al., 2013). In a recent study, children learned about farm and zoo animals that developed red dots and were treated with a specific diet: farm animals received a grain diet, and zoo animals received a grain-and-leaves diet. After children learned about the underlying probabilistic causal relationships, they had to choose which diets to administer to two new animals with red dots to make the dots disappear. The results show that children consider whether the effects observed in the animals can be attributed to grain alone or must involve an interaction of grain and leaves. In particular, when the grain diet had the same causal strength across the two contexts, children chose to feed the animals grain. In contrast, when the effects varied across contexts, and red dots disappeared more often with a grain-leaves-diet, they were sensitive to the differences in outcome due to the influence of the leaves and opted for the grain-and-leaves diet (Cheng et al., 2022).

Moreover, children seem to be quite open to the possibility that causal relationships may not be stable over time. For example, Sumner et al. (2019) presented 4- to 12-year-old children and adults with a dynamic game environment in which they had to identify a reward-generating monster out of four options, across 80 rounds. In one condition, the reward-generating monster was switched after 40 trials, so participants had to explore the four options again to find it. Adults took much longer to detect the target monster's change than children. This highlights the learning advantages of prolonged exploration, as it allows to detect consequential environmental changes that moderate causal relationships (Sumner et al., 2019).

More generally, young children are sensitive to the impact of background factors on causal relationships. In particular, children as young as 2 years begin to understand that factors such as social norms and moral beliefs can impact the causal behavior they observe (Chernyak & Kushnir, 2014; Kalish & Shiverick, 2004; Rakoczy et al., 2008; Smetana, 1981; Turiel, 1983). By the age of 3, they begin to understand that emotions influence behavior (Harris, 1989; Lagattuta & Wellman, 2001), and at the age of 4 children explain variations in observed causal behavior citing situational factors as reasons (Seiver et al., 2013).

Taken together, this evidence suggests that even young children possess the cognitive competencies required to engage in reasoning about the stability of causal relationships. Yet, to our knowledge, no study to date has examined how young children explore stable and unstable causal relationships. Any real-world agents with limited resources must select what new data to pursue. Do children prioritize reducing uncertainty associated with unstable causes by selectively testing them in novel backgrounds? Or are children equally interested in collecting new data about stable and unstable causes—perhaps targeting a higher level uncertainty about whether the causes are indeed reliably stable or reliably unstable across a broad range of backgrounds? Answering these questions promise to expand our understanding of the factors shaping children's active learning about the world.

The present study

We investigated how 5- to 7-year old children respond to the stability of probabilistic causal relationships across contexts. This age range captures a critical period in the development of skills relevant for exploration behavior (i.e., attention, memory, executive functions) (Diamond, 2013; Roebers et al., 2012). We developed a novel information-search task that allowed us to examine, on the one hand, how children dynamically explore different agents and backgrounds to learn more about their causal stability and, on the other hand, how the stability of the causal relationships under investigation impacts children's exploratory patterns. Children were introduced to two probabilistic causes (monsters) that were equated in average strength (probability of producing a lightning-bolt outcome) but varied in stability across backgrounds (planets). We measured whether children attributed outcomes to causes, backgrounds, or their combinations and whether they wanted to explore stable, unstable, or unknown causes in familiar and unfamiliar backgrounds.

The overall objective was to examine whether children are sensitive to causal stability and, if so, how it shapes their causal attributions and their exploration behavior. Specifically, first, how do children attribute outcomes to causes and backgrounds? If they notice and appreciate that backgrounds play an important moderating role in unstable relationships, they should attribute causal outcomes to the combinations of causes and backgrounds (monsters and planets) rather than to causes or backgrounds alone. Second, how do children select what causes to explore and intervene on, stable or unstable, in novel background contexts? Does their preference, if any, change with age? If children, like adults, prefer stability, they should intervene on stable causes (i.e., they should pick the stable monster). If they have not developed this preference yet, they should choose at chance. Yet another possibility is that children might have a preference opposite to that of adults, and favor unstable causes (i.e., they should pick the unstable monster). Third, what kind of information about causal relationships is a primary driver of children's exploration decisions, information about causal stability across backgrounds, or information about average causal strength (i.e. previously observed probability of an outcome associated with a given cause or background), or some other metric such as minimum or maximum causal efficacy observed so far?

Method

Participants

We recruited 60 children (27 female, M = 74.70 months; SD = 11.57 months; range: 60 to 95 months) through the participants' database of the Max Planck Institute in Berlin and tested them online via the Big Blue Button software. An additional 27 children were excluded from the analyses because they were too young (n=7), or failed to answer the comprehension check questions correctly (n = 20: 13 5-year-olds, 7 6-year-olds). Parents signed an informed consent form, and children agreed by giving verbal assent. The study was preregistered via OSF[1] and approved by the ethics committee of the Max Planck Institute for Human Development ethics committee in Berlin (N-2021-01). The sample size was determined by conducting a simulation-based a-priori power calculation, to detect a hypothesized effect of .15 (Cohen's d) for an interaction of age and task with .80 power and an alpha significance level of .05. The initially registered age range was 5 to 6 years old, but it was expanded to include 7-year-olds prior to data collection.

Design, Materials and Procedure

Participants sat next to their parents in front of a computer and were introduced to a game via screen share. The study consisted of three phases: familiarization, exploration without feedback, and free-exploration with feedback. An attribution question was presented twice, first after the familiarization phase, and second after the exploration phase.

Familiarization Children were introduced to a space-themed game in which they observed two types of causes (turquoise/yellow monsters) generating a probabilistic outcome (energy in the form of lightning bolts) on different backgrounds (red/blue planets). For example, on one learning trial, a group of 8 yellow monsters traveled to a red planet. Upon landing, some of them produced energy (visualized as overlaid lightning bolts), and children were asked to count the lightning bolts (with encouragement to re-count if they made an error). *Stable* monsters produced lightning bolts with a rate of 5/8 on both planets. *Unstable* monsters produced energy with a rate of 3/8 on one planet and 7/8 on the other planet. Importantly, on average, both the stable and unstable monsters produced the same amount of energy (10 out of 16 observations, see Planets 1 and 2 in Table 2.1). By the end of the familiarization phase, a child would have seen fewer lightning bolts

[1] Link: https://osf.io/2xb98

on one of the planets (8 bolts total, composed of 5 and 3 bolts produced by the stable and unstable monsters, respectively) compared to the other planet (12 bolts total, generated by the stable (5) and unstable (7) monsters); we refer to these planets as *low-energy* and *high-energy* planets, respectively (monster and planet colors counterbalanced across participants). Once children had observed one group of monsters visiting both planets, they completed a comprehension check, indicating whether the monsters produced the same energy on both planets or more energy on one of the two planets. Children completed one comprehension check for the stable and the unstable monsters, respectively. Children who failed these comprehension checks repeated the familiarization phase. If they failed to answer the comprehension checks after three familiarization rounds, they were excluded from the sample (n = 20; 10 female; M = 70.05 months; SD = 6.35). At the end of the familiarization phase, children were presented with a summary slide showing both monsters next to both planets, with the number of energy bolts they had produced on each planet.

Attribution questions After the familiarization phase, children completed the first causal-attribution task. They were presented with three statements attributing the outcome ("lightning bolts happen...") either to the causes ("because of the monsters"), or to the backgrounds ("because of the planets"), or to both (" because of the monsters and the planets"); each claim appeared in a speech bubble of a uniquely-colored unicorn. Children selected the unicorn they thought was right. At the end of the study, children were again presented with the same attribution question (unicorn colors counterbalanced within and between participants).

Exploration phase without feedback Children made six decisions, each involving a choice between two options. In three decisions, they chose between two causes, and in the remaining three they chose between two backgrounds (order pseudo-randomized). They did not receive any feedback about the outcomes of their choices. Out of the six decisions, two involved a novel element: either a novel cause or a novel background. The key decision trial (*novel planet*) assessed children's preference for intervening on a stable or unstable cause under conditions of background uncertainty: a child was presented with a novel planet (unfamiliar background), and was asked to decide whether to send a stable or unstable monster to this planet. A preference for exploring (un-)stable causes would manifest in selecting the respective monster in this task. (Note that, like many real-world decisions, this task can be construed as having elements of exploitation—applying prior knowledge to generate a desired outcome—and exploration—learning how a cause functions in a previously unexplored background. However, two features of this task maximize its exploratory character:

first, the two causes were equated in average causal strength, such that the expected probability of generating an outcome (lightning bolts) could not offer guidance for selecting one monster over the other in a new background; second, children did not receive any rewards or prompts to produce a high number of energy bolts at any point throughout the task; we say more on this in the Discussion.)

The second key decision trial examined whether children tracked the average causal strength of familiar backgrounds and used this information to inform their interventions. On this trial (*novel monster*), children were presented with a new monster and could decide whether they wanted to send it to the low or high-energy planet. If children aim to maximize the chances of producing the outcome based on the average causal strength they should pick the high-energy planet. Again, no incentives or prompts to produce a high number of energy bolts were given.

The remaining four decisions involved familiar combinations of causes and backgrounds that children had previously encountered during the familiarization phase. These questions allowed us to assess the extent to which children's choices were driven by general preferences for (in-)stability vs. by maximizing expected outcomes based on the previously observed average causal strength of each variable. On two *old planet* trials, children saw one planet (either the low or the high-energy planet) and chose whether to send there the stable or unstable monsters. If a preference for stability drives children, they should pick the stable monsters on both trials; if they prefer instability, they should consistently send the unstable monsters. If the causal strength instead drives their preferences, they should send the stable monster to the low-energy planet and the unstable monster to the high-energy planet to maximize outcomes. On the remaining two *old monster* trials, children saw one monster type (either the stable or unstable) and chose whether to send them to the low or the high-energy planet. Because causal stability offers no grounds for preferring one planet over the other, children relying on stability alone should choose at chance. If they rely on causal strength instead, they should always pick the high-energy planet. Since children received no feedback about the outcomes of their choices (i.e., they did not get to see what happened after the selected monsters traveled to the planets, etc.), they did not accumulate in this phase any new data about causal stability or average causal strength.

Free-exploration phase with feedback Children tried out different combinations of familiar and novel causes across familiar and novel backgrounds and observed the outcomes. We wanted to mirror an everyday situation where children encounter various causal relationships across different backgrounds and have the chance to explore the relationships freely, without explicit guidance or incentives. Therefore, we did not incentivize or encourage them to generate as much energy as possible. With this approach children may be more likely to explore and engage with the task in a more open-ended way, which can reveal their underlying cognitive processes and strategies without being narrowly focused on a particular goal. Children could choose between five different types of monsters (see Figure 2.1), send each monster type (in groups of eight) to one of six different planets of their choice, and observe how many of the eight monsters generated energy bolts on a particular planet. This task allowed us to examine what types of causes children were most interested to explore (and re-explore).

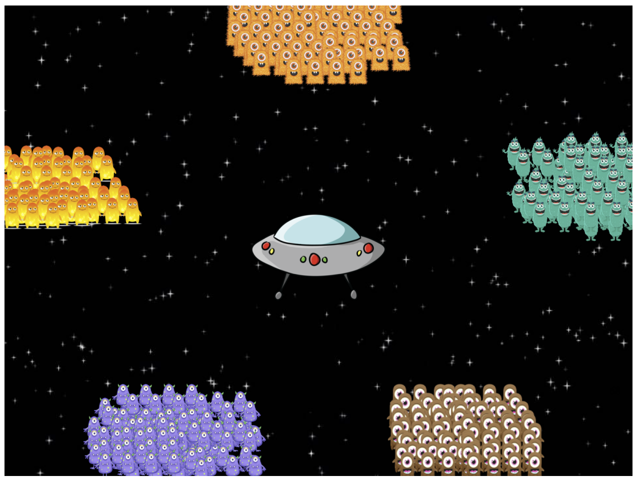


Figure 2.1. Screenshot from the free-exploration phase with feedback, with five cause options. Once a child selected one monster type, a group of eight monsters boarded the space shuttle and the child proceeded to select one planet to send the monsters to.

The five monster options included the stable and unstable monsters from the familiarization phase and three new monster types (monsters generated energy following one of the patterns shown in Table 2.1). The familiar monsters continued to produce energy displaying the previously observed patterns: the *old stable* monsters always generated energy with the rate of 5/8; the *old unstable* monsters alternated between generating 3/8 and 7/8 energy bolts across planets. The three new monster types included the *new stable low* monsters, which produced 1 energy bolt across the 8 monsters on all planets; the *new stable high* monsters produced 7 energy on all planets; finally, the *new unstable* monsters alternated between 0, 2 and 8 energy bolts depending on the planet they visited. Energy-production patterns were counterbalanced across monsters. The planets included the familiar red and blue planets and four novel planets (including one novel planet featured in the Exploration without feedback task). At the end of each exploration round, children counted the energy bolts produced. If they miscounted, they were encouraged to recount. Every four rounds, we asked children whether they wanted to "continue or stop" playing the game (wording counterbalanced within subjects: "stop or continue").

Table 2.1. Free-exploration phase with feedback: number of monsters generating energy bolts (out of 8 monsters), across different planets.

	Planet					
Monster	P1	P2	P3	P4	P5	P6
Stable	5	5	5	5	5	5
Unstable	3	7	3	7	3	7
New low	1	1	1	1	1	1
New unstable	8	2	0	8	2	0
New high	7	7	7	7	7	7

Results

Attribution Questions

On the first attribution question, the majority of the children (48.33%) attributed the outcome (energy bolts) to the combination of causes and backgrounds, i.e., monsters and planets, rather than to causes, i.e., monsters (26.67%), or backgrounds, i.e., planets (25.00%, $\chi^2(2) = 6.100$, $p = .047$). On the second attribution question, children's choices followed the same ordering, with 38.98% of children attributing the outcome to the combination of causes and backgrounds, 37.29% attributing it to causes, and 23.73% attributing it to backgrounds; however, these differences were not significant, $\chi^2(2) = 2.475$, $p = .290$. The difference between the two rounds of attribution questions was not significant ($\chi^2(2) = 1.399$, $p = .497$). To investigate whether the propensity to attribute outcomes to interactions changes with age, we re-coded responses into a binary variable, attributions to the cause x background interaction vs. a single factor (either causes or backgrounds). Age in months did not significantly predict these responses in a logistic regression, $p = .707$, $OR = 0.991$ [0.946 – 1.039].

Exploration phase without feedback

We began by examining choices on the *novel planet* trials, where children had the option of intervening on either the stable or unstable cause in a novel background (i.e., sending either stable or unstable monsters to a new planet they had no prior information about). Overall, 45% of the children sent the *stable* monsters to the new planet, which did not significantly differ from chance

(50%, p = .519, exact binomial test). A logistic regression predicting choices from age revealed no developmental change (p = .238, OR = 1.028 [0.982 – 1.076]) (see Figure 2.2).

We then turned to the two *old planet* trials, where children selected which monsters to send to the low and high-energy planets. This allows us to assess how children apply the prior evidence they gathered during the familiarization phase about causes and backgrounds in designing interventions. On average, when presented with the low-energy planet, half of the children chose the stable and the other half the unstable monsters (50%, p = 1.000, exact binomial test). When presented with the high-energy planet, 57% of the children preferred to send the stable monsters, which did not differ from chance (p = .366, exact binomial test). A logistic regression predicting monster choice from planet type (low vs. high-energy planet) and children's age in months revealed that age alone (p = .305, OR = 1.024 [0.979 – 1.071]) did not predict children's decisions. However, the type of planet presented (p = .067, OR = 97.194 [0.729 – 12958.617]) marginally predicted children's decisions. Most importantly, the interaction of age and the presented planet was significant (p = .050, OR = 0.937 [0.878 – 1.000]); as shown in Figure 2.2, with age children became more selective, sending the unstable monsters more to the low-energy planet, and the stable monsters more to the high-energy planet.

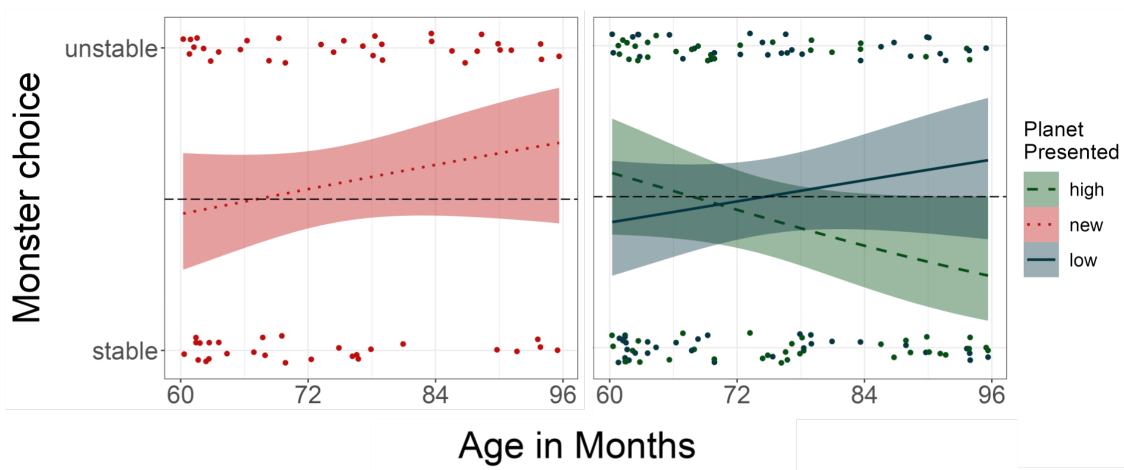


Figure 2.2. Children's choices between stable vs. unstable causes (monsters) in the exploration without feedback phase, when presented with the novel background (planet) *(left panel)*, or when presented with the old backgrounds (high-energy vs. low-energy planet) *(right panel)*. Each dot represents a child's choice between the stable and unstable causes. The lines indicate a fitted logistic regression.

Next, we examined responses from the *novel monsters* trial, where children chose to send new monsters they had no data about to either the low or high-energy planet. We found no evidence that children relied on average causal strength in this decision: 57% of children sent the new monsters to the high-energy planet, exact binomial test against chance 50%, p = .366. A logistic regression revealed that their choices did not vary with age (p = .266, OR = 0.974 [0.931 – 1.020]).

Children's choices on the two *old monsters* trials, where they were presented with either the stable or the unstable monsters and selected a planet (low or high-energy) to send each monster group to, also did not reveal significant preferences. When presented with the stable monsters, 55% of the children sent them to the high-energy planet (p = .519, exact binomial test). When presented with the unstable monsters, 55% decided to send them to the high-energy planet (p = .519). A logistic regression predicting planet choice from monster type (stable vs. unstable) and children's age in months revealed no effects of age (p = .616, OR = 0.989 [0.946 – 1.034]) or the type of monster (p = .846, OR = 1.606 [0.014 – 189.426]), and the interaction of age and monster type was not significant (p = .844, OR = 0.994 [0.933 – 1.059]).

Free-exploration phase with feedback

On average, children performed 8.26 (SD = 6.89) rounds of explorations. Most children (78%) tried at least some monsters more than once, and 38% of children re-explored all monsters. To investigate this further, we specified monster type as a predictor of whether children re-explored it. Overall, monster type predicted re-exploration behavior: the new low monster (always producing 1 energy bolt) (p = .027, OR = 3.725, [1.163 – 11.936]) and the new unstable monster (producing 0/2/8 energy bolts alternating) were significantly more likely (p = .006, OR = 5.287, [1.600 – 17.475]) to be re-explored than the old unstable monsters.

Discussion

We investigated whether 5- to 7-year-old children's active-exploration strategies are sensitive to the (in-)stability of causal relationships. We find that overall, in this age range, children can already appreciate the interactive nature of causal relationships. After they were presented with evidence that some causes act differently in different backgrounds, they attributed outcomes to a combination of causes and backgrounds rather than to either causes or backgrounds in isolation.

We find evidence that, with age, children use prior evidence more in designing interventions involving familiar combinations of causes and backgrounds that lead to low outcomes. In the exploration without feedback phase, older children tended to be more selective, using the *unstable* cause in the background context where this cause had been previously less effective (3/8, which is lower than the stable cause's performance of 5/8) and switching to intervening on the *stable* cause in the background where this cause had been less effective during the familiarization phase (5/8, which is lower than the unstable cause's performance of 7/8). This reveals an increasing capacity to track and integrate information about causes and backgrounds to guide exploration decisions.

While the results reported above are promising, we failed to find evidence that children's active exploration is directly guided by causal stability or by a preference for causal strength. This is surprising, given the prior empirical evidence that preschoolers are sensitive to stable causal relationships (Cheng et al., 2022). Instead we found that children's active exploration targeted causes with the lowest minimum observed causal efficacy. For example, in the exploration with feedback phase children repeatedly explored monster groups generating zero or one energy bolts in at least some backgrounds. This could be due to a variety of factors. One possibility is that children were simply drawn to low-energy outcomes for reasons beyond our study setup—perhaps they are budding environmentalists, who had learned that saving energy is crucial from their parents or at school. Another possibility is that the interest in exploring ineffective causes stems from children's expectations that these causes are unstable across backgrounds. Perhaps they were trying to find a background where these causes would turn out highly effective (looking for a "jackpot"). One way to examine this further would be to compare the exploration behavior of children who had and had not been exposed to unstable relationships beforehand. This lies beyond the scope of this paper. In our study, all children had witnessed unstable relationships in the familiarization phase, which likely made the possibility of contextual variability in causal strength more salient to all of them, which could make them seek fortuitous backgrounds for ineffective causes.

Children's general lack of preference for stable or unstable causes in exploration tasks can reflect several things. First, children may have been uncertain about what would be most beneficial to learn in this task. We did not offer incentives for generating outcomes, and the valence of the outcome was left ambiguous (we did not offer any guidance on whether it is better to produce as much energy as possible or to save energy); this openness could have resulted in high variability across children in our sample in terms of what each of them was trying to discover or achieve during exploration. Second, this may have been a challenging task with too many choice options. Since one must explore a cause in at least two backgrounds to know whether it is stable or unstable, the task of determining stability for five causes may have exceeded children's capacity. Reducing the free-exploration phase to four monsters and planets and ensuring all children gathered enough data about all causes might provide more precise insights in a future study. Third, our findings might mean that causal (in)stability does not matter in the context of exploration tasks (although Sumner, 2019, suggests otherwise). At this point, we do not have adult data for comparison; it is possible that while adults show stability preference in some tasks, they do not rely on it in exploration.

We are preparing a set of follow-up studies that address these possibilities by i) incentivizing children for each energy bolt they produce; ii) clearly stating that it is favorable to produce energy, for instance, by asking children to help the monsters restart their space shuttle by collecting as many energy bolts as possible; iii) running the study with older children (8- to 10-years-old) and adults and comparing the results between and within these age samples and the current sample; iv) requiring each child to explore all monsters at least twice, providing access to stability information for exploration decisions; v) implementing a computational approach to compare children's behavior against computational agents with a perfect preference for stability, instability, and causal strength.

In sum, children show signs of sensitivity to causal instability between 5 and 7 years of age (as revealed by their causal attributions). They become increasingly adept at tracking causal efficacy across multiple factors at the same time (causes and backgrounds), but they do not yet put this understanding to use to guide their exploration behaviors; at least, they do not show a blanket preference to explore stable or unstable causes; the possibility that causal (in)stability guides them in more subtle and indirect ways remains open and will be assessed in future studies.

References

Blanchard, T., Vasilyeva, N., & Lombrozo, T. (2018). Stability, breadth and guidance. *Philosophical Studies*, *175*(9), 2263–2283. https://doi.org/10/grm7qf

Cartwright, N. (1989). *Nature's capacities and their measurement.*

Cartwright, N. (2009). What are randomised controlled trials good for? *Philosophical Studies*, *147*(1), 59. https://doi.org/10/d9zf9c

Cheng, P. (2000). Causality in the mind: Estimating contextual and conjunctive power. In *Explanation and cognition* (pp. 227–253). The MIT Press.

Cheng, P., Liljeholm, M., & Sandhofer, C. (2013). Logical consistency and objectivity in causal learning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *35*.

Cheng, P., & Lu, H. (2017). Causal invariance as an essential constraint for creating representation of the world: Generalizing the invariance of causal power. In *The oxford handbook of causal reasoning* (pp. 65–84).

Cheng, P., Sandhofer, C. M., & Liljeholm, M. (2022). Analytic causal knowledge for constructing useable empirical causal knowledge: Two experiments on pre-schoolers. *Cognitive Science*, *46*(5). https://doi.org/10/grp9fh

Chernyak, N., & Kushnir, T. (2014). The self as a moral agent: Preschoolers behave morally but believe in the freedom to do otherwise. *Journal of Cognition and Development*, *15*(3), 453–464. https://doi.org/10/grp9fg

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10/b2m2

Harris, P. L. (1989). *Children and emotion: The development of psychological understanding.* Basil Blackwell.

Kalish, C. W., & Shiverick, S. M. (2004). Children's reasoning about norms and traits as motives for behavior. *Cognitive Development*, *19*(3), 401–416. https://doi.org/10/dqch48

Lagattuta, K. H., & Wellman, H. M. (2001). Thinking about the past: Early knowledge about links between prior experience, thinking, and emotion. *Child Development*, *72*(1), 82–102. https://doi.org/10/bf3fpm

Liljeholm, M., & Cheng, P. (2007). When is a cause the "same"?: Coherent generalization across contexts. *Psychological Science*, *18*(11), 1014–1021. https://doi.org/10/fb26x7

Lombrozo, T. (2010). Causal–explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology*, *61*(4), 303–332. https://doi.org/10/fvx8g3

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, *99*(2), 167–204. https://doi.org/10/fsdg8x

Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological Review*, *121*(3), 277. https://doi.org/10/f6ds5k

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology*, *44*(3), 875–881. https://doi.org/10/br6p85

Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: An explorative study on their interrelations and their role for school achievement. *Metacognition and Learning*, *7*(3), 151–173. https://doi.org/10/gfvw2c

Seiver, E., Gopnik, A., & Goodman, N. D. (2013). Did she jump because she was the big sister or because the trampoline was safe? Causal inference and the development of social attribution. *Child Development*, *84*(2), 443–454. https://doi.org/10/f4s8pt

Smetana, J. G. (1981). Preschool children's conceptions of moral and social rules. *Child Development*, 1333–1336. https://doi.org/10/d7cskx

Sumner, E., Li, A. X., Perfors, A., Hayes, B., Navarro, D., & Sarnecka, B. W. (2019). The exploration advantage: Children's instinct to explore allows them to find information that adults miss. https://doi.org/10/grqb4n

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.

Vasilyeva, N., Blanchard, T., & Lombrozo, T. (2018). Stable causal relationships are better causal relationships. *Cognitive Science*, *42*(4), 1265–1296. https://doi.org/10/gdp5bc

Woodward, J. (2006). Sensitive and insensitive causation. *The Philosophical Review*, *115*(1), 1–50.

Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, *25*(3), 287–318. https://doi.org/10/b2wbmr

Chapter 3

Preschoolers Select the Relevant Information When Looking for a Hidden Present

Serko, D., Li, Y., Swoboda, N. & Ruggeri, A. (2023).

Proceedings of the 45th Annual Conference of the Cognitive Science Society.

This paper can be accessed at:

https://escholarship.org/uc/item/6085s468

Abstract

Previous research suggests that children's information search remains largely inefficient until age 4. Here, we investigate the early emergence of children's information-search competence using a simplified version of Lindow's (2021) *finding-presents* game. Children (n = 86, 25- to 59-months old) had to find a present hidden in one of three closed boxes. All boxes were identical but for one feature (e.g., all boxes were blue and had a flower icon on top, but one box was round, one heart-shaped, and one squared). To identify the target box, children received three information cards revealing one feature of the target box (i.e., its color, shape, or icon). As the boxes differed in only one feature (e.g., their shape), only one information card contained the relevant information to the decision (i.e., the information card indicating the correct shape). Children could flip one information card to learn about one particular feature before deciding which box to open. This was our dependent measure. Our findings indicate that children as young as 2 years can efficiently search for information to guide their decisions and underline the importance of using age-appropriate paradigms.

Introduction

More information is not always better. Some information is indeed helpful, but some information—even when accurate—may be irrelevant, resulting in a waste of time and resources, or potentially generating confusion even. For example, instead of reading this paper, which will provide the reader with new evidence on young children's information search abilities, potentially relevant for their work, the reader could watch a compilation of the 100 cutest cat videos on YouTube—which would surely also provide a lot of interesting information, though probably (hopefully?) not as relevant. In this sense, being able to tell apart relevant, high-quality information from irrelevant and low-quality information, and relying on the former, is a crucial competence supporting learning in the social, digital, and physical world.

Work from the decision-making literature found that this sensitivity to the informativeness and relevance of different cues (i.e., pieces of information) develops rather late, reaching adult-like efficiency only by adolescence, or even later (Betsch et al., 2018; Davidson, 1991, 1996; Mata et al., 2011).

For example, Betsch et al. (2018) presented 5- to 10-year-old children and adults with a game in which they had to decide which cues to look up to find a treasure. Children were first familiarized with three animals, differing in their ability to predict the treasure's location correctly, and then with the animals' suggestions about the treasure's location. The authors found that preschoolers failed to integrate the probabilistic information about the animals' accuracy in their information-search decisions. This ability improved with age, with children beginning to show information-search strategies comparable to adults' by age 9.

In line with these results, research from educational psychology suggests that 4- to 6-year-old children have difficulties in understanding when enough information has been collected to be sure about something (referred to as *determinacy* or *indeterminacy* of evidence; Fay & Klahr, 1996; Klahr & Chen, 2003). In particular, children this age often *overestimate* the informativeness of the available evidence. For example, when presented with indeterminate evidence, children are overly optimistic about knowing the answer and tend to ignore that additional evidence would be required to support their answers (Klahr & Chen, 2003).

However, a growing body of work from developmental and cognitive psychology paints a much more optimistic picture, suggesting that the foundations required to support efficient information search may instead emerge very early in life. Research with infants indicates that systematic patterns of efficient information-seeking start emerging during the first months of life and become increasingly explicit and selective between the first and second year of life when infants can promptly and effectively signal their uncertainty and elicit information from the most informative sources available (for an overview, see De Simone & Ruggeri, 2022).

For instance, by 5 months of age, infants are already sensitive to the likelihood of a social partner being informative; that is, they look longer at partners who express willingness to convey information, for instance, by making eye contact, calling their name, and using infant-directed speech (Cooper & Aslin, 1990; Csibra & Gergely, 2009; Senju & Csibra, 2008).

This ability to discriminate partners by their informativeness sets the foundation for young children to *actively search* for information from their partners and in their environment. Studies using preferential looking as a measure for information search indicate that pre-verbal infants are sensitive to the novelty and quality of information. In particular, infants' looking-time increases when they encounter new objects (Kutsuki et al., 2007), when two novel objects are labeled with the same label (Hembacher et al., 2017; Vaish et al., 2011), or when they encounter something unexpected such as the disappearance of a puppet (Dunn & Bremner, 2017; Walden et al., 2007).

Beyond being selective in deciding what information and information sources are most likely to be informative, recent work suggests that infants look at other people to *actively solicit* information, suggesting that pretty much the same events and stimuli that trigger infants' perceptual interest (e.g., novelty of objects, violation of expectation, confounded evidence) also result in increased references to their social informants (Dunn & Bremner, 2017; Hembacher et al., 2017; Kutsuki et al., 2007; Vaish et al., 2011; Walden et al., 2007) and enhanced exploration (Stahl & Feigenson, 2015).

This early competence rapidly matures over the first years of life. Ruggeri et al. (2019) demonstrated that already by their third year of life, children are able to successfully tailor their information search strategies to the characteristics of the task they are presented with. In this study, children had to find an egg shaker hidden in one of four small boxes, which were, in turn, contained in two larger boxes. They were allowed to open *only one* large box, but they could shake one or both large boxes first if they wanted to. Crucially, before this test, children learned that either the egg was equally likely to be found in any of the four small boxes (uniform condition) or it was most likely to be found in one particular small box (skewed condition). Results show that preschoolers as young as 3 years successfully tailored their exploratory actions to the different likelihood distributions: Compared to the skewed condition, where children had a strong intuition as to where the egg shaker would be hidden, children in the uniform condition were more likely to shake a large box first. This way, they could hear which large box contained the small box with the egg shaker without risking opening the wrong one (Ruggeri et al., 2019).

How could we reconcile these findings from those from decision-making and education reviewed above, describing a much more protracted emergence of efficient information search patterns? We argue that these studies may have failed to capture children's early learning competence because they: (i) implemented paradigms that were too complicated or abstract for children to understand, relate to, or care about. For example, Ruggeri and Feufel (2015) compared the performance of 7-to 10-year-old children with adults in a 20-question game and found that they asked less informative questions when presented with professions rather than animals, highlighting the strong impact of domain-specific knowledge on question-asking competence (Ruggeri & Feufel, 2015; Ruggeri & Katsikopoulos, 2013); (ii) presented instructions, stimuli or tasks that required advanced math skills or verbal competences that just cannot be expected to be mastered until late childhood. For example, succeeding at the treasure hunt game by Betsch et al. (2018) requires a pretty sophisticated understanding of differences across probabilistic distributions that preschoolers (or at least some of them) may still be developing (Betsch et al., 2018); (iii) did not consider that children (and children of different ages, or Socio Economic Status) may be bringing in different assumptions to the task than what the researchers expected, potentially leading children to apply a different, yet ecologically effective, default strategy for active learning. For example, children may ask a question intended to confirm or rule out a hypothesis they believe is more likely than others, even though the researchers assume that all the considered hypotheses should be considered equally likely (Bramley et al., 2022).

Indeed, recent studies demonstrate how young children's question-asking performance can improve when presented with more child-friendly instructions and paradigms (Bonawitz et al., 2012; Domberg et al., 2020; Ruggeri et al., 2017; Ruggeri et al., 2019; Ruggeri et al., 2021; Swaboda et al., 2022). For instance, a recent study compared children's performance in a 20-question game and in a spatial-navigation task, in which they had to discover the path through a maze by removing masks covering its passages, and found that children searched more efficiently when they could make queries non-verbally (Swaboda et al., 2022). Along these lines, Lindow (2021) implemented a more child-friendly version of the treasure hunt game paradigm used in previous work (Betsch et al., 2014; Betsch et al., 2018) and found that, in simpler search environments, even 5- to 6-year-old children managed to select information effectively, compared to previous studies indicating ineffective information search until age 9 (Betsch et al., 2014; Betsch et al., 2018).

The current Study

In this project, we developed a novel version of the treasure hunt game developed by Betsch et al. (2014), Betsch et al. (2018) and then simplified by Lindow (2021), to examine the emergence of information-search efficiency in 2 - to 4-year-olds. We focused on this specific age range because we wanted to address younger children than in the initial study of Lindow (2021), while making sure that children understand basic verbal instructions and can indicate their information card choices verbally or by pointing at or crawling towards them. In contrast to Lindow's finding-presents game, our game version presented children with three boxes varying *on one specific feature* of the box, instead of four identical boxes with varying icons on top of them. In particular, we held two features constant (e.g., all boxes were blue and carried an icon of a flower on top of them), but we varied one feature (e.g., one box was round, one heart-shaped, and one squared). Three information cards provided information about the color, shape, or icon of the target box containing the present. Two information cards were irrelevant (information about color and icon did not help to disambiguate the location of the toy, as all boxes had the same color and the same icon), but one of the information cards contained the relevant information (the card indicating the specific shape of the target box allowed to find the present). Children indicated their information card choice verbally or by pointing, which allowed us to reduce verbal demands. Our design required no understanding of probabilities and avoided using distractor cards. We believe these simplifications allowed us to reduce verbal demands and thus target younger children while maintaining the overall structure of the task proposed by Lindow (2021). We hypothesized that children would select the informative cue card significantly above chance and that this ability would improve with age.

Methods

Participants

To ensure that children understood the task instructions and found the materials and procedure engaging, we piloted the experiment prior to data collection. The pilot sample included 54 participants (30 female; M = 41.16 months; SD = 9.77 months) tested at a local museum in Berlin, Germany.

In the final study, we tested 86 children between 25 and 59 months (46 girls, M = 41.50 months, SD = 8.81 months). They were recruited at a local museum or via the internal participant database of the Max Planck Institute for Human Development. We tested participants in the museum right after recruitment or in the lab. An additional 28 children (15 female; M = 34.78 months; SD = 8.48

months) were tested but excluded from the analysis because they were not concentrating on the task (n=8), they were too shy to interact with the experimenter (n=3), had language difficulties (n=2), failed to pass the training phase (n=1), because of experimenter error (n=8), parental intervention (n=2), or technical problems (n=4; in total: 17 2-year-olds, 8 3-year-olds, 2 4-year-olds, one participant did not provide a date of birth).

Written informed consent of legal guardians was obtained prior to participation. Children were asked for verbal assent before the study and received stickers as a reward for their participation after the study (see Design and Procedure). The study was approved by the ethics committee of the Max Planck Institute for Human Development (NC-2022-11). The sample size was determined by conducting a-priori power calculations via simulation for each planned statistical test. The most conservative estimate indicated an overall sample of 80 children to detect the estimated effect size (Cohen's h = 0.6) with 90% power using binomial logistic regression with a 0.05 criterion for statistical significance.

Materials

All materials were specifically built and consisted of three sets of cardboard boxes with corresponding information cards (see Figure 3.1 for pictures of all three sets). Each set of boxes consisted of three individual small boxes with removable lids, each with a particular color, shape, or icon on its lid. Within each set, all three boxes shared exactly two features but differed in one specific feature. Each set contained one distinguishing feature (see Figure 3.1). For example, in one set all boxes were blue, had a flower icon on top, but differed in the shape of the box (see example Set 2 in Figure 3.1).

Each set of boxes was accompanied by a set of information cards, which showed all available variants of each feature on their backside as many times as they occurred among the boxes, and the feature variant of the target box on their front side (see Figure 3.1). For example, in Set 2, the *color card* showed three blue splashes of color (since there were three blue boxes) on its back and one splash of blue on its front. The *icon card* showed three flowers on its back (since all three boxes had flowers on top) and one flower on its front. The *shape card* showed a square, a circle, and a heart on its back, and one of these shapes indicated the critical image to find the critical box on its front (e.g., a circle if the target box was round-shaped). One box in each set contained a feather or sticker as a present.

Design and Procedure

We presented children with three boxes, of which one contained a present. The boxes were identical in two features (e.g., all three boxes had the same color and the same sticker on top of them) but differed in one feature (e.g., each had a different shape; as an example, see Set 2 in Figure 3.1). To find out which box contained the present, children could pick one of three information cards, each revealing one of the target box's features (i.e., the color, shape, or icon on top of the box). Only one of the cards revealed the crucial feature necessary to find the target box (in this example, the card identifying the correct shape of the box is the informative card as it is the only feature that allows inferring the target box). Children were allowed to flip only one card but had up to two attempts per test. The experiment consisted of a training phase, familiarizing children with the boxes and the cue cards, and two tests. Sets and target boxes were counterbalanced.

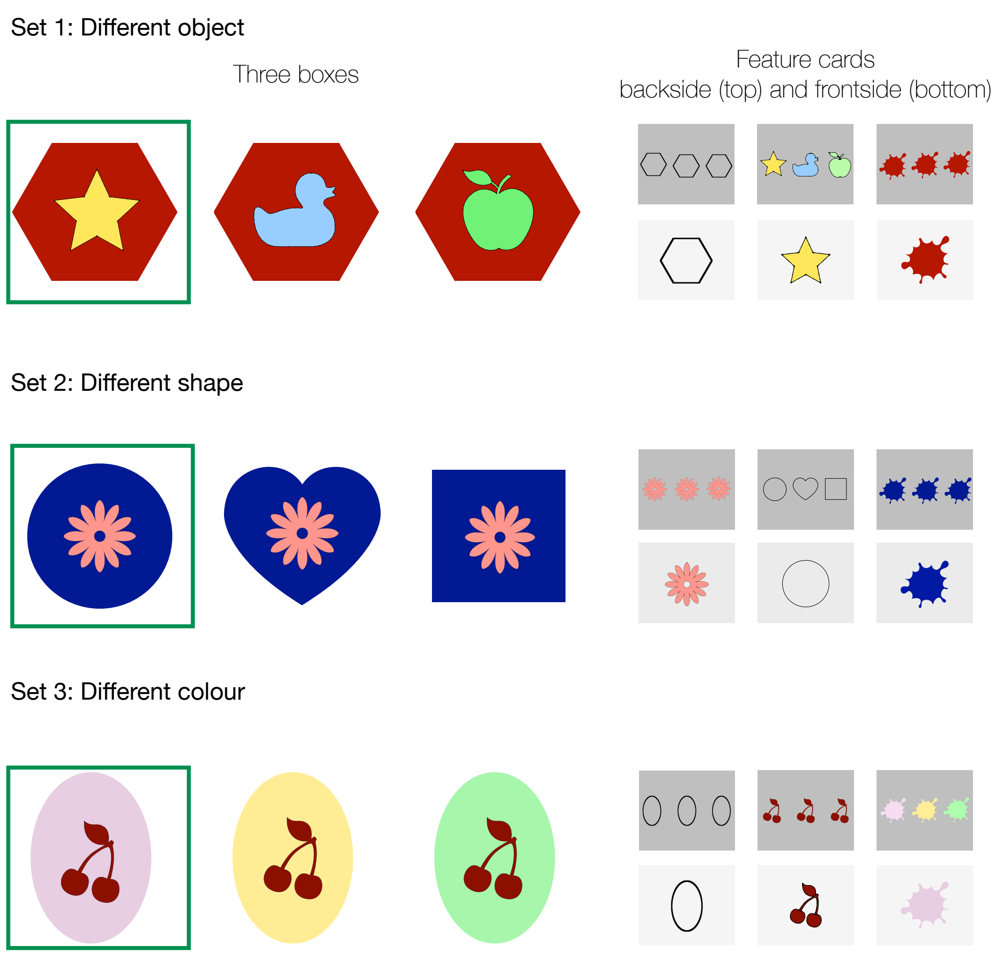


Figure 3.1. Picture showing all sets of boxes and the corresponding information cards used in the experimental procedure. A sticker or a feather were placed in one box of each set as a present.

Training phase. Children were presented with one set of boxes placed on a blanket on the floor (see Figure 3.1, box sets were counterbalanced between participants). They were told that there was a small present hidden in one of the boxes and that the goal of the game was to find out which box contained it. Next, the experimenter familiarized the children with the features of boxes, saying *"Look, all boxes have the same color/shape/icon. They are all [...]."*, for the shared features, or *"Look, all boxes have a different color/shape/icon. This box is [...], this one is [...], and this one is [...]."*, for the one differentiating feature. The differentiating feature was presented either first or last (counterbalanced between participants).

The experimenter then took the cue cards and shuffled them, saying *"We do not know yet in which of the boxes the present is hidden. But to find out, I brought these cards with me. They can help us finding out what the box with the present looks like."*. Next, she placed one by one the cue cards in front of the child, saying, *"Look, this card tells us the color/shape/icon of the box with the present* (see Figure 3.1). Once all cards had been placed down, she turned over the three cards one after the other, from right to left, to demonstrate how they revealed the features of the target box. The procedure for each card was identical: First, the experimenter turned over the card and said, *"When we turn over this card, we know the color/shape/icon of the box with the present."* and pointed out the revealed feature (*"Look, the box with the present is [color/shape/icon]*), and then asked children to indicate all boxes possessing that feature (*"Can you show me all the [color/shape/icon] boxes?"*). If children failed to answer (e.g., because they were shy), the experimenter pointed at the boxes one by one, asking children if that box had the queried feature. For each revealed card, the experimenter emphasized whether the feature was shared by all boxes (*"So, all boxes are [color/shape/icon], right?"*) or different across all boxes (*"Only this box is[color/shape/icon], right?"*).

Once all cards had been turned over, the experimenter summarized what they had learned about the features of the box (*"Now we know that the box with the present is [color/shape/icon]"*) and asked children to point at the target box. If children failed to identify the target box, the experimenter repeated the summary of the features and highlighted once more how they indicated the target box. This was repeated until children were able to successfully indicate the correct box (verbally or by pointing). One child (age 37 months) failed to do so on their own even after several explanations and was excluded from the analysis.

First test. The first test presented children with the *same* set of boxes used in the training phase. The experimenter removed the cards and hid a new present in one of the three boxes while the child looked away. Next, she told the children that there was a new present in one of the boxes and that it could be in the same or a different box than before (we counterbalanced the order in which we said "same" and "different" between participants). Then, as before, she said that children could not know where the present was hidden and that they could look up the cue cards to find out, which she shuffled and placed in front of the child, saying *"Would you like to know about the color, shape, or icon of the box with the surprise?"* (following the order of placement).

Children were then allowed to look up *only one* card, which the experimenter commented, revealing the related feature. If children looked up the cue card with the differentiating feature, she prompted children to point to the target box and retrieve the present.

If children looked up one of the features shared by all boxes, the experimenter commented, saying *"Ah, the box with the present is [color/shape/icon]. But all boxes are [color/shape/icon], aren't they? So you can not really know yet where the present is. Let's try again."*. In this case, the experimenter reshuffled the cards and repeated the procedure.

Second test. To investigate children's information search abilities across different contexts, we conducted a second test. In the second test, we used a different set of boxes than the one used for training and in the first test, with a new differentiating feature. The experimenter introduced the new boxes as she did before and then moved to the test phase, which procedure was identical to the first test.

Results

In the first test, 52 out of 86 (60.47%) children picked the relevant cue card on their first attempt. An exact binomial test revealed that children's choices significantly differed from chance (33%, $p < .001$, binomial test). A logistic regression analysis with age in months as a predictor revealed no significant effect of age ($p = .390$, OR = 1.02 [0.97 – 1.08]).

In the second test, 52 out of 86 (60.47%) children picked the relevant cue card on their first attempt. An exact binomial test revealed that children's choices significantly differed from chance (33%, $p < .001$, binomial test). A logistic regression analysis with age in months as predictor revealed a significant effect of age ($p = .003$, OR = 1.09 [1.03 – 1.16]), indicating that older children were more likely to look at the relevant informative cue card compared to younger children.

To analyze children's ability to pick the correct information card at the first attempt in *both tests*, we created a dummy variable indicating success in both tests. 32 out of 86 children (37%) picked the correct card on the first attempt in both tests. An exact binomial test revealed that children's choices

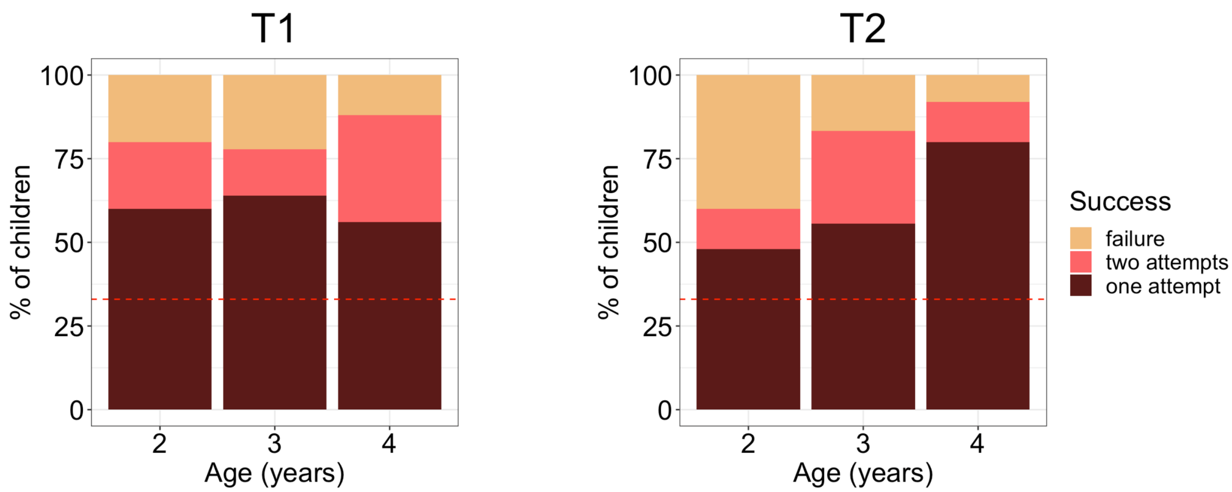


Figure 3.2. Proportion of children looking up the informative cue card by age group in the first (left) and second test (right). Colors indicate whether children looked up the informative cue card on the first or second attempt or not at all. The dashed red line indicates the chance level (33%).

significantly differed from chance (11%, $p < .001$, binomial test). A logistic regression analysis with age in months as a predictor to choose the correct information card at the first attempt in both tests revealed a significant effect of age ($p = .019$, OR = 1.07 [1.01 – 1.13]), indicating that older children were more likely to look up the informative cue cards in both tests compared to younger children.

We controlled for the possibility that children's performance was influenced by the specific set they were presented with, thereby making sure that children were not more likely to pick the correct information card because one of the features was more salient to them than others. A logistic regression with the target feature (color vs. shape vs. icon) as a predictor and success in the first test as the dependent variable revealed no significant effect (object: $p = .492$, OR = 0.64 [1.17 – 2.25]; shape: $p = .413$, OR = 0.54 [0.12 – 2.49]). We found similar results when considering success in the second test as the dependent variable (object: $p = .893$, OR = 1.10 [0.29 – 4.31]; shape: $p = .859$, OR = 0.89 [0.25 – 3.09]).

Discussion

In this study, we investigated the emergence of 2 - to 4-year-old children's information-search efficiency. To do so, we developed a simplified version of the *finding-presents* game implemented by Lindow (2021). Our findings indicate that all age groups performed significantly above the chance level. Even the youngest children in our sample engaged in efficient information search by selecting the informative information cards. When children were presented with a new set of boxes and information cards (i.e., in the second test phase), performance for all age groups remained above chance level, with 4-year-olds performing near ceiling (see Figure 3.2).

Previous studies have stressed young children's ineffective information search (Betsch et al., 2018; Davidson, 1991, 1996; Fay & Klahr, 1996; Herwig, 1982; Klahr & Chen, 2003; Mata et al., 2011; Ruggeri & Feufel, 2015; Ruggeri & Lombrozo, 2015; Ruggeri et al., 2016). However, our results add to a growing body of literature highlighting young children's emerging abilities to search for information efficiently across different contexts, once tested using paradigms that are sufficiently simple, clear, and child friendly (Bonawitz et al., 2012; Domberg et al., 2020; Ruggeri et al., 2017; Ruggeri et al., 2019; Ruggeri et al., 2021; Swaboda et al., 2022). Indeed, our findings provide strong evidence that even 2-year-olds are competent active learners, able to select the relevant and informative cues they need to support their decisions. This work further highlights the importance of developing age-appropriate paradigms that capture children's early competence in order to gain a more fair and comprehensive picture of their emerging information-search abilities.

Our task addressed three shortcomings of previous-research designs, which often present i) children with tasks that are not suitable for the age groups targeted (e.g., 7- to 9- year-old children ask less informative questions in a 20-question game when they have to guess professions rather than animals (Ruggeri & Feufel, 2015), ii) rely on an advanced understanding of math and probabilities (Betsch et al., 2018), or iii) ignore the fact that children may have different assumptions about the task structure and goals than what expected by the researchers (Bramley et al., 2022). In particular, we designed a task that young children would find simple, familiar, and engaging. We minimized verbal and computational task demands and made the task structure and assumptions as explicit and straightforward as possible—also making sure children had a clear understanding of the game rules and goals by the end of the familiarization phase.

At the same time, our task controlled for potential confounds (e.g., a preference for a particular color, shape, or icon; a particular sensitivity to one of the features over the others) by counterbalancing the stimuli sets, thereby ensuring that children based their inferences on the task structure, rather than on more superficial aspects of the task. We further confirmed analytically that children's performance did not differ depending on the particular set they were presented with. Moreover, the fact that 63% of children (54 out of 86) selected the correct information card in *both* rounds (with 37% of children selecting the correct card at the first attempt in both rounds), and therefore across different sets of boxes, suggests that children's information-search skills are robust and adaptive across contexts.

Our results also raise the question of whether the foundations of efficient information-search competence may emerge even earlier: Can infants differentiate between *relevant* and *irrelevant* information, and if so, how can we capture this ability experimentally?

To investigate this question, we are currently piloting an eye-tracking paradigm with 12- to 20-month-old infants, using a finding-presents game similar to that described in this study. Infants will be presented with four boxes, one on each corner of a screen, presenting different patterns or different shapes. In the middle of the screen, infants will be presented with one cue card that, when flipped, will reveal the pattern or the shape of the one box containing the present. Crucially, this cue card will be informative or uninformative, depending on whether the feature differentiating among the target boxes is the shape or the pattern. We hypothesize that infants' pupil dilation and looking time will differ between informative and uninformative trials, indicating their sensitivity to the relevance of the information provided.

References

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica*, *153*, 74–86. https://doi.org/10/f6vd8w

Betsch, T., Lehmann, A., Jekel, M., Lindow, S., & Glöckner, A. (2018). Children's application of decision strategies in a compensatory environment. *Judgment and Decision Making*, *13*(6), 514–528. https://doi.org/10/grqb3t

Bonawitz, E. B., van Schijndel, T. J., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, *64*(4), 215–234. https://doi.org/10/fxm7wp

Bramley, N. R., Jones, A., Gureckis, T. M., & Ruggeri, A. (2022). Children's failure to control variables may reflect adaptive decision-making. *Psychonomic Bulletin & Review*, *29*(6), 2314–2324. https://doi.org/10/grp9ks

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development*, *61*(5), 1584–1595. https://doi.org/10/cf6xv7

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10/dmst37

Davidson, D. (1991). Developmental differences in children's search of predecisional information. *Journal of Experimental Child Psychology*, *52*(2), 239–255. https://doi.org/10/b4jmd3

Davidson, D. (1996). The effects of decision characteristics on children's selective search of predecisional information. *Acta Psychologica*, *92*(3), 263–281. https://doi.org/10/crg9x7

De Simone, C., & Ruggeri, A. (2022). Searching for information, from infancy to adolescence. In I. C. Dezza, E. Schulz, & C. M. Wu (Eds.), *The drive for knowledge: The science of human information-seeking.* (pp. 77–100). Cambridge University Press.

Domberg, A., Koskuba, K., Rothe, A., & Ruggeri, A. (2020). Goal-adaptiveness in children's cue-based information search. *Proceedings of the Annual Conference of the Cognitive Science Society*, *42*.

Dunn, K., & Bremner, J. G. (2017). Investigating looking and social looking measures as an index of infant violation of expectation. *Developmental Science*, *20*(6). https://doi.org/10/gchx2t

Fay, A. L., & Klahr, D. (1996). Knowing about guessing and guessing about knowing: Preschoolers' understanding of indeterminacy. *Child Development*, *67*(2), 689–716. https://doi.org/10/crvw5h

Hembacher, E., deMayo, B., & Frank, M. C. (2017). Children's social referencing reflects sensitivity to graded uncertainty. *Cognitive Science*.

Herwig, J. E. (1982). Effects of age, stimuli, and category recognition factors in children's inquiry behavior. *Journal of Experimental Child Psychology*, *33*(2), 196–206. https://doi.org/10/dfd86w

Klahr, D., & Chen, Z. (2003). Overcoming the positive-capture strategy in young children: Learning about indeterminacy. *Child Development*, *74*(5), 1275–1296. https://doi.org/10/fb4z5f

Kutsuki, A., Egami, S., OGURA, T., Nakagawa, K., Kuroki, M., & Itakura, S. (2007). Developmental changes of referential looks in 7-and 9-month-olds: A transition from dyadic to proto-referential looks. *Psychologia*, *50*(4), 319–329. https://doi.org/10/dsjr5g

Lindow, S. (2021). On searching and finding: The development of information search abilities. *Cognitive Development*, *58*, 101011. https://doi.org/10/grcfkx

Mata, R., von Helversen, B., & Rieskamp, J. (2011). When easy comes hard: The development of adaptive strategy selection. *Child Development*, *82*(2), 687–700. https://doi.org/10/cp2jx4

Ruggeri, A., & Feufel, M. (2015). How basic-level objects facilitate question-asking in a categorization task. *Frontiers in Psychology*, *6*, 918. https://doi.org/10/f7jmf7

Ruggeri, A., & Katsikopoulos, K. V. (2013). Make your own kinds of cues: When children make more accurate inferences than adults. *Journal of Experimental Child Psychology*, *115*(3), 517–535. https://doi.org/10/f46hvj

Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, *143*, 203–216. https://doi.org/10/f7n6j2

Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2016). Sources of developmental change in the efficiency of information search. *Developmental Psychology*, *52*(12), 2159. https://doi.org/10/f9gbx8

Ruggeri, A., Sim, Z. L., & Xu, F. (2017). "Why is Toma late to school again?" Preschoolers identify the most informative questions. *Developmental Psychology*, *53*(9), 1620. https://doi.org/10/gcjpjg

Ruggeri, A., Swaboda, N., Sim, Z. L., & Gopnik, A. (2019). Shake it baby, but only when needed: Preschoolers adapt their exploratory strategies to the information structure of the task. *Cognition*, *193*, 104013. https://doi.org/10/grqb4j

Ruggeri, A., Walker, C. M., Lombrozo, T., & Gopnik, A. (2021). How to help young children ask better questions? *Frontiers in Psychology*, *11*. https://doi.org/10/gm8czq

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668–671. https://doi.org/10/bj743g

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*(6230), 91–94. https://doi.org/10/f66td4

Swaboda, N., Meder, B., & Ruggeri, A. (2022). Finding the (most efficient) way out of a maze is easier than asking (good) questions. *Developmental Psychology*, *58*(9), 1730–1746. https://doi.org/10/grn9jp

Vaish, A., Demir, Ö. E., & Baldwin, D. (2011). Thirteen- and 18-month-old infants recognize when they need referential information. *Social Development*, *20*(3), 431–449. https://doi.org/10/fhgsxt

Walden, T., Kim, G., McCoy, C., & Karrass, J. (2007). Do you believe in magic? Infants' social looking during violations of expectations. *Developmental Science*, *10*(5), 654–663. https://doi.org/10/dgk9wg

Chapter 4

Children Strategically Decide What to Practice

Serko, D., Leonard, J., & Ruggeri, A. (under Review).

Child Development.

Abstract

Adjusting practice to different goals and task characteristics is pivotal for learning, but it is unclear how this ability develops. Across 2 preregistered experiments, 190 children aged 4-8 years (106 female) and 31 adults played an easy and a difficult game and were informed that they would later be tested on either the easy, the difficult, or a randomly chosen game. Before the test, participants had to choose one of the two games to practice. When participants knew which game they would be tested on, participants chose to practice the soon-to-be tested game. Critically, when facing a randomly chosen test, children and adults chose to practice the difficult game, suggesting that even 4-year-olds can prepare for an uncertain future.

Introduction

Imagine that you are back in primary school, and every Friday you get either a math or a French test. On some weeks, the teacher tells you in advance which subject you will be tested on. On the weeks you know you will get a math test, you focus on studying math, and on the weeks you know you will get a French quiz, you focus on studying French. However, on some other weeks, the teacher does not tell you in advance which subject you will be tested on. How should you prepare? The answer depends on your skill set. Given an unknown upcoming test, you should strategically decide to practice the subject that is harder for you to minimize your chances of failing. If you are really good at math and confident that you will be able to do well on a math test without studying very much, you should focus on studying French. However, if you are really good at French, you might want to focus on studying math. In other words, strategic practice choices depend on both the goal (e.g., what test you need to prepare for) and the task characteristics (e.g., what task is more difficult for you).

Although a primary school Math or French quiz may seem trivial, decisions about where and how we practice build up over time to determine what we learn and who we become. This idea aligns with Ericsson's notion of deliberate practice, which posits that repeated actions aimed at improving a specific skill are crucial for achieving high performance (Ericsson et al., 1993). Practice involves the deliberate allocation of attention and efforts towards specific tasks or goals, incorporating a range of cognitive and behavioral processes. These processes include executive function, prospective memory and future thinking, metacognition and self-monitoring (Ericsson, 2008; Ericsson & Harwell, 2019; Ericsson et al., 1993; Ericsson et al., 2009). However, while for Ericsson "practice" refers to gradual improvement through the repetition of an action (Ericsson, 2008; Ericsson & Harwell, 2019; Ericsson et al., 1993; Ericsson et al., 2009), the example above specifically refers to a flavor of practice that is more strategic and active. In this paper, we focus on what we term "ecological active practice"—the strategic choices students make regarding how to invest their time and effort in preparing for the future. This concept aligns more closely with the literature on active and self-directed learning than with traditional educational views of practice. In this context, we situate our concept within the ecological active learning framework (see Ruggeri, 2022). Ecological active learning involves actively exploring and learning by recognizing and leveraging the specific structure and features of a learning task or environment. Learners adapt their exploratory and learning strategies to maximize efficiency and effectiveness, taking into account task goals, characteristics, available resources, and their own prior knowledge and abilities. Ecological active practice, in turn, refers to a strategic, personalized approach in which students make deliberate choices about how to invest their time and effort to prepare for future challenges.

Considering the critical role of ecological active practice in learning, it is surprising how little we know about its developmental trajectory. Gaining insight into its development is essential, not only for informing theoretical discussions about the relative advantages of active versus instructed learning across different stages of development (Bonawitz et al., 2011; Gureckis & Markant, 2012; Klahr & Nigam, 2004; Piaget, 1930), but also for practically understanding when and how to best scaffold children's learning to help them develop into competent and independent adults. Here, we explore whether children aged 4 to 8 are able to adapt their practice choices based on the goals (e.g., which task they will be tested on) and task characteristics (e.g., task difficulty) to maximize rewards and minimize losses.

Prior research suggests that adults adapt their practice choices appropriately based on both goals and task characteristics (Baranes et al., 2014; Kornell & Metcalfe, 2006; Locke & Latham, 2002; Metcalfe, 2011; Metcalfe & Kornell, 2005; O'Doherty et al., 2017; Ten et al., 2021). For example, Ten et al. (2021) presented participants with games that varied in difficulty and told them to play for a given number of trials. With no external constraints, adults spent their time playing easier games, on which they made fast progress (see also, Baranes et al., 2014). However, when participants were instructed to learn all games because they would eventually be tested on them, they were more likely to spend their time playing more difficult games (Ten et al., 2021). Similarly, when preparing for a test of novel word pairs, adults studied items of intermediate difficulty and avoided spending time studying items they already knew or that were very difficult (Metcalfe & Kornell, 2005). In short, adults effectively chose to practice more difficult items before a test. Given the maturity of practice decisions in the adult state, we turn our attention to when in development children possess this capacity.

Prior developmental work paints a contradictory picture of whether and in what situations children possess the cognitive capacities to strategically engage in active practice (Brinums et al., 2018; Casey & Redshaw, 2022; Cimpian, 2017; Davis et al., 2016; Flavell et al., 1970; Magid et al., 2018; Metcalfe & Finn, 2013; Wang & Bonawitz, 2022). On the one hand, research on decision-making and metacognition suggests that, unlike adults and older children, younger children's practice choices are not adaptive, in that they do not systematically take into account goals and task characteristics (Brinums et al., 2018; Brinums et al., 2021; Casey & Redshaw, 2022; Flavell et al., 1970; Metcalfe

& Finn, 2013). For example, preschoolers and third graders do not allocate more study time to a memory test item they previously did poorly on (Flavell et al., 1970; Metcalfe & Finn, 2013), and it is not until fifth grade (i.e., around 10 years of age) that children begin to behave like adults, devoting extra study time to yet-to-be-learned items before a test (Metcalfe & Finn, 2013). Furthermore, when 4- to 7-year-old children are encouraged to practice one of three games before a test, only 6-and 7-year-old children strategically practice the game that they will be tested on (Brinums et al., 2018). Overall, this line of work suggests that the ability to tailor one's practice strategies to the goals and characteristics of a given task may develop rather late in childhood, between ages 6 and 10.

On the other hand, a growing body of research with younger children on persistence, exploration, and information search provides compelling evidence that even toddlers can make adaptive learning choices in response to goals and task characteristics (Bridgers et al., 2020; Davis et al., 2016; Leonard et al., 2017; Leonard et al., 2020; Lucca et al., 2020; Magid et al., 2018; Ruggeri, 2022; Ruggeri et al., 2019; Rule et al., 2023; Wang & Bonawitz, 2022). Infants and children put more effort into a task when evidence suggests it is difficult (e.g., they observe an adult who needed to put effort into succeeding Leonard et al. (2017) and Leonard et al. (2020), Lucca et al. (2020)). Toddlers are more persistent in their search when there is more information to be gathered with their actions (Ruggeri et al., 2023) and adapt their search strategy to the task characteristics to maximize their information gain (Ruggeri et al., 2017; Ruggeri et al., 2019). Studies have shown that when the goal is to play for fun rather than to win, children aged 5 to 10 are more likely to choose a more challenging version of a game (Rule et al., 2023). Also, in simpler, more constrained and controlled paradigms, even 4-year-olds demonstrate forward-thinking by selecting objects that help them solve future tasks (Suddendorf & Moore, 2011) and by choosing to practice the game they know they will later be tested in a forced-choice task, despite being unable to explicitly why practice is important (Davis et al., 2016). This line of research suggests that, by the preschool years, children are already able to adapt their actions based on specific goals and in response to task characteristics, such as its difficulty and information structure.

One key difference between the two lines of work discussed above is the targeted age range, which influences the task designs and paradigms used. Specifically, research on metacognition and decision-making typically tests children aged 4 to 11 and often involves paradigms that rely heavily on verbal instructions and impose high memory demands. These tasks sometimes require children to follow multi-step procedures and remember specific rules (e.g., Brinums et al., 2018; Casey & Redshaw, 2022; Flavell et al., 1970; Metcalfe & Finn, 2013), which may disadvantage younger children.

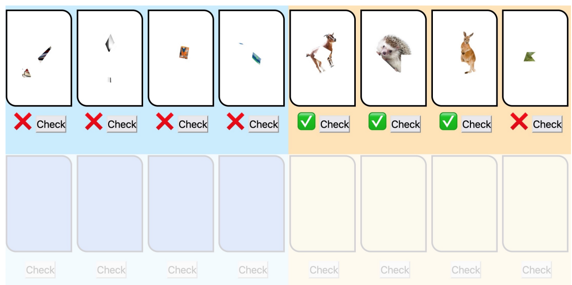
This raises the possibility that selective active practice abilities may emerge in children earlier than previously thought when presented with minimally demanding paradigms. However, beyond task demands, two key gaps remain in the literature. First, no prior work has explicitly manipulated both task characteristics (e.g., task difficulty) and goals (e.g., what to prepare for) within the same paradigm in children. As a result, it remains unclear whether and when children can use *both* factors to inform their *practice* choices. Second, no prior work has explored children's active practice choices in the face of uncertainty. This is crucial to examine, as it provides valuable insights into how children develop decision-making skills, navigate uncertain situations while minimizing potential losses, and maximize learning opportunities and efficiency.

Importantly, prior work suggests that the ability to prepare for mutually exclusive possibilities emerges at age 4 (Davis et al., 2016; Redshaw et al., 2018), suggesting that even young children can reason about uncertain situations (Coughlin et al., 2015; Ghetti et al., 2013; Goupil et al., 2016; Hembacher & Ghetti, 2014; Lyons & Ghetti, 2013; Redshaw et al., 2018; Suddendorf et al., 2017). Thus, it is possible that even preschool-age children may use their understanding of alternative future outcomes to strategically decide what to practice.

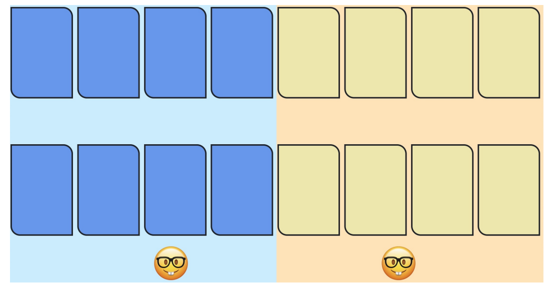
Here, we set out to answer whether children make ecological active practice choices. In particular, across two preregistered experiments (see OSF-links below), we investigated 4- to 8-year-old children's and adults' ability to tailor their active practice choices to the goals (i.e., preparing for a known or unknown test) and characteristics of a given task (i.e., the difficulty of the task). This relatively wide age range was strategically selected to bridge the two lines of research reviewed above. By testing children across this age range, we aimed to capture both the emergence and refinement of active practice, which may be influenced by developmental changes in metacognitive abilities (e.g., Fleur et al., 2021; Whitebread & Neale, 2020), executive functions (e.g., Best & Miller, 2010; Lee et al., 2014), and information-search abilities (e.g., Poli et al., 2024; Ruggeri, 2022). Additionally, we included an adult sample to provide a performance benchmark.

We specifically designed our experiments to be engaging and understandable for children across the age range targeted. Specifically, we used child friendly games (guess the picture and block building), employed memory aids, ample visual examples, and many comprehension check questions to make sure children understood our procedure (see Methods).

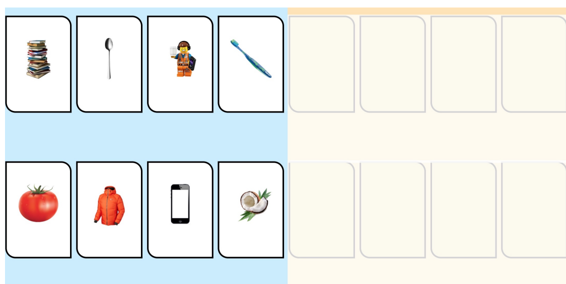
Experiment 1 implemented a within-subjects design where children (4-to 8-year-old) and adults played an online guessing game in which they had to guess pictures of familiar animals and objects that were partially occluded. Experiment 2 implemented a between-subjects design where children (4-to 5-year-old) played a hands-on, minimally-verbal brick-building task. In both studies participants were familiarized with an easy and a difficult version of a game, and were then informed about a later test, in which they would be presented with either the easy (*Test-Easy* condition), the difficult (*Test-Difficult* condition), or a randomly chosen game (*Test-Random* condition). Before entering the test, children were asked which of the two versions of the game they would like to practice.



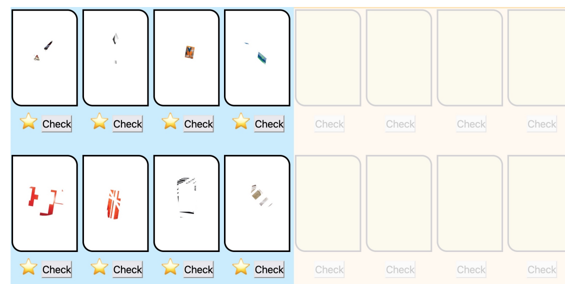
(a) Familiarization phase.



(b) Information about later test (*Test-Random* condition).



(c) Practice choice: difficult game.



(d) Test.

Figure 4.1. Screenshots from Experiment 1. (a) Children were presented with two games, where they had to identify partially-occluded pictures of animals or objects (counterbalanced across participants). They were first familiarized with the games by playing the top four cards of each game. The easy game (in this example presented to the right) included three very easy-to-guess (80% visible) pictures and one difficult-to-guess picture (20% picture visible), whereas the difficult game (in this example presented to the left; sides were counterbalanced across participants) included four very difficult-to-guess pictures. During the familiarization rounds, the feedback (green tick for a correct answer vs. red cross for an incorrect answer) was displayed under the corresponding cards. The games were *not* explicitly labeled as easy or difficult by the experimenter to avoid biasing children. After the familiarization phase, we told children that they would eventually be tested on the easy game (*Test-Easy* condition), on the difficult game (*Test-Difficult* condition), or on a randomly chosen game (*Test-Random* condition; conditions manipulated within subjects in counterbalanced order). In the *Test-Easy* and *Test-Difficult* condition, an emoji below the game they would be tested on provided a memory aid for children. (b) In the *Test-Random* condition, the emoji was placed below both games, and the experimenter emphasized that for the time being they could not know in which of the two games they would eventually be tested. We then asked children to decide which of the two games they wanted to practice before being tested. (c) Practice setup, where children practiced all the 8 images of the game they selected (the difficult game in this example). (d) Test setup, in which children were tested on all 8 images of the game they were assigned to (in this example, the difficult game; note that half children in the *Test-Random* condition were eventually tested on the easy game, whereas the other half were eventually tested in the difficult game). At test, a golden or black star below an image indicated a correct or incorrect guess. For each correct answer in the tests children received one sticker, for each incorrect answer they lost one sticker.

We predicted that participants would make ecological active practice choices, tailored to the goal structure and task difficulty in both experiments with the aim to improve performance on future tests. Specifically, we hypothesized that participants would choose to practice the easy game in the *Test-Easy* condition and the difficult game in the *Test-Difficult* condition. Our critical condition of interest was the *Test-Random* condition: we hypothesized that participants would choose to practice the difficult task in the *Test-Random* condition to minimize potential losses. Indeed, they should expect to do relatively well on the easy test even without practice, but to perform poorly in the difficult test if they had not practiced it (see the hypothesis rationale section in the SOM for more details). We tested a large age range in Experiment 1 to explore *when* children make adaptive active practice choices. In Experiment 2, we specifically targeted preschool-age children to see if even young children could make adaptive active practice choices with very minimal task demands. Sample sizes for both studies were calculated using a-priori power analysis on a simulated data set (see Methods sections and SOM). Preregistration, data, and analyses are available on the Open Science Framework (Overall project: https://tinyurl.com/4bj4dfb3; Experiment 1: https://tinyurl.com/33xptwus; Experiment 2: https://tinyurl.com/364tzkv9). Please note: Study 1 in OSF corresponds to the pilot of Experiment 1 in this paper, while Study 2 in OSF corresponds to Experiment 1 in this paper.

Experiment 1

Methods

Participants

Participants in Experiment 1 were 115 4- to 8-year-old children (64 female; M = 76.97 months; SD = 17.29 months; Range: 48 to 107 months) and 31 adults (23 female, M = 29.42 years; SD = 10.25 years; Range: 19 to 70 years). No ethnic or socio-economic status data were collected, but the population from which the sample was drawn is approximately 71% ethnic German, 11% other European, 9% Middle Eastern, 3% Asian, 2% Afro-German or Black African, and 4% other or unspecified, and encompasses a wide range of socio-economic backgrounds. We recruited participants from the database of the [blind for review] and the [blind for review]. Eight additional children were excluded from the analyses because they did not want to participate in the games ($n = 2$), they were outside our age target ($n = 3$), because of technical malfunction ($n = 1$) or missing demographic data ($n = 2$). Additionally, participants' data were excluded on rounds in which they failed to answer a comprehension check question ($n = 7$ rounds) which in some analyses leads to a small deviation in reported total trials. We also tested one additional adult, who had to be excluded from the analyses because the equipment failed to record the session.

The study was approved by the IRB of the [blind for review]. Prior to the beginning of the experimental session, adults and children's parents signed an informed consent form online, and we asked children to give verbal consent to participate. To estimate the sample size prior to data collection, we performed a power analysis on a simulated dataset (see SOM), which indicated that at least 80 children in total had to be tested to detect a difference between conditions with 80% power and 0.8 estimated effect size, with a 0.05 criterion for statistical significance. We tested more children than suggested by the power analysis to ensure an even age distribution within the sample.

Design

Experiment 1 consisted of three rounds, across which we manipulated the conditions (*Test-Easy, Test-Difficult, Test-Random*) within subjects (order of presentation counterbalanced). Each round included a familiarization, a practice, and a test phase (see Fig. 4.1). We ensured through extensive pilot testing ($N = 146$) that the instructions, materials and goals of the task were understandable for young children. Pilot data suggested that all children know the animals and objects used as stimuli. The games were designed using HTML, CSS, and JavaScript and are available online (http://xyzzz.de/paper/follow_up/openlink.html). We conducted the experiment online using the Webex conferencing tool. Webex is an online conferencing tool similar to Zoom. Webex is a secure video conferencing platform that meets the data protection standards required by the author's research organization. Children were seated next to their parents while the experimenter shared their screen. Before beginning the procedure, the experimenter confirmed via the parent that the stimuli were clearly visible to the children and that the procedure was appropriately displayed in full-screen mode. Parents were reminded not to interfere with their children's responses during the experiment. Note that a meta-analysis by Chuey et al. (2022) found that effect sizes in developmental studies conducted online were comparable to those in in-person studies, underscoring the reliability of our online testing methodology.

Familiarization phase. The experimenter introduced participants to two guessing games: an *Animal*-pictures game (8 cards) on one side of the screen, and an *Object*-pictures game (8 cards) on the other side (side counterbalanced). Participants had to guess what the first 4 pictures (top row) of each game represented, and received visual feedback on their performance (green tick for a correct answer, red cross for an incorrect answer), but were neither told the correct names for unknown items, nor given explicit feedback about the difficulty of the games (see Fig. 4.1a, SOM; see also Lyons and Ghetti (2011) who used a similar approach). All pictures showed animals and objects that participants were familiar with (e.g., a cat, a dog, a slice of pizza), but we varied the degree to which each picture was actually visible by covering part of the picture with white geometrical shapes: The easy game included three very easy-to-guess (80% visible) pictures and one difficult-to-guess picture (20% picture visible), whereas the difficult game included only very difficult-to-guess pictures.

We added one difficult-to-guess picture to the easy game, so that i) it would be necessary to practice the easy game in the *Test-Easy* condition to be able to guess all items correctly at test, and ii) children would still perceive the easy game as challenging and fun (see Serko et al., 2022). To confirm that the easy and difficult games differed in difficulty as intended, we compared children's performance in both games during familiarization (see also the detailed analysis of children's performance in each game during familiarization and at test in the SOM). As expected, in the large majority of easy-game trials children were able to correctly guess 3/4 items (87%; 294 out of 338 trials), whereas only in a few difficult-game trials children were able to correctly guess at least one of the items (4%; 15 out of 338 trials; for details see SOM). A paired t-test revealed a significant difference in performance between the easy and difficult game ($t(337) = 126.10$, $p < .001$, 95% CI [2.84, 2.93]).

Practice phase. After the familiarization phase, we provided participants with information about the test phase. In the *Test-Easy* and *Test-Difficult* conditions, we informed participants that they will eventually be tested on all eight cards of the easy or the difficult game, respectively. During this explanation of the test phase, we visually highlighted the side of the screen corresponding to the game they would be tested on by placing a smiley below the game and partially dimming the other side of the screen.

In the *Test-Random* condition, we told participants that we would test them on a randomly chosen game. To emphasize this, the experimenter showed participants two cards that matched the colors of the cards in the two games. After shuffling the cards, the experimenter explained that she would randomly draw one of the cards which would later indicate the target test game. The experimenter then randomly drew one of the cards without looking at it and without revealing it to the participant and set it aside. In order to further emphasize the uncertainty about the later test, we included a smiley icon beneath each game, highlighting the fact that the specific game for testing was not yet determined (see Fig. 4.1b).

We told child participants that they would win one sticker for each correct answer and loose one sticker for each incorrect answer. We told adult participants that the highest performers would enter a lottery over a €50 Amazon voucher. We informed participants that they could choose to practice one of the two games before the final test. Specifically, the experimenter said: "Before you take the test, you can choose one of the two games that you would like to practice again. You can practice the [color of the game] game [point to the left side], in which you were [depending on game difficulty: already pretty good vs. not very good], or you can practice the [color of the game] game [point to the right side], in which you were [depending on game difficulty: already pretty good vs. not very good]. Which game would you like to practice again before taking the test?" Once participants made their choice, they entered the practice phase in which they again guessed the four practice cards and the four new cards (8 cards in total) of the chosen game and received corrective feedback (see Fig. 4.1c).

Test phase. At test, participants had to guess all eight cards of the easy or difficult game in the *Test-Easy* and *Test-Difficult* conditions, respectively (see Fig. 4.1d). In the *Test-Random* condition, all participants were tested on the difficult game.

Results

Adults' Active Practice Choices

As predicted, adults effectively adapted their active practice choices to the task characteristics and goals: A logistic mixed-effects model predicting adults' active practice choice (easy or difficult game) by condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy* as baseline) as a fixed effect and participants' ID as a random effect revealed main effects of the *Test-Difficult* condition (p = .002, OR = 0.100 [0.024– 0.417]) and the *Test-Random* condition ($p < .001$, OR = 0.055 [0.011 – 0.266]). Exploratory follow-up analyses revealed that as expected only 26% of the adults in the *Test-Easy* condition (n = 8/ 31) selected to practice the difficult game, which is significantly lower than chance (one-tailed 50% binomial test, p = .005). In the *Test-Difficult* and *Test-Random* condition 71% (n = 22/31) and 81% (n = 25/31) of the adult participants, respectively, chose to practice the difficult game. Both proportions were significantly greater than chance (one-tailed 50% binomial test,*Test-Difficult* condition: p = .015; *Test-Random* condition: $p < .001$). A chi-square test revealed no difference in active practice choice between the *Test-Difficult* and *Test-Random* conditions ($\chi^2(1) = 0.352$, p = .553). Thus, as predicted, we found that adults chose to practice the difficult games more often in the *Test-Difficult* and *Test-Random* conditions compared to the *Test-Easy* condition.

Children's Active Practice Choices

Overall, we found that children—like adults—effectively adapted their active practice choices to fit the goals and task characteristics. A logistic mixed-effects model predicting children's active practice choices (easy or difficult) with condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy* as baseline) and age in months as fixed effects and participants' ID as a random effect revealed main effects of the *Test-Difficult* condition ($p < .001$, $OR = 0.197$ [0.102 – 0.380]) the *Test-Random* condition ($p < .001$, $OR = 0.226$ [0.119 – 0.430]). Exploratory follow-up analyses revealed that as expected in the *Test-Easy* condition, 39% of the children ($n = 44/113$) selected the difficult game (significantly lower than chance, $p = .012$, one-tailed 50% binomial test). As expected, in the *Test-Difficult* condition and *Test-Random* condition 71% ($n = 79/111$) and 68% ($n = 78/114$) of the children, respectively, selected the difficult game. Both proportions were significantly greater than chance (one-tailed 50% binomial test,*Test-Difficult* condition: $p < .001$; *Test-Random* condition: $p < .001$). A chi-square test revealed no difference in task choice between the *Test-Difficult* and *Test-Random* conditions ($\chi^2(1) = 0.092$, $p = .761$) indicating that in both conditions children made similar active practice choices. The analysis also revealed a main effect of age in months ($p < .001$, $OR = 0.966$ [0.949 – 0.983]), indicating that older children were more likely to choose to practice the difficult game.

To further explore age effects, we conducted an exploratory logistic mixed-effects model, predicting children's active practice choices by condition, age in months, and their interaction as fixed effects, and participants' ID as a random effect. The analysis revealed a significant interaction of the *Test-Difficult* condition and age ($p < .001$, $OR = 0.204$ [0.092 – 0.453]), and the *Test-Random* condition and age ($p = .013$, $OR = 0.440$ [0.231 – 0.838] see Fig. 4.2 and Fig. 4.3), indicating that older children were more likely to practice the difficult game in these conditions. Note that this interaction model fits the data better (AIC: 397.0) compared to the main effects model presented above (and preregistered; AIC: 412.7), without the interaction. In an additional model, we added round as a main effect predictor to investigate whether children's performance improved over the course of the experiment (i.e., as they played more rounds). However, adding round as predictor did not improve the model fit (AIC: 414.6) and round did not contribute to predicting children's practice choices.

Finally, we performed an exploratory analysis to identify the age at which children begin making ecological active practice choices. Specifically, we compared the number of children within each age group, binned by years, who made adaptive active practice choice with the number of children who did not, using a one-sample proportions test. We found that among 4- (54.93%; 39 out of 71 trials) and 5-year-old (52.86%; 37 out of 70 trials) children performance was not significantly different from chance (two-tailed 50% binomial test, p = .202 for 4-year-old children; p = .360 for 5-year-old children). By age 6, the majority of children (66%; $n = 43$ out of $n = 65$ trials) were able to make adaptive practice choices (two-tailed 50% binomial test, p = .013; see Fig. 4.2 and SOM for a complete breakdown of adaptive practice choices across age groups).

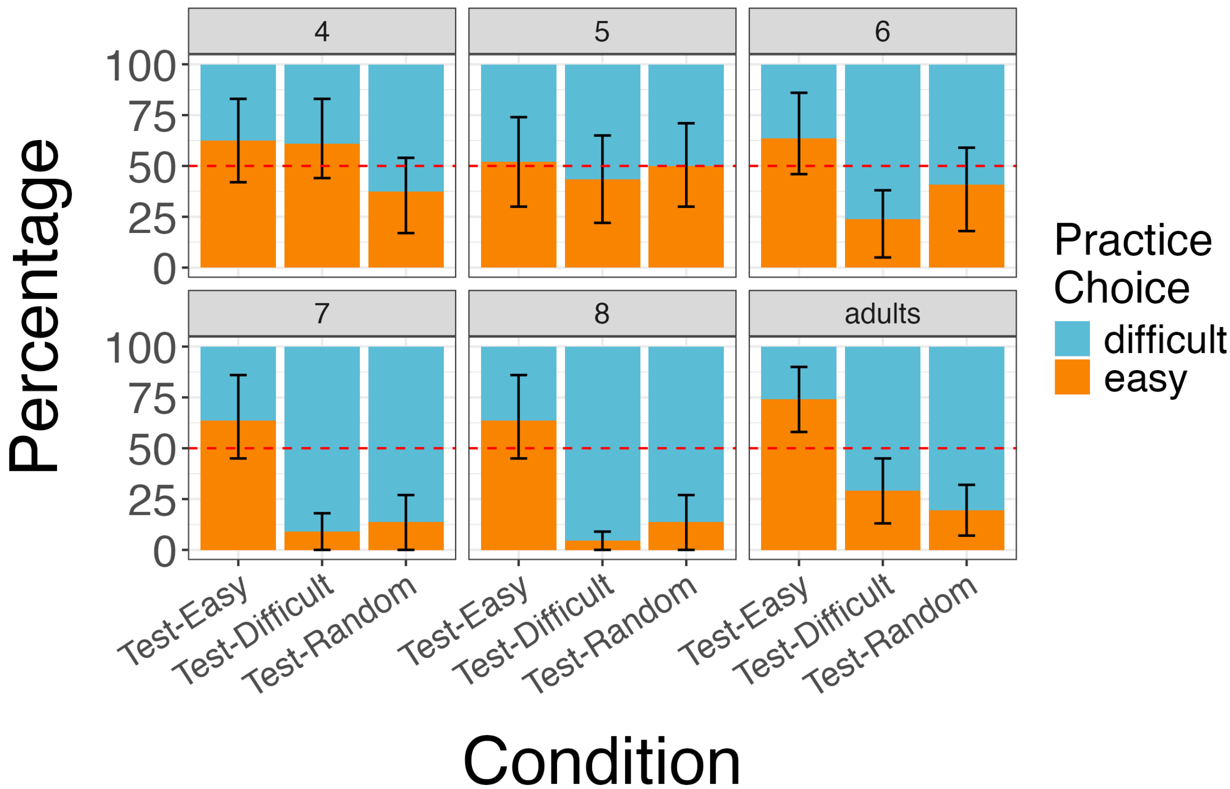


Figure 4.2. Illustration of participants' active practice choices by condition (*Test-Easy, Test-Difficult, and Test-Random*) in Experiment 1 faceted by age in years (adults as a separate group). Bars represent the percentage of easy (blue) and difficult (orange) active practice choices for each condition. Error bars indicate 95% confidence intervals.

Interim summary

The findings from Experiment 1 indicate that both adults and children aged 6 and above make active practice choices that adapt to the goals and characteristics of a given task: They choose to practice the task that they will be eventually tested on, and importantly, when they don't know which task they will be tested on, choose to study the more difficult task to make up for their potential losses. For the 4- and 5-year-old children, performance did not significantly differ from chance, leaving it unclear whether their active practice choices were intentional or merely random (see Fig. 4.2 and SOM).

There are several reasons why younger children may have had difficulties adapting their active practice choices to the task goal and characteristics.

First, although we aimed to minimize linguistic and cognitive demands, we may not have reduced them sufficiently to accommodate the needs of younger children. The game relied on remembering the names of the objects learned during the practice round, which may have been much more challenging for younger children, whose memory is still developing (Gathercole, 1998; McCormack & Atance, 2011). Indeed, additional analyses of children's performance indicate that younger children, compared to older children, had difficulties with guessing the difficult items at test even when they had practiced them, indicating that these items were difficult to remember for them (see SOM for details). The study design also required children to understand the concept of "randomness." In Experiment 1, we illustrated randomness by shuffling in front of the children two cards, each representing the easy or the difficult game. The experimenter then told children that, at test, they will pick one of the cards and find out which game children will be tested on (see Methods). Despite this, it is possible that children may have failed to understand that they had a 50% chance of being tested in either the easy or the difficult game. Furthermore, we did not implement a comprehension check assessing whether children understood the given goal of the task (see Locke & Latham, 2002). Thus, it could be the case that younger children chose what to practice based on their own goals (e.g., perhaps to maximize their fun), rather than the given task goals. Finally, the experimental session in Experiment 1 took about 20 minutes to complete, which may have been too long for young children.

Experiment 2

Experiment 2 aimed to test whether 4- and 5-year-old children can make adaptive active practice choices when presented with a less demanding version of the paradigm used in Experiment 1. To address the limitations mentioned above, we made several changes to the paradigm. First, we made the task more child-friendly and less cognitively demanding by using a brick-building task that didn't rely on verbal memory and allowed children to indicate their choices by pointing. We also conducted the tests in person to ensure maximum engagement. Additionally, we provided explicit feedback regarding the difficulty level of each game, labeling them as "easy" or "difficult" after children had familiarized themselves with each game. Furthermore, we explicitly demonstrated the concept

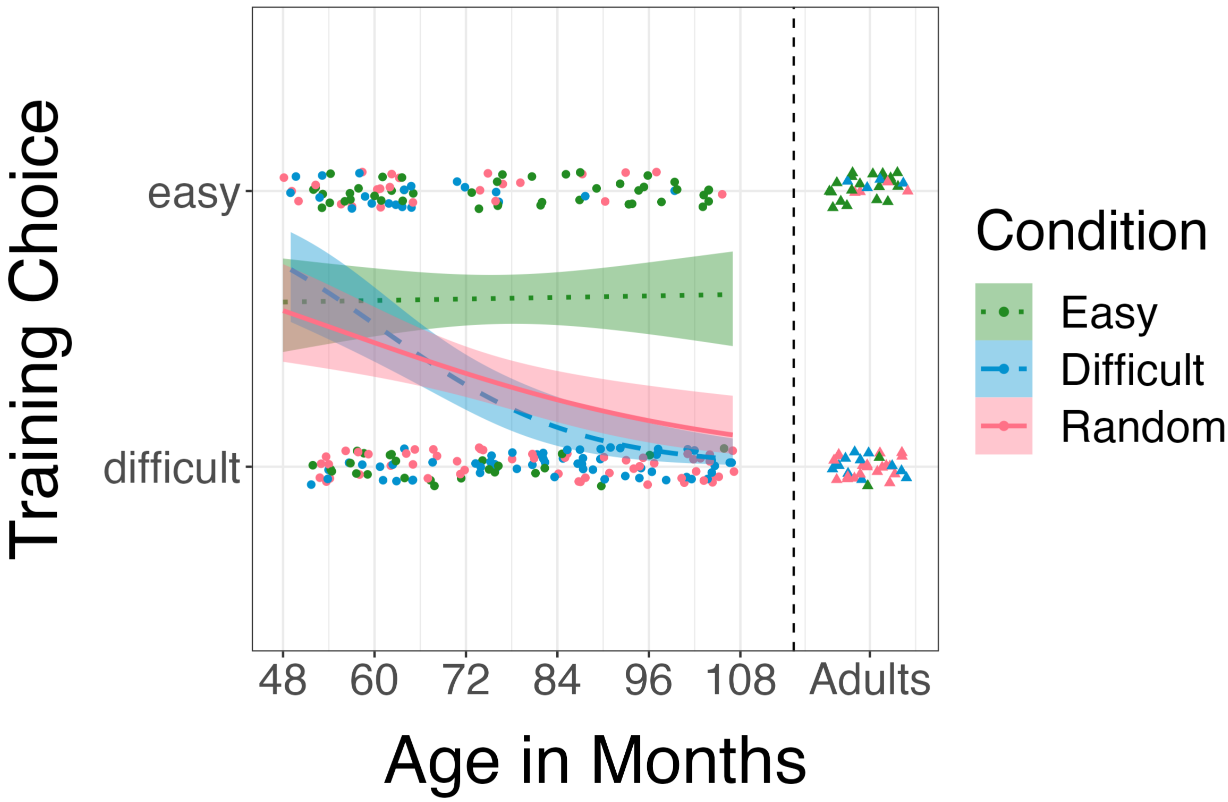


Figure 4.3. Illustration of participants' active practice choices, displayed by age (in months, with adults as a separated group on the right) and condition. Dots indicate individual active practice choices in a single game. The three lines are fitted logistic regressions with a 95% confidence interval by condition.

of randomness to young children by blindly picking between a box with easy-to-stack or difficult-to-stack blocks and confirmed that children understood this manipulation (see Design for details). We also implemented a series of comprehension checks to make sure children fully understood the goal and the characteristics of their assigned condition. We predicted that with these extra scaffolds, even 4- and 5-year-old children would make adaptive active practice choices. Finally, we changed our paradigm to be between-subjects, which reduced the average length of the experimental session from 20 minutes to 6 minutes, and reduced cognitive demands by only presenting children with one condition.

Methods

Participants

Participants were 75 (n = 25 per condition) 4- to 5-year-old children (42 female, M = 59.72 months; SD = 6.79 months; Range: 48 to 71 months). We recruited and tested participants in the public Zoo of [blind for review]. No ethnic or socio-economic status data were collected, but the population from which we obtained the sample was the same as in Experiment 1. We recruited 36 additional children (22 female, M = 56.09 months; SD = 7.87 months; Range: 37 to 71 months), but excluded them from further analysis due to preregistered exclusion criteria: failure to answer the comprehension questions correctly (n = 13), difficulties or failure to understand the instructions or to perform the brick building task (n = 9), experimenter error (n = 4), because they were too shy to interact with the experimenter (n = 3), they were outside our age range (n = 3), because of technical failure (n = 3), or because they watched another child's active practice-choice making (n = 1).

We acknowledge a high exclusion rate in our study, which we attribute to several factors. Beyond the typical challenges of working with young children—such as shyness, distraction, or inattentiveness—two additional factors likely contributed. First, Experiment 2 was conducted at the zoo, providing strong ecological validity by placing children in an engaging, educational environment. However, the lively setting, with exotic animals and other visitors, also made it harder for children to concentrate on the tasks and follow instructions. As a result, the higher exclusion rate for failing comprehension checks and struggling with the block task may stem from the distractions inherent to this stimulating environment.

Second, we implemented different comprehension-check questions to ensure that all included children understood the task instructions. If a participant failed a comprehension check, the experimenter corrected them and repeated the question. Participants were excluded from the analysis if they failed to answer correctly after three attempts. Specifically, before beginning the main task, we asked children to: 1) identify which of the two games was easier and which was more difficult for them, 2) confirm whether they knew the blocks they would use during the test (*Test-Easy* and *Test-Difficult* conditions) or did not know them (*Test-Random* condition), 3) confirm that the goal was to build the tallest tower to win stickers, and 4) verify that they understood they had a limited number of attempts during the test phase.

Consequently, we excluded 6 participants for incorrectly identifying the game difficulty and 7 participants for failing to accurately indicate their knowledge about the game they would later be tested on. While this exclusion process was stringent, we believe it was essential to maintain the integrity of our results, ensuring a more accurate understanding of the developmental trajectory in children's active practice choices. Also note that only 13 participants were excluded due to failing a comprehension check questions, while the remaining 23 exclusions were due to factors such as difficulty building the towers during familiarization, shyness in interacting with the experimenter, age ineligibility, experimenter error, or observing another child's practice choice.

The sample size was determined by conducting a simulation-based a-priori power calculation to detect the hypothesized effect with 80% power and a 0.4 difference between the *Test-Easy* condition on the one side and the *Test-Random* condition on the other side, with a 0.05 criterion for statistical significance. The difference of 0.4 between conditions was based on the results of Experiment 1. This sample-size calculation indicated a sample of 50. Given that we have a third condition (*Test-Difficult* condition), we tested 25 additional children, for a total of 75 children.

Before beginning the experimental session, parents signed an informed consent, and children were asked to give verbal consent to participate. The study was approved by the ethics committee of the [blind for review].

Design

Children played two building-block games: In the easy game they had to build a tower with three easy-to-staple cuboid blocks, whereas in the difficult game they had to build a tower of three oddly-shaped, difficult-to-staple blocks (color and order were counterbalanced across participants; see Fig. 4.4; see scripts in the SOM). After a set of comprehension checks (see familiarization phase below), we informed children that we would eventually test them on the easy game (*Test-Easy* condition), the difficult game (*Test-Difficult* condition), or a randomly chosen game (*Test-Random* condition), and that they could win stickers depending on the height of the tower they built: one sticker for a three-blocks tower, two stickers for a six-blocks tower. However, if the tower collapsed, children would not win anything. Children could then choose which of the two games they wanted to practice before entering the test phase.

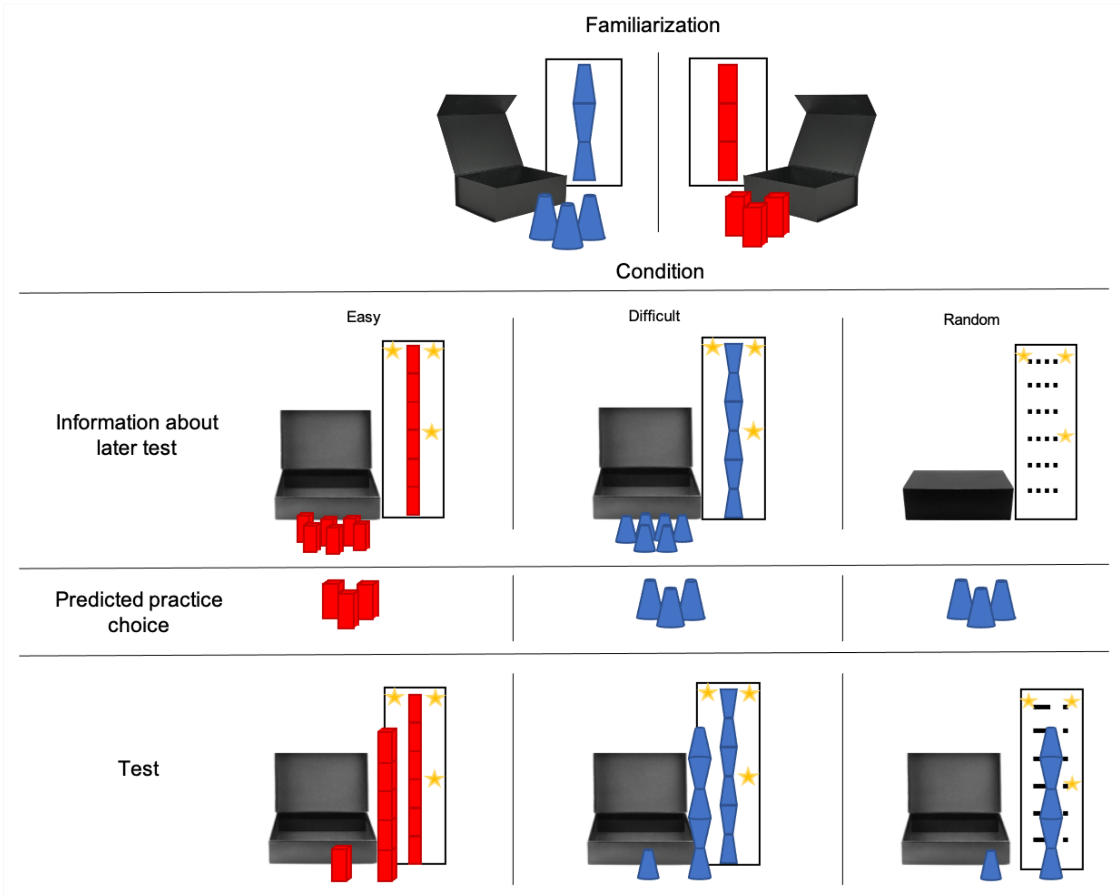


Figure 4.4. Illustration of the procedure of Experiment 2. Children were presented with two games, in which they had to construct towers using blocks. They first familiarized themselves with each game. The easy game (in this example presented on the right) featured three easy-to-assemble cuboid blocks, while the difficult game (in this example presented on the left; sides and colors were counterbalanced across participants) included three oddly shaped and difficult-to-build blocks. The order and color of the two games were counterbalanced across participants. After the familiarization phase, we told children that they would eventually be tested on the easy game (*Test-Easy* condition), on the difficult game (*Test-Difficult* condition), or on a randomly chosen game (*Test-Random* condition), and that they could win stickers depending on the height of the tower they built. Children were told that if they managed to build a tower out of three blocks at test, they would win one sticker; If they managed to build a tower out of six blocks they would win two stickers (indicated with stars). If the tower collapsed, children would not win anything. We then asked children to decide which of the two games they wanted to practice before being tested. At test, children in the *Test-Easy* condition were given 6 easy blocks. Children in the *Test-Difficult* condition were given 6 difficult blocks. In the *Test-Random* condition children were given a randomly chosen box containing 6 blocks of either the easy or difficult to stack blocks. A golden star on the level of three stacked blocks indicated that children won 1 sticker, two golden stars on the level of 6 stacked blocks indicated that children won 2 stickers.

We ensured through extensive pilot testing (N = 125) that our instructions were clear, and that young children could perform the tasks. All materials used in the experiment were novel and specifically built for the experiment. To ensure that building a tower with regular blocks was indeed easier than with oddly-shaped blocks, we measured how long it took children to build a tower with each set of blocks. To avoid a bias towards larger objects, the blocks used in both games were built to have the exact same height (63mm; for details see SOM).

Familiarization phase. Children sat on a blanket next to the experimenter in front of two covered identical boxes. The experimenter uncovered one of the boxes and introduced the children to three identical blocks and a picture of a three-blocks tower built out of the three blocks which children had to reproduce using the given blocks. The experimenter then repeated the same procedure with the other box. In the easy game, the blocks were easy-to-stack regular cuboid blocks; in the difficult game, the blocks were difficult-to-stack oddly-shaped blocks (see Fig. 4.4 and SOM).

After children successfully built each tower, the experimenter gave explicit feedback on the difficulty of the two games ("wow that was easy" or "wow that was difficult") and asked whether they believed they could get better with practice at that particular game. Irrespective of their response, the experimenter always said that children could get better at each game with practice. After this, we asked children to indicate which game was easy and which one was difficult for them. If they failed to answer correctly (n = 14), we asked them to rebuild both towers and repeated the comprehension check question. If they again failed to answer correctly (n = 6), they were excluded from the study. Children faced no time restriction to build each tower and we included only children in the sample who managed to build both towers (n = 2; excluded for failing to build the difficult tower). The order of presentation of the games (easy or difficult), as well as the color of the blocks (blue or red) was counterbalanced. As expected, during familiarization, children needed significantly less time to build the three brick tower with the easy blocks (M = 8.47 seconds, SD = 4.38) compared to building the difficult tower (M = 20.58 seconds, SD = 16.58). A Welch two-sample t-test revealed a significant difference in performance between the easy and difficult game ($t(81.985) = -6.036$, $p < .001$).

Practice phase. Children were randomly assigned to one of three conditions: *Test-Easy, Test-Difficult, and Test-Random.* In the *Test-Easy* and the *Test-Difficult* conditions, children were presented with an open box (covered until then) containing six blocks of the same color and shape as those presented in the easy or difficult game, respectively (see Information about later test in Fig. 4.4). Next to the box, there was a picture illustrating a tower made out of six of those blocks: at the three-blocks height there was one star and at the six-blocks height there two stars indicating that children would win 1 or 2 stickers depending on their performance. We informed children that, at test, they would have to build a tower out of the blocks in that box to win stickers: they would win one sticker if they managed to build a three-blocks tower, two stickers for a six-blocks tower.

In the *Test-Random* condition, children were presented with two *closed* identical boxes (covered until then). The experimenter opened the two boxes, one by one, and showed that they contained either the easy-to-staple or the difficult-to-staple blocks (order counterbalanced). The experimenter then closed both boxes, began shuffling them behind a curtain with their eyes closed and explained children that at the count of three she would pick one random box with her eyes closed, which she did. Pilot testing indicated that when the experimenter kept her eyes closed during the procedure (as opposed to keeping them open), children clearly understood that she was selecting a box at random, rather than intentionally choosing a specific box for them to be tested on. The closed box was placed next to the illustration showing the required tower height for children to construct during the testing phase (see Fig. 4.4). Children in all conditions were told that at test they would have only one shot at building the tower, and that if the tower collapsed they would not win anything.

Before entering the test phase, we asked children which one of the two games they wanted to practice in order to prepare for the test. This choice was our main dependent variable. Children then practiced with the three blocks of the chosen game until they managed to build a three-blocks tower (see Fig. 4.4).

Test phase. At test, children were given the box with the test blocks and were asked to build a six-blocks tower. The results indicated that children who encountered the easier blocks, on average, built significantly taller towers (M = 5.83 blocks, SD = 0.65) compared to those who were presented with the difficult blocks (M = 4.43 blocks, SD = 1.34; $t(29.832) = 4.602$, $p < 0.001$, Welch two-sample t-test).

Results

Practice choices

As predicted, 4- to 5-year-old children effectively adapted their active practice choices to the task characteristics and goals (see Fig. 4.5). A logistic regression model predicting children's active practice choices (easy or difficult) with condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy* as baseline) revealed main effects of the *Test-Difficult* condition ($p < .001$, OR = 0.097, [0.024 – 0.339]) and the *Test-Random* condition ($p < .001$, OR = 0.118, [0.030 – 0.403]). An exploratory analysis adding age in months to the model, revealed that the reported main effects of conditions hold, and indicated no significant effect of age. In the *Test-Easy* condition only 20% of the children (n=5/25) selected the difficult game, which was significantly lower than chance (p = .002, one-tailed 50% binomial test). In the *Test-Difficult* and *Test-Random* conditions , 72% (n=18/25) and 68% (n=17/25) of the children, respectively, selected the difficult game. The proportion of the *Test-Difficult* (p = .022) condition was significantly greater than chance and the *Test-Random* (p = .054) condition was trending to be significantly greater than chance. Also note that the CIs of the *Test-Random* condition do not touch the 50% chance rate which can be interpreted as further support for a significant effect (see Fig. 4.5). A chi-square test revealed no difference in task choice between the *Test-Difficult* and *Test-Random* conditions ($\chi^2(1)$ = 0.000, p = 1.000). Thus, even preschool-age children can adapt their training choices based on what they will eventually be tested on to maximize rewards and minimize their losses on tasks with minimal demands.

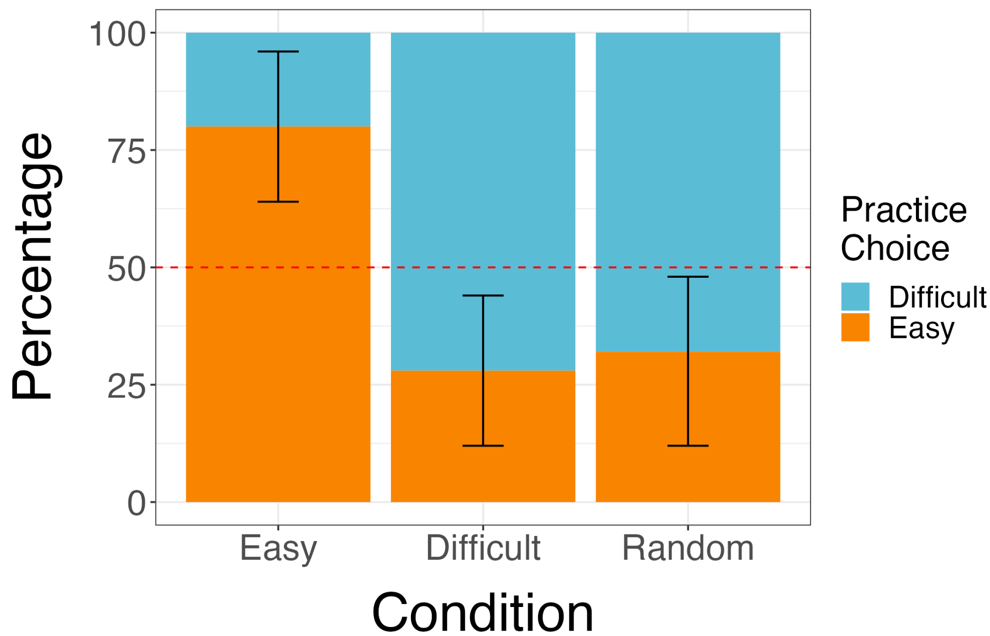


Figure 4.5. Illustration of participants' active practice choices by condition in Experiment 2. Each bar represents one of three conditions: *Test-Easy, Test-Difficult*, or *Test-Random.* The y-axis shows the percentage of participants' choices between the easy and difficult games. The blue color indicates that participants chose to practice the difficult game, while the orange color indicates participants chose to practice the easy game. The red dotted line indicates the 50% chance level. The error bars extending above and below the bars represent the 95% confidence intervals.

Discussion

We find that 4- to 8-year-old children are capable of making efficient ecological active practice choices. When aware of the task they will be tested on, they selectively practice that task. Critically, when uncertain about the task they will face, they choose to practice the more difficult one. This suggests that, like adults, even preschoolers adapt their active practice strategies when faced with uncertainty, focusing on tasks they are less proficient in rather than those they already excel at.

In Experiment 1, we found evidence of developmental changes in children's ecological active practice choices. In particular, children did not begin to selectively practice the game they would soon be tested on until age 6. Moreover, it was not until age 7 that children behaved like adults, choosing to practice the more difficult game when unsure which game they would be tested on. This finding is consistent with research on the developmental trajectory of question asking by Ruggeri and Lombrozo (2015) and Ruggeri et al. (2017), Ruggeri et al. (2019), which indicates that by ages 7 and 10, children adapt their strategies as readily as adults, even though their baseline performance differs.

However, in Experiment 2, we found that even preschool-age children can adapt their active practice choices to different conditions when presented with a simplified task and provided with adequate scaffolding. This is particular noteworthy, as previous work was inconclusive about whether young children have the cognitive capacities to adapt their active practice choices to given task goals and characteristics (Brinums et al., 2018; Cimpian, 2017; Magid et al., 2018; Metcalfe & Finn, 2013; Wang & Bonawitz, 2022).

Why were preschool-age children able to engage in effective active practice in Experiment 2, but not in Experiment 1? One possibility is that Experiment 2 had lower task demands compared to Experiment 1. Specifically, Experiment 2 was substantially shorter and less complicated than Experiment 1 because it employed a between-subjects design with a child-friendly block-building task. Additional, Experiment 2 was conducted in person (unlike Experiment 1, which was online), with explicit labeling of game difficulty and a clear demonstration of randomness—where the experimenter selected a box blindfolded behind a curtain. We also ensured comprehension by including only children who passed a series of checks. Any of these factors could have contributed to the children's success. Future research is needed to pinpoint which elements are most critical in helping young children adapt their practice choices effectively.

More broadly, our research adds to the ongoing discourse about whether and how self-directed, active learning enhances learning depth and quality when contrasted with instructed learning (Bruner et al., 1976; Kuhn, 2000; Montessori, 1912/1964; Piaget, 1930). In particular, recent work shows that allowing children and adults to actively control their learning experience (e.g., decide what to study, in what order and for how long) improves their memory of the learned materials (Chi, 2009; Gureckis & Markant, 2012; Markant & Gureckis, 2014; Markant et al., 2016; Ruggeri et al., 2019).

While our findings indicate that even young children can make active learning choices that prepare them for the uncertainty of the future, we acknowledge that our study does not investigate the efficiency of these decisions compared to instructed learning decisions in terms of their impact on learning outcomes. Future research should directly compare children's learning outcomes in active learning settings with those in instructed learning environments to draw more explicit conclusions about the effects of active versus instructed learning on children's learning processes.

Our findings also contribute to debates concerning young children's *wishful thinking* (Bernard et al., 2016; Lipko et al., 2009; Schneider, 1998; Wente et al., 2019). For example, Wente et al. (2019) demonstrated that 3-to 5-year-old children were likely to overestimate the occurrence of a low-probability event (like drawing a rare card) when they were promised a reward upon the event's occurrence, compared to when the reward was guaranteed regardless. This finding suggests that young children's predictions are often biased by their desires. However, if the participants in our study had been primarily motivated by wishful thinking, they should have expected in the *Test-Random* condition to be tested in the easy game (which was way less likely to result in a disappointing failure, and offered the perspective of winning more stickers), and therefore should have practiced the easy game. However, we find that the large majority of children in the *Test-Random* condition chose to practice the difficult game. Thus, our results suggest that children do not always let their desires control their beliefs about the future, and therefore bias their actions aimed to prepare for an unknown future state. This is interesting also in light of recent evidence suggesting that preschoolers struggle to align their actions to the probability of future events (Crimston et al., 2023).

While our results offer promising insights and potential for future research avenues, we also recognize a few limitations of our Experiments. First of all, our participant pool was predominantly comprised of European children. This might reduce the generalizability and therefore the broader applicability of our findings to other cultural contexts. Future studies should prioritize the inclusion of a more diverse sample and control for variability in socioeconomic status (SES) and parental education. Second, the current results do not allow us to monitor individual differences leading to performance variation. Future work should identify specific individual factors, including cognitive competencies which may mediate children's promptness and competence to make adaptive active practice choices. Third, our experimental design was relatively narrow, centered around a single task (either guessing item names or building a tower) presented at two varying difficulty levels (easy and difficult versions of a game). However, the realm of active practice choices encompasses a much wider spectrum of complexity. A promising avenue for future research would explore children's active practice choices within and across different domains and types of tasks —for instance, deciding to practice a physical task versus delving into a cognitive task. Fourth, we used comprehension check questions to ensure that all children understood the task. These checks were designed to confirm a basic understanding, not to select only the most attentive or advanced participants. Future research could explore varying the complexity of these checks to better assess how different levels of attentiveness or cognitive ability impact children's active practice choices. Fifth, we provided explicit feedback to participants, in order to control their inferences about task difficulty. Future work should explore both how children represent and reason about their own abilities in the absence of performance feedback, and whether they can use these internal assessments to prepare for future unknowns.

The decisions individuals make about what to practice and learn, accumulate over the course of their lives. Eventually, these active practice choices end up shaping people's competencies and expertise, their likelihood "*to be well prepared*" for whatever may come—in the words of a 4-year-old girl explaining why she chose to practice the difficult game in the *Test-Random* condition in Experiment 2. Our work shows that the ability to effectively adapt ecological active practice choices to specific goals and task characteristics emerges already by 4 years of age. In this sense, young children seem to be equipped with the tools they need to prepare for the unknown.

References

Baranes, A. F., Oudeyer, P.-Y., & Gottlieb, J. (2014). The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration. *Frontiers in Neuroscience*, *8*, 317. https://doi.org/10/gfkpgg

Bernard, S., Castelain, T., Mercier, H., Kaufmann, L., Van der Henst, J.-B., & Clément, F. (2016). The boss is always right: Preschoolers endorse the testimony of a dominant over that of a subordinate. *Journal of Experimental Child Psychology*, *152*, 307–317. https://doi.org/10/ghn6hn

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child development*, *81*(6), 1641–1660. https://doi.org/fmbqgm

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322–330. https://doi.org/10/ds2mqz

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. https://doi.org/10/gq92d5

Brinums, M., Imuta, K., & Suddendorf, T. (2018). Practicing for the future: Deliberate practice in early childhood. *Child Development*, *89*(6), 2051–2058. https://doi.org/10/gfj2wh

Brinums, M., Redshaw, J., Nielsen, M., Suddendorf, T., & Imuta, K. (2021). Young children's capacity to seek information in preparation for a future event. *Cognitive Development*, *58*. https://doi.org/10/gtphnh

Bruner, J. S., Jolly, A., & Sylva, K. (1976). *Play: Its role in development and evolution*. Penguin Harmondsworth.

Casey, C. G., & Redshaw, J. (2022). Developmental changes in children's choices of practice strategies. *Learning and Instruction*, *80*. https://doi.org/10/gtphnk

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105. https://doi.org/10/dcbxk3

Chuey, A., Boyce, V., Cao, A., & Frank, M. C. (2022). Conducting developmental research online vs. in-person: A meta-analysis. https://doi.org/10/n3nd

Cimpian, A. (2017). Early reasoning about competence is not irrationally optimistic, nor does it stem from inadequate cognitive representations. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation* (2nd ed.). Guilford Press.

Coughlin, C., Hembacher, E., Lyons, K. E., & Ghetti, S. (2015). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science*, *18*(6), 957–971. https://doi.org/10/f7vdbp

Crimston, J., Redshaw, J., & Suddendorf, T. (2023). What are the odds? Preschoolers' ability to distinguish between possible, impossible, and probabilistically distinct future outcomes. *Developmental Psychology*, *59*(10), 1881–1891. https://doi.org/10/gtphnm

Davis, J. T. M., Cullen, E., & Suddendorf, T. (2016). Understanding deliberate practice in preschool-aged children. *Quarterly Journal of Experimental Psychology*, *69*(2), 361–380. https://doi.org/10/gh7x9t

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: A general overview. *Academic Emergency Medicine*, *15*(11), 988–994. https://doi.org/10/d2v2n7

Ericsson, K. A., & Harwell, K. W. (2019). Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future research. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02396

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. https://doi.org/10/d7gc4g

Ericsson, K. A., Nandagopal, K., & Roring, R. W. (2009). Toward a science of exceptional achievement. *Annals of the New York Academy of Sciences*, *1172*(1), 199–217. https://doi.org/10/cxdpk8

Flavell, J. H., Friedrichs, A. G., & Hoyt, J. D. (1970). Developmental changes in memorization processes. *Cognitive Psychology*, *1*(4), 324–340. https://doi.org/10/dt7tpf

Fleur, D. S., Bredeweg, B., & van den Bos, W. (2021). Metacognition: Ideas and insights from neuro- and educational sciences. *npj. Science of Learning*, *6*(1), 1–11. https://doi.org/10/gkgk4r

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(1), 3–27. https://doi.org/10/c6cdwz

Ghetti, S., Hembacher, E., & Coughlin, C. A. (2013). Feeling uncertain and acting on it during the preschool years: A metacognitive approach. *Child Development Perspectives*, *7*(3), 160–165. https://doi.org/10/f47t2m

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. https://doi.org/10/f8gq87

Gureckis, T. M., & Markant, D. B. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, *7*(5), 464–481. https://doi.org/10/f399xq

Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science*, *25*(9), 1768–1776. https://doi.org/10/f6ks7j

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661–667. https://doi.org/10/bst2b4

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology-Learning Memory and Cognition*, *32*(3), 609–622. https://doi.org/10/btm7x5

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*(5), 178–181. https://doi.org/10/cnp5fp

Lee, M. C. Y., Chow, J. Y., Komar, J., Tan, C. W. K., & Button, C. (2014). Nonlinear pedagogy: An effective approach to cater for individual differences in learning a sports skill. *PLoS ONE*, *9*(8). https://doi.org/10/f6spqz

Leonard, J. A., Lee, Y., & Schulz, L. E. (2017). Infants make more attempts to achieve a goal when they see adults persist. *Science*, *357*(6357), 1290–1294. https://doi.org/10/gb2cwh

Leonard, J. A., Sandler, J., Nerenberg, A., Rubio, A., Schulz, L. E., & Mackey, A. P. (2020). Preschoolers are sensitive to their performance over time. *Proceedings of the Annual Conference of the Cognitive Science Society*, *42*.

Lipko, A. R., Dunlosky, J., & Merriman, W. E. (2009). Persistent overconfidence despite practice: The role of task experience in preschoolers' recall predictions. *Journal of Experimental Child Psychology*, *103*(2), 152–166. https://doi.org/10/bzt5vw

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation - a 35-year odyssey. *American Psychologist*, *57*(9), 705–717. https://doi.org/10/b7spwn

Lucca, K., Horton, R., & Sommerville, J. A. (2020). Infants rationally decide when and how to deploy effort. *Nature Human Behaviour*, *4*(4), 372–379. https://doi.org/10/gq6jbr

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*(6), 1778–1787. https://doi.org/10/bkn4tg

Lyons, K. E., & Ghetti, S. (2013). I don't want to pick! introspection on uncertainty supports early strategic behavior. *Child Development*, *84*(2), 726–736. https://doi.org/10/f4s8dx

Magid, R. W., DePascale, M., & Schulz, L. E. (2018). Four-and 5-year-olds infer differences in relative ability and appropriately allocate roles to achieve cooperative, competitive, and prosocial goals. *Open Mind*, *2*(2), 72–85. https://doi.org/10/grqb4h

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94. https://doi.org/10/f5s4jq

Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education*, *10*(3), 142–152. https://doi.org/10/f833ck

McCormack, T., & Atance, C. M. (2011). Planning in young children: A review and synthesis. *Developmental Review*, *31*(1), 1–31. https://doi.org/10/frncxc

Metcalfe, J. (2011). Metacognitively guided study in the region of proximal learning. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Artificial intelligence in education* (pp. 1–2). Springer.

Metcalfe, J., & Finn, B. (2013). Metacognition and control of study choice in children. *Metacognition and Learning*, *8*(1), 19–46. https://doi.org/10/f4qkjg

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*(4), 463–477. https://doi.org/10/bstxnf

Montessori, M. (1912/1964). *The montessori method*. Schocken.

O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making (S. T. Fiske, Ed.). *Annual Review of Psychology*, *68*, 73–100. https://doi.org/10/gf2stt

Piaget, J. (1930). *The child's conception of causality* (Vol. 5). Routledge.

Poli, F., Li, Y. L., Naidu, P., Mars, R. B., Hunnius, S., & Ruggeri, A. (2024). Toddlers strategically adapt their information search. *Nature Communications*, *15*(1). https://doi.org/10/n3nb

Redshaw, J., Leamy, T., Pincus, P., & Suddendorf, T. (2018). Young children's capacity to imagine and prepare for certain and uncertain future outcomes. *PLoS ONE*, *13*(9). https://doi.org/10/gd6stg

Ruggeri, A. (2022). An introduction to ecological active learning. *Current Directions in Psychological Science*. https://doi.org/10/grbccz

Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition*, *143*, 203–216. https://doi.org/10/f7n6j2

Ruggeri, A., Sim, Z. L., & Xu, F. (2017). "Why is Toma late to school again?" Preschoolers identify the most informative questions. *Developmental Psychology*, *53*(9), 1620. https://doi.org/10/gcjpjg

Ruggeri, A., Stanciu, O., Pelz, M., Gopnik, A., & Schulz, E. (2023). Preschoolers search longer when there is more information to be gained. *Developmental Science*. https://doi.org/10/gsbgvt

Ruggeri, A., Swaboda, N., Sim, Z. L., & Gopnik, A. (2019). Shake it baby, but only when needed: Preschoolers adapt their exploratory strategies to the information structure of the task. *Cognition*, *193*, 104013. https://doi.org/10/grqb4j

Rule, J., Goddu, M., Chu, J., Pinter, V., Reagan, E. R., Bonawitz, E., Gopnik, A., & Ullman, T. (2023). Fun isn't easy: Children choose more difficult options when "playing for fun" vs. "trying to win". https://doi.org/10.31234/osf.io/q7wh4

Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science*, *1*(2), 291–297. https://doi.org/10/bmkpxx

Serko, D., Leonard, J. A., & Ruggeri, A. (2022). Developmental changes in children's training strategies. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *44*(44).

Suddendorf, T., Crimston, J., & Redshaw, J. (2017). Preparatory responses to socially determined, mutually exclusive possibilities in chimpanzees and children. *Biology Letters*, *13*(6). https://doi.org/10/n3m8

Suddendorf, T., & Moore, C. (2011). Introduction to the special issue: The development of episodic foresight. *Cognitive Development*, *26*(4), 295–298. https://doi.org/10/c94fqn

Ten, A., Kaushik, P., Oudeyer, P.-Y., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications*, *12*(1), 1–10. https://doi.org/10/gm467m

Wang, J., & Bonawitz, E. (2022). Children's sensitivity to difficulty and reward probability when deciding to take on a task. *Journal of Cognition and Development*. https://doi.org/10/grmj56

Wente, A., Goddu, M., Garcia, T., Posner, E., Flecha, M., & Gopnik, A. (2019). Young children are wishful thinkers: The development of wishful thinking in 3- to 10-year-old children. *Child Development*, *91*. https://doi.org/10/gf6294

Whitebread, D., & Neale, D. (2020). Metacognition in early child development. *Translational Issues in Psychological Science*, *6*(1), 8–14. https://doi.org/10/gtphpx

Chapter 5

Discussion and Conclusion

5.1 Discussion

This thesis began with an old question: *can children direct their own learning?* Despite a growing body of empirical research on children's cognitive development, there appears to be scant agreement among educators, teachers, and researchers as to whether and when children can be expected to autonomously and efficiently adapt their learning. Therefore, educational approaches vary greatly, ranging from guided traditional methods to alternative models that emphasize the importance of self-guided learning. In this thesis, I focus on broadening our understanding of how children adapt their learning behaviors across three dimensions critical for self-guided learning: *task structures*, *goals*, and their *abilities*. Understanding from what age children can successfully master their learning across these dimensions and where they may benefit from guidance can help educators provide young learners with the necessary tools to navigate their learning efficiently and successfully.

To advance our understanding of children's adaptation across these three crucial dimensions, I introduce the idea of adaptation as a match between the requirements of the environment and an individual's abilities (see Todd & Gigerenzer, 2012). Such a perspective allows to understand young children's learning decisions as attempts to adapt to their environments with the cognitive capacities they possess. To assess young children's adaptations across three critical dimensions (structures, goals, and their abilities), I created behavioral studies that allow me to investigate their free exploration, information search, and practice decisions. My work provides evidence that young children's adaptive learning manifests across the three dimensions relevant to self-guided learning. At the same time, my results suggest that young children gain particular benefit from scaffolding, whereas older children can navigate their learning decisions more autonomously.

Summary of Chapters 2—4

In Chapter 2, I investigate children's capacity to adapt their exploration to the structural complexities inherent in their environments. I show that school-aged children adapt their exploration behaviors to environmental complexity. Specifically, my results indicate that older children increasingly incorporate prior learning about causal structures into their exploration in a free exploration task. Moreover, in a forced-choice task, older children showed a preference for specific cause and background combinations. Understanding which causal relationships we can trust to remain stable and which change across contexts is a critical component for understanding the complexities of our environment. My work shows that school-aged children may already be equipped with the necessary tools to succeed at this task.

In Chapter 3, I shift my attention toward investigating how young children adapt their information searches to different structures when facing an explicit goal. My results indicate that children as young as two years old successfully select the information that allows them to disambiguate the location of a hidden toy, even when presented with new target materials. These results suggest that previous work (Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018; Lindow, 2021), which suggested a much later emergence of children's information search abilities, may have significantly underestimated these abilities in young children. My work offers a new perspective on the early emergence of cognitive abilities in young children, enabling them to successfully navigate information structures.

In Chapter 4, I use a unified framework to investigate how children make adaptive practice decisions with regard to *task structures*, *goals*, and their *abilities*, instead of investigating these dimensions separately. I demonstrate that, under the right conditions, starting from the age of 4, children can make adaptive practice decisions. When they receive less scaffolding, it takes until the age of six for these adaptive abilities to be expressed behaviorally. These results show that children can engage in deliberate practice to *prepare* themselves for the future from a young age, even in the light of uncertainty about a future test, but young children may need support to succeed at these tasks.

Implications

With the above results in mind, let us return to the three questions raised at the beginning of this thesis:

1. Do children adapt their learning to the task *structures* they encounter?
2. Do children adapt their learning to achieve different *goals*?
3. Can children adapt their learning to their *abilities*?

My results indicate that, much earlier than expected (from age 2), children can successfully navigate the information structure they are presented with to reach an explicit goal, and even preschoolers (at the age of 4) integrate information about task structures, goals, and their performance that allows them to make deliberate practice decisions. However, at that age, scaffolding (in the form of child-friendly materials, reduced verbal demands, and asking the right comprehension-check questions) is important for young children to succeed in these tasks. Once children enter school, their deliberate practice decisions reflect their ability to integrate information for the multiple dimensions (structures, goals, and their abilities) proposed as central for self-guided learning and their free exploration becomes more attuned to the task structures they face.

These findings have implications for developmental research and educational practices. My results suggest that previous divergent findings (e.g., Betsch & Lang, 2013; Betsch et al., 2014; Betsch et al., 2018; Lindow, 2021; Metcalfe & Finn, 2013) regarding the developmental trajectory of children's adaptive learning may stem from methodological choices rather than children's cognitive capacities. Experimental settings should therefore present children with tasks that are age-appropriate and cognitively suitable, avoiding confounds from memory demands, overly complex instructions, or non-intuitive materials. Key cognitive functions, such as executive function, working memory, and inhibitory control—all relevant for engaging in self-guided learning—are still developing in young children (Blackwell & Munakata, 2014; Chatham et al., 2009; Chevalier et al., 2014; Chevalier et al., 2015; Roebers, 2017), underlining the importance of controlling for these potential confounds. Providing children with memory aids, child-friendly materials, or comprehension check questions can therefore lead to a more precise evaluation of their cognitive abilities. These results should provoke further questions, such as: What are other cognitive capacities of young children waiting to be discovered? And how can we improve, or—to stay consistent with the terminology introduced in this thesis—*adapt* our research methods in a way that they provide us with a more accurate picture of what young children are capable of?

Beyond these questions, which are relevant for research, my work has implications for educational practice. In particular, the results presented in Chapter 3 and Chapter 4 show that 2-year-olds can disambiguate between relevant and irrelevant information when searching for a present, and 4-year-olds can make adaptive practice decisions based on the assessment of the task structures and goals they face, and their individual abilities. Both findings highlight that young children can engage in very sophisticated learning decisions provided they have the appropriate amount of scaffolding, whereas older children may already possess the necessary tools to thrive in more open and self-guided learning environments. A promising way to apply these findings in educational practice could be to provide young learners with structured learning environments, allowing them to direct their learning within these settings, and progressively reduce these constraints for older children.

What could this mean more specifically? Young children's successful adaptation in both studies underscores how questions can facilitate children's adaptive learning. Consider, for example, the second experiment presented in Chapter 4. Here, comprehension check questions reminded children of the task structures and goals they face (e.g., "What is the task you will later face?" or "How many attempts do you have?") and their prior performance ("Which game was easy/difficult for you?"). Along with the right study materials, these questions likely enhanced young children's performance. Thus, educators can facilitate young children's self-guided learning by posing questions that heighten their awareness of task structures, goals, and their abilities. For example, educators might inquire which tasks children find challenging or straightforward, whether they recall the task goal, and how they previously performed on a task. Scaffolds like these can support children's self-guided learning even before formal education begins, so that once children enter primary school, they might already be equipped with the necessary tools to successfully engage in self-guided learning. Identifying the right balance between age-appropriate scaffolding and self-guidance will be the work of future research, for which I propose ideas in the next chapter.

Open Questions and Future Research Directions

First, I have presented evidence that children across all age groups can (to varying degrees) adapt their behaviors to structures, goals, and their abilities. However, the results do not allow us to assess individual differences in children's adaptation. Within each age group, I found considerable variation in the extent to which children adapted to the respective tasks they faced, indicating that some of the youngest children in our samples may have mastered tasks that some of the older children were struggling with. Although these results are merely anecdotal, they highlight the importance of further research to investigate individual differences in children's ability to adapt their behaviors. After all, my results indicate how groups of children behaved *on average* in the respective task they were facing, but not individually. However, in the real world, humans do not exist on average. Consequently, it is possible that some children profit from guidance whereas others may need more learning autonomy. Moreover, it is possible that the same child may profit from scaffolding in one dimension, and from self-guided learning in another. Future research should aim to capture individual differences of children to open up the potential for tailored educational practices where each child can prosper according to their individual abilities and needs.

Second, and connected to the previous point, I would like to remind the reader that I have treated structures, goals, and a person's abilities as separate, but intertwined dimensions in which adaptation can manifest. However, my work leaves open whether it is appropriate to conceptualize and investigate adaptation as a general capability with different subcomponents, or whether we would be better advised to think of the different dimensions of adaptation presented here as independently operating cognitive capacities. In other words, are some children generally better at adapting their learning behaviors to all three dimensions, or should we expect some children to be good at adapting their learning decisions to structures, others to be proficient at adapting to goals, and still others to excel at adapting to their abilities? Moreover, can we expect that a child's ability to assess their abilities and make adaptive learning decisions based on such assessments will transfer to their ability to adapt their learning to structures and goals, and vice versa? Although it is not within the scope of this research to answer these questions, my experiments open the possibility of addressing them in the future. A promising avenue for future research could involve presenting children with separate tasks that assess different cognitive dimensions of adaptation (e.g., the tasks presented here) and then analyzing whether the children excel in all of the tasks or only in some, and how children's capacities for adapting to the different dimensions are related. This would allow us to investigate where adaptation abilities overlap and where they diverge, and whether it makes sense to think of adaptation as a general construct.

Finally, my work indicates that starting from age 4, children can adapt their practice choices to the task goals and the probabilistic structure of a task based on an assessment of their performance. However, it remains unclear how accurately children estimate their abilities, knowledge, strengths, and weaknesses to then apply these insights in practice. In future studies, I plan to explore the relationship between children's ability to reflect upon the limitations of their knowledge and their *active learning about themselves*. A potential approach for investigating this would be to allow children to perform different tasks without providing performance feedback and then manipulate and combine various variables such as cooperation and competition among potential partners, the difficulty of games, and whether children face explicit goals or play for fun. We could then measure under which conditions children want to learn about their performance and what aspects of their performance they seek knowledge about. This approach would enable us to study how children direct their active learning towards themselves, as well as whether and when they use these estimates to make future-directed learning decisions.

5.2 Conclusion

The research presented in this dissertation offers compelling evidence that preschool-aged children are capable of directing their learning, particularly when supported by appropriate scaffolding. Older children appear to possess the requisite skills to employ more sophisticated adaptive learning strategies. This thesis significantly enriches our theoretical and empirical understanding of how children adapt their learning across three dimensions central for self-guided learning: the *task structures*, *goals*, and their *abilities*. My work introduces novel and age-appropriate experimental methods relevant to researchers interested in capturing children's learning behaviors. Moreover, the results obtained could inform educational practice. Specifically, my findings underscore that young children are likely to benefit most from learning environments characterized by clear constraints, including explicit objectives, straightforward instructions, child-friendly materials, and guiding questions. In contrast, school-aged children might already be equipped with the necessary tools to navigate more open learning environments. My work advances our understanding of children's adaptive learning and provides new insights into the evolving capabilities of young learners.

5.3 References

Betsch, T., & Lang, A. (2013). Utilization of probabilistic cues in the presence of irrelevant information: A comparison of risky choice in children and adults. *Journal of Experimental Child Psychology*, *115*(1), 108–125. https://doi.org/10/f4vxqn

Betsch, T., Lang, A., Lehmann, A., & Axmann, J. M. (2014). Utilizing probabilities as decision weights in closed and open information boards: A comparison of children and adults. *Acta Psychologica*, *153*, 74–86. https://doi.org/10/f6vd8w

Betsch, T., Lehmann, A., Jekel, M., Lindow, S., & Glöckner, A. (2018). Children's application of decision strategies in a compensatory environment. *Judgment and Decision Making*, *13*(6), 514–528. https://doi.org/10/grqb3t

Blackwell, K. A., & Munakata, Y. (2014). Costs and benefits linked to developments in cognitive control. *Developmental Science*, *17*(2), 203–211. https://doi.org/10/f5t7rn

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences*, *106*(14), 5529–5533. https://doi.org/10/b5x9gz

Chevalier, N., Chatham, C. H., & Munakata, Y. (2014). The practice of going helps children to stop: The importance of context monitoring in inhibitory control. *Journal of Experimental Psychology: General*, *143*(3), 959–965. https://doi.org/10/f56d69

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience*, *27*(6), 1125–1136. https://doi.org/10/gc5ss6

Lindow, S. (2021). On searching and finding: The development of information search abilities. *Cognitive Development*, *58*, 101011. https://doi.org/10/grcfkx

Metcalfe, J., & Finn, B. (2013). Metacognition and control of study choice in children. *Metacognition and Learning*, *8*(1), 19–46. https://doi.org/10/f4qkjg

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review*, *45*, 31–51. https://doi.org/10/gbqdb7

Todd, P. M., & Gigerenzer, G. (2012). *Ecological rationality: Intelligence in the world*. Oxford University Press, USA.

Appendix

This appendix provides additional details, analyses, and supporting evidence that complement the findings and discussions presented in Chapter 4 of this dissertation.

Experiment 1

We first present a hypothesis rationale that participants in our experiment might have used to decide which game to practice when facing different conditions. We then examine the child and adult samples separately, addressing the manipulation check and comprehension check questions. First, we demonstrate that our statistical analyses remain robust even when excluding rounds where children's answers during the familiarization phase (i.e., how many items they guessed correctly) deviated from our expectations (e.g., they guessed fewer or more items correctly than 3/4 in the easy game or more items correctly than 0/4 in the difficult game). Second, we show that our statistical analyses remain significant even when including trials where children failed to answer the comprehension check question correctly, in which they had to indicate which of the two games they performed better in. Finally, we provide a comprehensive overview of the analyses performed on both child and adult samples for each round.

Hypothesis Rationale

In Experiment 1, participants practice choices may have followed the following rationale: In the familiarization phase, participants played an easy game consisting of three easy and one difficult to guess item (we implemented one difficult to guess item in order to make the game easy, but not to easy, see Methods for a detailed explanation) and one difficult game consisting of 4 difficult to guess items. We predicted that after having familiarized themselves with the easy and the difficult game and when facing the *Test-Easy* condition they should practice the easy game, where they scored 3/4 (i.e., 3 out of 4). This would significantly improve their performance at test and allow them to reach the maximum score at test (8/8). If they practiced the difficult game instead, they would achieve a minimum score of 3/8 (the items they had already guessed correctly during the familiarization phase) and a maximum of 7/8, if they were able to guess all the test items. Note that if participants projected the same proportion of easy/difficult-to-guess items in the test as in the familiarization set (3/4, with one difficult item), they should expect to achieve a score of 6/8 on average at test if they trained the difficult game. When facing the *Test-Difficult* condition, participants should practice the difficult game, where they scored 0/4, to increase their likelihood of reaching the maximum score on the test (8/8). If they practiced the easy game instead, they would achieve a minimum average score of 0/8 and a maximum average score of 4/8, if they were able to guess all the test items. Note that if participants projected the same proportion of easy/difficult-to-guess items in the test as in the familiarization set (0/4, with 4 difficult items), they should expect to achieve an average score of 0/8 at test if they trained the easy game. When facing the *Test-Random* condition participants should practice the difficult game. In this way, if they were tested in the easy game (50% chance), but had practiced the difficult game they would achieve a minimum score of 3/8 (those they had already guessed during familiarization) and a maximum score of 7/8—or 6/8 if participants projected the same proportion of easy/difficult-to-guess items in the test as in the familiarization set (3/4, with one difficult item); if they were tested in the difficult game (50% chance) and had practiced the difficult game they would achieve the maximum score (8/8) assuming perfect recall of the practiced items. If they practiced the easy game, they would achieve the maximum score 8/8 if tested on the easy game, but a score of 0 if tested on the difficult game.

Child Sample

Manipulation Check

We checked whether we successfully manipulated game difficulties. We expected children to guess 3/4 items correctly in the easy game, and 0/4 items correctly in the difficult game (see Table A 1 and Table A 2. In the analyses we report in the main paper we included children regardless of whether they performed the familiarization phase as we expected. Note that this inclusion had no effect on the statistical results we found.

Table A 1. Children's Cumulative Performance in the Easy Game Familiarization Across Three Rounds

Items Guessed	0/4	1/4	2/4	3/4	4/4
Count	0	4	30	299	12
Percentage (%)	0	1.16	8.80	86.67	3.48

Table A 2. Children's Cumulative Performance in the Difficult Game Familiarization Across Three Rounds

Items Guessed	0/4	1/4	2/4	3/4	4/4
Count	330	15	0	0	0
Percentage (%)	95.65	4.35	0	0	0

Comprehension Check

After children played the easy and the difficult game, they performed a comprehension check in which they had to indicate the game at which they were better. Participants successfully completed the comprehension check in 238 out of 245 rounds (97%). In the analyses of the main paper we excluded 7 rounds in which participants failed to answer the comprehension check correctly. This removal had no effect on the statistical results we found.

Children's Practice Choices by Round

Round 1

We analyzed children's practice choices in round 1. A logistic mixed-effects model predicting children's practice choices (easy or difficult) with condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy* as baseline) and age in months as fixed effects and participants' ID as a random effect revealed main effects of the *Test-Difficult* condition (p = .028, OR = 0.314 [0.108 – 0.855]) the *Test-Random* condition (p = .014, OR = 0.289 [0.103 – 0.759]), and age in months (p = .002, OR = 0.506 [0.318 – 0.772]).

Looking at the interaction effects of the interaction of condition and age revealed no significant effects (*Test-Difficult* condition and age in months: p = .264, OR = 0.524 [0.147 – 1.528]; *Test-Random* condition and age in months: p = .840, OR = 0.895 [0.292 – 2.609]).

Round 2

Next, we analyzed children practice choices in round 2. A logistic mixed-effects model predicting children's practice choices (easy or difficult) with condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy* as baseline) and age in months as fixed effects and participants' ID as a random effect revealed main effects of the *Test-Difficult* condition (p = .001, OR = 0.163 [0.052 – 0.466]) the *Test-Random* condition (p = .089, OR = 0.449 [0.175 – 1.120]), but no effect of age in months (p = .459, OR = 0.859 [0.571 – 1.282]).

Looking at the interaction effects of conditions and age revealed a significant effect of the interaction of condition and age (*Test-Difficult* condition and age in months: $p < .007$, $OR = 0.140$ [0.029 – 0.530]; *Test-Random* condition and age in months: $p = .010$, $OR = 0.256$ [0.084 – 0.691]).

Round 3

Next, we analyzed children practice choices in round 3. A logistic mixed-effects model predicting children's practice choices (easy or difficult) with condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy* as baseline) and age in months as fixed effects and participants' ID as a random effect revealed main effects of the *Test-Difficult* condition (p = .002, OR = 0.200 [0.069 – 0.534]) the *Test-*

Random condition ($p < .001$, $OR = 0.106$ [0.030 – 0.033]), and age in months ($p = .002$, $OR = 0.475$ [0.290 – 0.748]).

Looking at the interaction effects of conditions and age revealed a significant effect of the interaction of condition and age (*Test-Difficult* condition and age in months: $p = .007$, $OR = 0.902$ [0.831 – 0.967]; *Test-Random* condition and age in months: $p = .054$, $OR = 0.923$ [0.839 – 0.994]).

Children's Practice Choice by Age Group

Table A 3. Children's Active Practice Choice by Age Group

Age in Years	N Adaptive Practice	N per Age Group	Percent (%)
4	39	71	54.93
5	37	70	52.86
6	43	65	66.15
7	53	66	80.30
8	54	66	81.81

Children's Test Scores

Overall performance

Table A 4. Study 3 Exp 1 - Average number of points achieved by children in different conditions and age groups over three rounds

	Age in Years				
Condition	4	5	6	7	8
Test-Easy	7.38	7.00	7.55	7.46	7.59
Test-Difficult	2.17	2.57	4.91	6.00	6.09
Test-Random	2.44	1.71	3.41	5.36	5.64

Round 1

Table A 5. Study 3 Exp 1 - Average number of points achieved by children in different conditions and age groups in Round 1.

	Age in Years				
Condition	4	5	6	7	8
Test-Easy	7.66	7.00	7.57	7.18	7.46
Test-Difficult	2.09	3.43	3.00	5.00	6.57
Test-Random	3.57	1.13	3.70	6.14	4.25

Round 2

Table A 6. Study 3 Exp 1 - Average number of points achieved by children in different conditions and age groups in Round 2.

	Age in Years				
Condition	4	5	6	7	8
Test-Easy	7.38	7.25	7.80	7.62	8.00
Test-Difficult	4.33	3.00	6.44	6.88	5.5
Test-Random	2.25	1.00	2.50	4.17	5.70

Round 3

Table A 7. Study 3 Exp 1 - Average number of points achieved by children in different conditions and age groups in Round 3.

	Age in Years				
Condition	4	5	6	7	8
Test-Easy	7.20	6.71	7.40	8.00	7.57
Test-Difficult	1.56	1.38	4.12	5.70	6.29
Test-Random	1.00	3.00	4.50	5.56	6.25

Practice choices by media split

To further investigate the observed age effects, we performed exploratory analyses within a younger and older age group via a median split by age (cutoff at 75 months; age 6.25 years). A logistic mixed-effects model predicting practice choices (easy or difficult) in the older age group with condition (*Test-Test-Easy, Test-Test-Difficult, Test-Test-Random; Test-Test-Easy* as baseline) and age in months as fixed effects and participants' ID as a random effect revealed a significant effect of the *Test-Test-Difficult* condition (p < .001, OR = 0.044 [0.012 – 0.161]) and the *Test-Test-Random* condition (p < .001, OR = 0.125 [0.046 – 0.344]), but not age in months (p = .195, OR = 0.762 [0.505 – 1.150]). In the older age group, 36% of the children (n = 19/53) selected the difficult game in the *Test-Test-Easy* condition (significantly lower than chance; p = .026; two-tailed 50% binomial test). In the *Test-Test-Difficult* condition and *Test-Test-Random* condition 93% (n = 49/53) and 81% (n = 43/53) of the older children, respectively, selected the difficult game. Both proportions were significantly greater than chance (*Test-Test-Difficult* condition: p < .001; *Test-Test-Random* condition: p < .001, two-tailed 50% binomial test). A chi-square test revealed no difference in task choice between the *Test-Test-Difficult* and *Test-Test-Random* conditions ($\chi^2(1) = 2.058$, p = .151). The same model did not reveal any significant main effects of condition or age in the younger age group (see Fig. A 1).

Retrospective Evaluation

After each test, we reminded children of their practice choices during the practice phase and asked whether they believed their choices were a good idea (yes or no). We introduced a binary dummy variable to assess the alignment between their practice decisions and their retrospective evaluations. We assigned a value of 1 to cases where participants made adaptive practice choices (e.g., practicing the easy game in the *Test-Easy* condition, or the difficult game in the *Test-Difficult* and *Test-Random* conditions) and indicated that these choices were retrospectively beneficial. Similarly, a value of 1 was assigned if participants made non-adaptive practice decisions and retrospectively indicated them as unfavorable. In all other cases, when there was no alignment between practice decisions and retrospective evaluations, we assigned a value of 0.

A logistic mixed-effects model, predicting a match between practice choices and their retrospective evaluation (coded as 0/1) by condition, age in months, and their interaction as fixed effects, and participants' ID as a random effect revealed a significant interaction interaction effect of the *Test-Difficult* condition and age (p = .005, OR = 1.095 [1.027 – 1.167]); as well as the *Test-Random* condition and age (p = .004, OR = 1.089 [1.028 – 1.153]). Indicating that with age children in both condition were more accurate in retrospectively evaluating the practice choices they had made.

Adult Sample

Manipulation Check

We first checked whether we successfully manipulated game difficulties (in the easy game participants should guess 3/4 items, in the difficult game they should guess 0/4 items; see Table A 8 and Table A 9).

Table A 8. Study 3 Exp 1 - Comparison of adults' performance in the familiarization with the easy game

Items Guessed	0/4	1/4	2/4	3/4	4/4
Count	0	0	7	80	6
Percentage (%)	0	0	7.53	86.02	6.45

Table A 9. Study 3 Exp 1 - Comparison of adults' performance in the familiarization with the difficult game

Items Guessed	0/4	1/4	2/4	3/4	4/4
Count	73	19	0	1	0
Percentage (%)	78.49	20.43	0	1.07	0

Comprehension Check

After participants played each game they performed a comprehension check in which they had to indicate the game in which they were better at. All participants correctly indicated the easy game. All participants answered the comprehension check questions correctly.

Adult's Practice Choices by Round

Round 1

A logistic mixed-effects model predicting participants' practice choices (easy or difficult) with condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy* as baseline) and age in years as fixed effects and participants' ID as a random effect revealed main effects of the *Test-Difficult* condition ($p < .001$, $OR = 0.032$ [0.005 – 0.208]) and the *Test-Random* condition ($p < .001$, $OR = 0.018$ [0.003 – 0.116]).

Round 2

We first looked at adults' behavior in round 2. A logistic mixed-effects model predicting participants' practice choices (easy or difficult) with condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy*

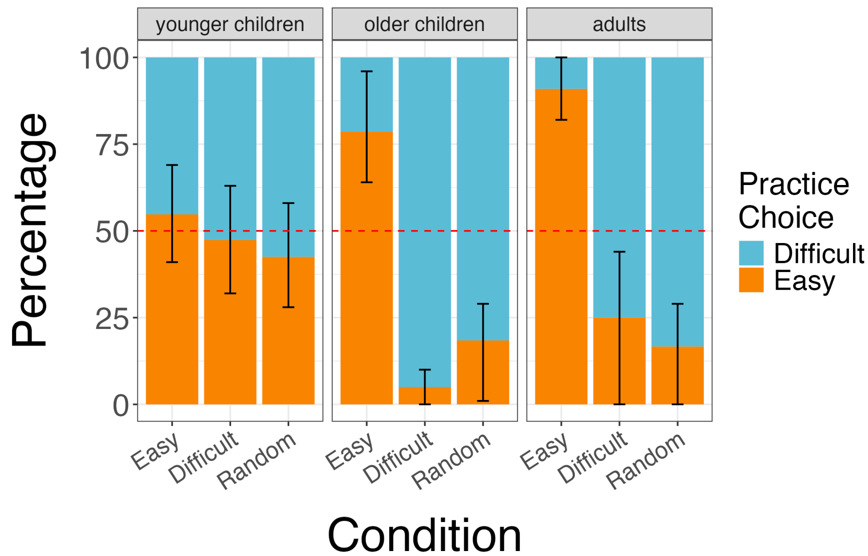


Figure A 1. Illustration of participants' practice choices by condition in Experiment 1. Participants' practice choices, displayed by age group (young children vs old children (cutoff at 75 months) vs adults) and condition. Bars represent *the Easy, Difficult*, and *Random* conditions. The red dotted line indicates the 50% chance level. The error bars extending above and below the bars represent the 95% confidence intervals, indicating the range within which the true population means are likely to fall.

as baseline) and age in years as fixed effects and participants' ID as a random effect revealed main effects of the *Test-Difficult* condition (p = .001, OR = 0.625 [0.266 – 0.984]) and the *Test-Random* condition (p < .001, OR = 0.769 [0.451 – 1.088]).

Round 3

Next, we analyzed participants practice choices in round 3. A logistic mixed-effects model predicting participants' practice choices (easy or difficult) with condition (*Test-Easy, Test-Difficult, Test-Random; Test-Easy* as baseline) and age in years as fixed effects and participants' ID as a random effect revealed main effects of the *Test-Difficult* condition (p = .007, OR = 0.285 [0.001 – 0.273]) and the *Test-Random* condition (p = .003, OR = 0.020 [0.001 – 0.182]).

Retrospective Evaluation

We introduced the same binary dummy variable to assess the alignment between practice decisions and participants' retrospective evaluations as in the child sample. A logistic mixed-effects model, predicting a match between adults' practice choices and their retrospective evaluation (coded as 0/1) by condition as a fixed effect, and participants' ID as a random effect revealed no significant effects.

Experiment 2

Manipulation Check

We first analyzed how much time it took children to build the easy and the difficult tower during familiarization phase and whether these times differed significantly between the easy and the difficult games. As expected, children needed less time to build the easy tower ($M = 8.47 seconds$, $SD = 4.38$) compared to building the difficult tower ($M = 20.58 seconds$, $SD = 16.57$; $t(81.985) = -6.036$, $p < 0.001$, Welch two-sample t-test).

Comprehension Check

We excluded a total 13 children because they failed to answer one or multiple of the comprehension check question correctly even after we repeated the relevant part of the procedure with them for a maximum of three times. In particular, we excluded 7 children because they claimed to know in which game they would be tested in the *Test-Random* condition, even after demonstrating the procedure to them several times. We excluded 6 children who failed to indicate the easy game as easy and the difficult game as difficult.

Retrospective Evaluation

We introduced the same binary dummy variable to assess the alignment between practice decisions and participants' retrospective evaluations as in Experiment 1.

A logistic mixed-effects model, predicting a match between practice choices and children's retrospective evaluation (coded as 0/1) by condition revealed no significant effects.

Declaration of Independent Work

I hereby declare that:

- I completed this doctoral thesis independently. Except where otherwise stated, I confirm that the work presented in this thesis is my own.
- Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
- I have not applied for a doctoral degree elsewhere and do not have a corresponding doctoral degree.
- I have acknowledged the Doctoral Degree Regulations which underlie the procedure of the Department of Education and Psychology of Technical University Munich.
- The principles of Technical University Munich for ensuring good academic practice have been complied with.

[redacted]

Daniil Serko

Berlin, 05.04.2024