Active learning in infancy and adulthood: individual strategies for information sampling

Active learning in infancy and adulthood: individual strategies for information sampling

Cécile Gabrielle Gal

Department of Psychological Sciences

Birkbeck, University of London

A thesis submitted for the degree of

Doctor of Philosophy

2022

Supervised by:

Dr Teodora Gliga
University of East Anglia

Dr Marie L. Smith
Birkbeck, University of London

Dr Eng. Raul C. Mureșan
Transylvanian Institute of Neuroscience
Neurodynamics

Evaluated by:

Prof. Gaia Scerif
University of Oxford

Prof. Natasha Z. Kirkham
Birkbeck, University of London

Originality statement

I, Cécile Gabrielle Gal, hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at the University of London or any other educational institution, except where due acknowledgment is made in the thesis. Any contribution made to the research by others, with whom I have worked at University of London or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation, and linguistic expression is acknowledged.

COVID impact statement

Part of this PhD took place during the COVID pandemic. Beyond general complications and added stress that the pandemic brought for all, this event luckily did not impact much on the flow of the studies. The data recordings for both infant studies were set to be finished on the 13th of March 2020, a week before the UK was first put into lockdown. Recordings for the 15-month-olds study (Chapters 5 and 6) were completed that week, while some of the appointments taken for the 10-month-olds study (Chapters 3 and 4) were understandably cancelled, hence the number of participants for this study did not reach the targeted number (31 usable datasets rather than the targeted 35). The number of recordings obtained before the laboratory was shut down for about a year was deemed close enough to the target for it not to affect the viability of the study (more details on this issue are reported in the Methods section of Chapter 3).

Abstract

Humans are astounding learners. They don't passively absorb information but actively engage in the process: they select information according to their own characteristics e.g., their state of knowledge, abilities, needs or goals, which has the potential to profoundly impact on how individuals experience the world. This PhD aimed to characterise different ways in which agents tailor their sampling of information to fit their priors' strength, attentional skills, learning progress or executive functioning. In a first study (Chap. 2) conducted with adults, we looked at the influence of priors (prior access to informative stimuli) on visual objects recognition and exploration. Priors enabled participants to guide their fixations to quantitatively more informative locations when exploring ambiguous stimuli. However, presenting stimuli of varying ambiguity levels in a random fashion destroyed this ability to guide exploration with specific priors. In a second study (Chap. 3 and 4) using electroencephalography (EEG), we showed that 10-month-old infants' parent-reported trait attention was linked to their processing of an information stream in which visual distractors interrupted an ongoing movie. Importantly, we found that infants' trait sensory processing as reported by parents was not only linked to their engagement with the task, but also to their brain response to distractors, linking together several levels of individual differences in information processing. At the brain level, we found a crucial role of occipital high-frequency gamma-range EEG activity and, for the first time in infants, of its alignment with lower-frequency activity for blocking the processing of distractors vs. the ongoing video. These results bring in new and valuable information for theories of how the brain processes information and implements attentional mechanisms early on during life. Finally, in a last study, we looked at the influence of learning progress (Chap. 5) and executive functions (Chap. 6) on how 15-month-old infants learn and explore. We showed that learning progress at the start of a matching-rule learning task but not the achieved score per se, predicted how long infants would stay on the task. We also investigated the existence of overarching exploration strategies in infants by looking at exploratory depth's and breadth's stability within the visual modality and between the visual and manual modalities, as well as their link with individual differences in executive function. We only found evidence for stability in exploratory breadth between two sets of visual trials. This visual exploration breadth was positively correlated with participants' age and visual working memory. Overall, this thesis brings together studies with different age groups and techniques which all point to the fact that individuals actively shape their own sampling of information in a deterministic fashion that suits their personal state and abilities.

Acknowledgements

This thesis would never have seen the day without the support of my team of supervisors, Teodora Gliga, Marie Smith and Raul Mureșan. Teea, thank you for bringing this team and this project together, for your infallible availability and for critically questioning this work every step of the way. Your analytical mind and your tireless guidance vastly improved my work and pushed me to learn so much. Marie, thank you for welcoming me into your team, for your careful eye throughout this project and for gently checking in when I was staggering. Raul, you have greeted me into your lab as in a family. I am most grateful for your generosity in making my experience in Romania the best possible. A piece of me certainly turned Romanian while in Cluj! Thank you for introducing me to complicated methods with contagious confidence in my ability to grasp them and a lot of pedagogy.

Thanks also go to my funder, the European Commission's Marie Skłodowska-Curie Actions, and all the people who helped me with this work. I am thinking especially of the participants and families who took part in the studies, the researchers Elena Piccardi and Ioana Țincaș whose prior work was instrumental, and the interns wo helped on this project: Francesca Penza, Diniyah Rehman, Rebecca Simon, Cristina Vella, Cara Hatton, Odessa Hamilton and Cristina Pelea. I am also terribly thankful to the research assistants who stepped in when I needed to quickly adapt: Francesca, who amazed with her eagerness to always learn and take on more responsibilities (good luck with your PhD applications!), as well as Tamsin Osborne and Isabel Quiroz, who selflessly made time to help me with testing. I must also thank Leslie Tucker, Berta Hortiguela and the INTERLEARN team for helping me to swiftly adapt my plans when last-minute changes came.

I am also so grateful for all the people that I met along this path and who quickly became the loveliest partners in crime throughout this adventure. To my INTERLEARN teammates Louisa Gossé, Susanne de Mooij, Alicja Brzozowska and Giada Guerra: your intelligence, kindness and camaraderie has been more empowering than I could say. I am very proud to have been part of our cohort! I am also so grateful for the people I met at the CBCD who kept me in cheerful company: Aude Carteron, Jessica Massonié (thank you for being my mentor!), Emily Thomas, Kathryn Bates, Ana Portugal, Viktoria Csink, and all the others. Finally, I warmly thank my Cluj team for the warmth with which they welcomed me in their life: Medorian Gheorghiu (thank you for even giving me a home!), Ana-Maria Ichim, Loredana Dan, Harald Bârzan, Andrei Ciupariu, Mădălina Mastacan, Adriana Dăbâcan, Vlad Moca, George Grosu, Cătălin Coltău, Richard Ardelean, Cosmina Pavel and Laura Rus. I must also acknowledge the short- and long-distance support that my former Dual Masters cohort still brings to me. Thank you for being such a bright and bighearted group

who keeps inspiring me as a person and a researcher every day. Special mention to Lindsay Rondot, Stella Guldner, Eva Islimye, Annie Gaule, Jesse Geerts, Cassandra Visconti and Margot Tirole for being there for me each in their own way through life and neuroscience ups and downs. I would also like to thank my supervisors in former projects, without whom I night never have built the confidence to follow this scientific path. Special thanks to Xavier Houard for offering me my first internship in a science laboratory and teaching me so generously, and to Maria Chait and Nicolas Barascud for mentoring me on my first Neuroscience project in such a kind and conscientious way.

I would never have reached this point without the loving support of all my non-neuroscience friends who kept me happy and grounded: Gia, Diana, Kim, Massimo, Émilie, Chloé, as well as Lea, Ana and Eryck who kept me moving, and all the others who shared my daily life through this journey. Special mention to Léo for all the energy you gave me, and Daphné for taking me under your wing and getting me to finish this thesis. To Voicu: thank you for keeping my heart happy and my head under a roof. I cannot wait to meet you on the other side.

Finally, le plus grand merci to my brilliant and most supportive family. Your strength, brightness and bigheartedness gently push me to better myself every day. Thank you to my parents Catherine and Olivier, my grand-parents Mireille, Jean-Claude and Pierre-Yvan, and my brothers Louis and Raoul for your visits and your warm welcomes when I came home, and your caring support in every way. The last word goes to my sister Jacqueline, for always remaining my biggest source of inspiration for courage, grace and kindness. Thank you for relentlessly lifting me up in your own way.

Table of contents

List of tables

List of figures

List of abbreviations

ADHD	Attention Deficit and Hyperactivity Disorder
ANOVA	Analysis of Variance
AOI	Area of Interest
ASD	Autism Spectrum Disorder
ASQ	Ages and Stages Questionnaire
ECBQ	Early Childhood Behavior Questionnaire
EcoG	Electrocorticogram
EEG	Electroencephalography
ERP	Event Related Potential
fMRI	Functional Magnetic Resonance Imaging
fNIRS	Functional Near-Infrared Spectroscopy
IBG	Infant Behavior Questionnaire
ICA	Independent Component Analysis
ICC	Intraclass Correlation Coefficient
ISI	Inter-Stimulus Interval
ITSP	Infant/Toddler Sensory Profile
IQ	Intelligence Quotient
LCD	Local Contour Density
LDD	Local Dots Displacement
LP	Learning-progress
MEG	Magnetoencephalography
SCA	Scaled Correlation Analysis

S.D. Standard Deviation

S.E. Standard Error

TF Time-frequency

Chapter 1: General introduction

Humans are constantly learning about the world around them: in a highly complex and ever-changing environment, they need to sample and process the information around them, build knowledge and adapt their behaviour. Through their actions, they actively sample information to learn from: they can orient their gaze to a source of information or on the contrary gaze away from it, orient their head towards a source of sound, move in the direction of the source, grasp a related object, touch it, etc. There is a whole range of actions to be taken that offer people the possibility to act not just as passive observers, but really as active learners. Even young infants have been shown to adapt their actions to modulate their stimulation. In an influential experiment with 4-month-old infants, Piaget (1952) showed that when attaching infants' hand to a rattle of which they had no previous knowledge, they slowly learned that shaking the rattle provided increased multisensory stimulation, leading them to intensify their hand movements for more stimulation from the rattle. Actively sampling information and choosing what to engage with thus appears to be a key ability of humans that is already present soon after birth. This brings the question of what it is that drives an individual's urge to actively sample one type of information over another or to remain focused on one source of information rather than moving on to another.

1.1. Theories of active learning

The study of how people actively learn and autonomously engage with information has a long history in Psychology. Theories have been proposed from a variety of angles and in several waves, often following influential streams of thoughts of the time. Until the 1960's, accounts essentially stemmed from a behaviourist point of view and placed a strong emphasis on the reward-related aspect of the question, focusing on defining the nature of this reaction to or drive for information, while later theories stirred the debate away from this issue and sought to characterise types of information content associated with active information-seeking behaviours. Although this separation into two types of theories simplifies the debates into two extremes and does not fully

reflect the complexity of the ideas, which often touched upon both issues, it helps to facilitate the identification of their main message. Early reward-based theories will be presented first, followed by descriptions of later theories related to information-content.

1.1.1. Reward-based theories

1.1.1.1. Classical reinforcement-learning theories

The emergence of behaviourism in the 1900s laid the ground for the first systematic and quantitative studies of behaviour, moving away from older introspective approaches to Psychology. Behaviourism bore the general idea that individuals mostly react to information in their environment and learn by association of concurrent events. This was shown in Pavlov's classical conditioning experiments, in which a neutral stimulus, a bell sound, was repeatedly presented to a dog in combination with another loaded stimulus, food. Eventually, the dog learned to associate one with the other and began to salivate in expectation of food when presented with the bell alone (Pavlov & Thompson, 1910). Skinner took this further and showed that individuals can learn to repeat or avoid any type of actions (operant) that they take themselves depending on their outcome (Skinner, 1938). His experiments, mainly conducted with rats and pigeons, showed that rewarding outcomes reinforced behaviours, which the animals were more likely to repeat, while punishing outcomes weakened these behaviours. The same idea of a "Law of Effects" (Thorndike, 1927) was also expressed by Thorndike following his puzzle-box experiments with cats. He placed cats in boxes from which they could escape by pressing a lever and placed a piece of meat in front of the box. While the cats first had no knowledge of the box and the action of pressing the lever to get out of it, they quickly learned the association between the lever-press action, the door opening and the food reward, eventually becoming faster to escape with each repetition (Thorndike, 1898). This reinforcement learning framework (Sutton & Barto, 1998) based on reactions to rewards and punishments was extremely fruitful for explaining a large part of how individuals behave. For example, it is at the core of the Nobel-rewarded prospect theory, which operationalises the comparative effects of gains and losses on people's emotions and decisions, and shaped the field of economics (Kahneman & Tversky, 1979). However, it did not fully encompass the complexity of human behaviour when it comes to how they seek information. Indeed, conditioning settings suppose that agents only react to the value of the outcome, with no account made of the role that their internal states might play. Yet,

humans are notoriously curious and happen to spend their energy on gathering information even in cases when there seems to be no obvious external reward, a situation which cannot be explained by reward-punishment reinforcement.

1.1.1.2. Drive theories

Later research slowly stepped away from the behaviourist concept that individuals' actions can be explained by conditioning responses by adding a new element to the idea: the notion that human behaviour can be explained by internal drives to fulfil core needs through actions on the environment, and not merely by reactions to external events in the environment in the classical behaviourist sense. Hull was the first to articulate a drive theory of learning, which he described as a specific biological deficit in information that the organism needs to try and attenuate (Hull, 1943). Much as hunger or thirst are reduced by eating and drinking, the drive for learning could be reduced by gathering information. Numerous variations on this idea were formulated. Harlow instead talked about a drive for *manipulation*, following his experiments of complex puzzle-solving with monkeys (Harlow, 1950). Montgomery studied rats' exploratory choices when placed in mazes with two-choice arms and came to the conclusion that they were driven by a need for *exploration* (Montgomery, 1954). Fowler also studied rats' behaviour in mazes but instead proposed a drive to *fight boredom* (Fowler, 1965). Hebb, based on his studies of sensory deprivation in young dogs and human adults, proposed a drive for *optimal arousal* (Hebb, 1955, 1958). Through investigations of humans' aesthetical preferences, Berlyne came up with a similar drive for *intermediate complexity, novelty and surprise* (Berlyne, 1950, 1954).

The list of drive theories is long and attests to the field's preponderant tendency throughout the 1950's and 1960's to focus much of the debate on defining the precise nature of this drive, coming from different perspectives and methods. A lot of effort was also expended on determining the valence of the reward aspect of the information drive: are agents avoiding an aversive state of lacking information, or are they seeking information to reach a pleasant state of knowledge? Recent evidence using a lottery task and self-report measures of curiosity and happiness suggests that this drive for information might rather arise in an effort to avoid an aversive state of uncertainty, but the authors concede that the debate remains open today (van Lieshout et al., 2021).

Ultimately, focusing the research on these issues left much of the story missing. For example, it did not leave much room for linking internal characteristics of the individual to specific properties of the external stimuli, nor for explaining the ongoing aspect of the process: what happens when the drive is fulfilled? Can one really reach satiation with information, as with primary needs such as food or sleep? Importantly, White critiqued the drive approach for being unable to explain exploratory and playful behaviours, which do not appear to stem for a primary need nor lead to satiation with the consumption of the information (White, 1959).

Some of these drive theories did focus less on the reward-based perspective though, and brought new perspectives on how to characterise the information content of the stimuli associated with the drive, and not merely their valence. Such theories help to build stronger predictions about information-driven behaviours, for example which of several sources of information a learner decides to sample when faced with several options. We come back to those in the following section 1.2.

1.1.1.3. Evidence that information itself is rewarding in the brain

Both types of reward-based theories stick to a valence-oriented, unidirectional approach: on the one hand, drive theories largely restrict information seeking behaviours to being a reflex for fulfilling an internal need, and on the other hand, former conditioning accounts limit them to a reaction to a stimulus depending on its valence. Still, both ideas proved very helpful in shedding light on the reward-related aspect of information seeking behaviours.

Numerous works in various species provided evidence that individuals are driven by information-based stimuli and find reward in seeking such stimuli, be it to avoid an aversive state or reach an appetitive one (e.g. Dashiell, 1925; or Nissen, 1930 with rats; Butler, 1953 with monkeys; Hebb, 1958 with humans; or even Darchen, 1957 with cokroaches). This notion is backed by the more recent accumulation of evidence from neuroscience studies for an activation of the dopaminergic reward system, a group of brain regions involved in the valuation of material gains, in the case of information-related gains (Gottlieb et al., 2016).

In an important fMRI study with humans, Kang et al. presented participants with trivia questions for which they had to rate how curious they were to know the answer during an anticipatory period, before the answer was revealed. They found that during the anticipatory period, regions

of the reward network were activated: the left caudate nucleus, bilateral inferior frontal gyrus, and loci in the putamen and globus pallidus. Additionally, their behavioural data showed that participants were willing to trade money to get the answers to the trivia questions about which they were the most curious, proving that information could be valued against material gains, one of the most typical tokens for rewards (Kang et al., 2009). More fMRI studies since then also showed evidence for the activation of classical reward and valuation-related regions in anticipation of information in another trivia paradigm (Gruber et al., 2014) or cued monetary lottery tasks (Charpentier et al., 2018; Kobayashi & Hsu, 2019).

Two experiments with monkeys expanded on these findings, linking the reward-network activations to information-related cellular activations within the dopaminergic pathway (Bromberg-Martin & Hikosaka, 2009; Blanchard et al., 2015). Indeed, these studies showed that when they anticipated incoming visual information, monkeys' cells activated both in the midbrain's efferent dopaminergic pathway, and in the orbitofrontal area of the pre-frontal cortex, which receives dopaminergic afferences. Moreover, similarly to humans' willingness to trade money for information, monkeys were also willing to trade a valued material reward, juice, against information.

Taken together, these findings support the idea that in primates, including humans, the same reward system signals both the value of material rewards and information-related rewards, and that information itself can be rewarding and drive active learning behaviours.

1.1.2. Information-content based theories

A second wave of theories, actually starting with some of the drive theorists mentioned earlier and continuing into contemporary research, shifted the focus from the response to the stimulus onto the information-content itself. These theories seek to address the question of what makes an information suitable to fulfil an individual's drive for information, in other words, how the combination between information properties and the learner's state defines an optimal information-content.

1.1.2.1. Medium-level theories

1.1.2.1.1. Optimal arousal

One of the first attempts to explain information-seeking behaviours by linking the internal state of the agent with the properties of the stimuli, came from the intermediate arousal theory evoked earlier (Hebb, 1958). Hebb's work on sensory deprivation led him to believe that individuals do not merely respond to stimulation but come to seek or avoid it themselves, depending on their current state of arousal and in relation to their own ideal homeostatic level of arousal. Arousal homeostasis could be achieved either by approaching a source of stimulation in case of low arousal, such as sensory deprivation, or by avoiding it if the individual is already over-aroused. This idea implies that the same stimulation can be either positively or negatively valued depending on the state of the agent, which is an important point for explaining individual differences in response to the same stimulation. It also entails that stimulation intensity is an important factor, and that all types of stimulation will not necessarily satisfy the return to homeostasis. In this sense, Hebb's theory distinguished itself from other standard drive theories because it did not involve the reduction of a tension that always pulls in the same direction, but the attainment of an ideal, middle-ground state. However, Hebb's theory did not offer much indication as to how individuals choose between stimulations of similar intensities but different information content, and what the determinants for individuals' optimal arousal level were.

1.1.2.1.2. Optimal incongruity

A similar idea of a middle-ground target was developed by Hunt, who proposed that individuals seek information of optimal incongruity (Hunt, 1963, 1965). His approach of child development through stages, building on Piaget's theory of developmental stages (Piaget, 1952), led him to believe that seeking optimally incongruous events is essential for child development: it enables children to process relevant information that they can compare to the standard, the average stimulus that they have encountered so far. Optimally informative stimulation enables children to gradually build a repertoire of appropriate responses stage after stage, moving from a basic orienting reflex at birth to more complex approach actions in later developmental stages.

1.1.2.1.3. Moderate discrepancy

Hunt's concept of optimal incongruity was later refined into a moderate discrepancy theory (McCall & McGhee, 1977). While incongruity lies in a comparison with a standard, discrepancy is linked to a comparison with the individual's expectations. This is a slightly more powerful approach since it allows individuals to form expectations about new items by combining information from a variety of relevant past experiences, without having to first establish the standard based on similar items only. This gives the theory more room for explaining how individuals engage with novel objects. Another similar theory is Kagan's idea that learners aim to reduce their uncertainty, which is also defined as the comparison between their expectations and sensory input (Kagan, 1972). In both approaches, individuals end up forming expectations closer to the reality that they sample.

1.1.2.1.4. Intermediate collatives

Finally, one last take on this topic which can be grouped with the former accounts is Berlyne's idea that individuals seek information of intermediate level in terms of "*collative variables*": "*a group of stimulus properties to which we commonly refer by such words as "novelty", "change", "surprisingness", "incongruity", "complexity", "ambiguity", and "indistinctiveness""* (Berlyne, 1965, p. 245). Berlyne's approach is interesting because it admits that there are in fact a lot of ways to express the informativeness of an item.

1.1.2.1.5. Conclusion on medium-level theories

All in all, the theories mentioned here helped to identify what could characterise information that drives interest or not. However, they present the inconvenience of leaving the definition of what is intermediate quite open to interpretation. Also, establishing this middle-ground level requires a comparison relative to elements of reference, which cannot always be achieved.

1.1.2.2. Information-gap theories

Following this last idea, and in an effort to reconcile former accounts, Loewenstein formulated a knowledge-gap theory of information seeking (Loewenstein, 1994). He proposed that individuals are mainly driven towards information which fills in a *gap* in their knowledge. This gap can be due to incongruent, discrepant, novel, surprising or information of any type, as long as the learner identifies a missing piece in their representation of the world. Lowenstein originally proposed

that people are driven to close this knowledge gap in an attempt to avoid the aversive feeling that it bears, however it has also been proposed to come from a positive drive to experience the pleasure of closing the gap (Litman, 2005). Information-gap theories are close but deviate from discrepancy theories in that the learner does not have to form expectations for the gap to exist, which has the potential to explain how they interact with a bigger range of information. However, learners still need prior context to approach information: they need to hold knowledge on the topic beforehand so that they can even identify the gap. This does not fully explain how people get interested in new topics. Indeed, when they engage in exploratory behaviours and move away from known tasks, learners' uncertainty actually increases with the acquisition of new partial knowledge with new information gaps (Gottlieb et al., 2016).

1.1.2.3. Learning-progress theories

The most recent accounts on how individuals seek information suggest that they are in fact driven by their learning progress itself: they interact with information that they expect to maximise learning progress (Kaplan & Oudeyer, 2007b; Schmidhuber, 1991). This recent theory borrows from older accounts that proposed a drive for competence or mastery (White, 1959; Deci & Ryan, 1985), as well as Csikszentmihalyi's theory of flow (Csikszentmihalyi, 1990) which supposes that people are driven by intermediate challenge. Indeed, in order to maximise learning progress, the difficulty should rather be intermediate, such that it is neither too easy and leads to no further progress compared to the level already achieved, nor too hard and leaves the agent overwhelmed and incapable of any progress. This point recalls former medium-level accounts, on which this new learning-progress theory adds.

One of the main additions to medium-level accounts is the idea of a positive feedback loop between learning and information seeking (see Figure 1.1). Former accounts did not explain how the learning process could be sustained over an extended period of time. Because they presupposed a need or a gap to fulfil, they also implied that the action of gathering information mainly reduced this state, rendering the drive for gathering information less intense. However, individuals are shown to engage in learning activities over long periods of time. In the learning-progress framework, this is predicted by the fact that optimal activities foster learning progress, which remains motivating over time, regardless of how much information has been sampled. If more information allows for steeper progress, the drive can even increase with more information.

Inversely, when subjects misidentified ambiguous information for content from which they expected to learn, the absence of progress over time releases the drive.

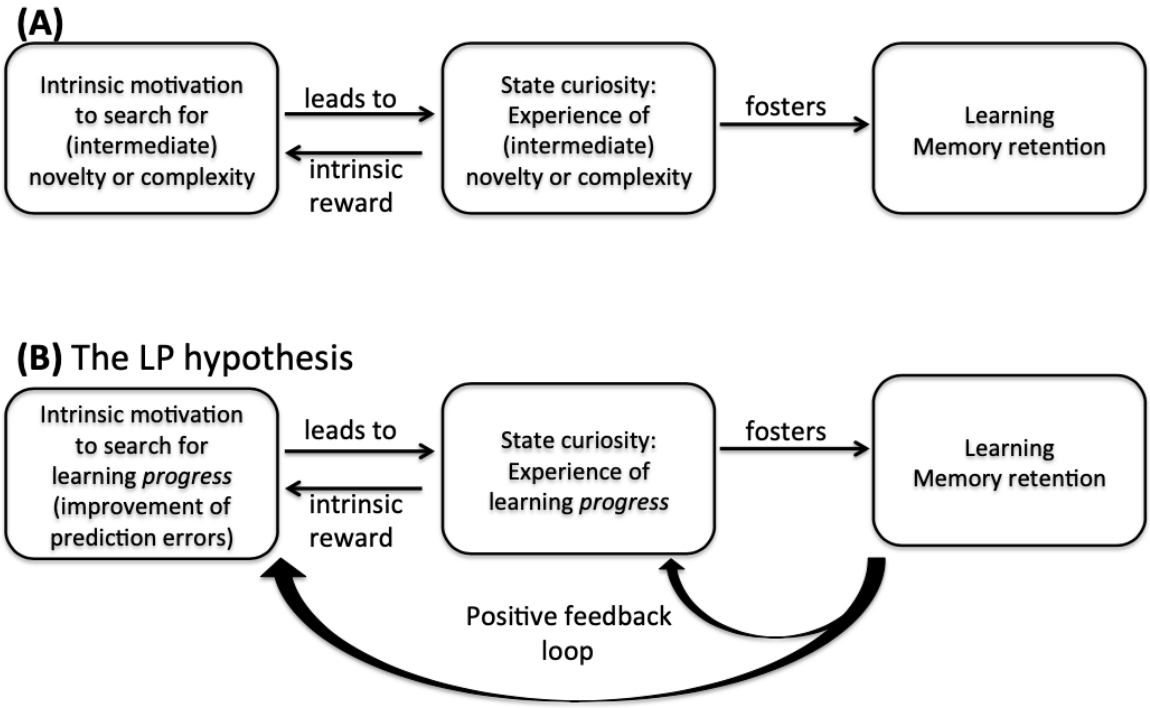


Figure 1.1: comparing the learning progress (LP) hypothesis (B) with other accounts of a drive for intermediate novelty or complexity (A); taken from Oudeyer et al. (2016).

This is particularly powerful for explaining how individuals explore in open-ended environments, where the possibilities are numerous and their potential for progress not always easy to identify before engaging with the information. This learning-progress drive enables agents to explore their options without staying stuck at extremes that happen to be either too easy or too hard after they started to engage. Furthermore, it explains how individuals hierarchise information and learn in an incremental manner. For example, having learned how to control actions involving one muscle decreases the associated learning progress, but it also opens the possibility to learn more complex actions combining muscles, which the learner can now move on to learn, without having to drop out altogether once the first step is achieved.

In an embodied robotic experiment, Oudeyer et al. implemented these principles into an algorithm controlling the actions of a robot placed on an infant's playground (Oudeyer et al., 2005). The robot explored the playground and received feedback from another robot, in a pattern that closely resembled human infants' own exploratory patterns and developmental course. They

found that the robot "*first explores simple activities, and then progressively shifts to more complex learning experiences, effectively self-generating and self-organizing a learning curriculum adapted to the current constraints of the learning system, and at the same time constraining learning and shaping the developmental trajectory*" (Oudeyer et al., 2005). Moreover, Oudeyer at al. repeated this experiment several times, starting with a clean algorithm given the same functions and starting point each time. They showed that there was a variability in these general patterns, due to the stochasticity of the process, and that the range of variability resembled the range of individual differences in infant development. However, this hypothesis was not directly tested with infants. Additionally, in a recent experiment with human adults, Ten at al. showed that participants' likelihood to select a game over another when freely playing on a platform was dependent on their learning progress (Ten et al., 2021).

1.1.3. A general word on active learning theories

Through the first section of this general introduction, the main theories of active learning have been listed and discussed in relation to each other. Strengths and weaknesses were identified, but it is worth noting that these theories largely do not contradict each other. For example, the latest learning-progress theory was tested using robots which were actually implemented with a type of reinforcement-learning algorithm (Oudeyer et al., 2005; Sutton & Barto, 1998), based on the early conditioning accounts presented at the start of this section. The learning-progress take on this classical framework does step away from the typical implementations of the past though, in the sense that Oudeyer's reinforcement-learning algorithm is intrinsically motivated by the agent's progress itself, rather than by external rewards or punishments as in classical implementations (Chentanez et al., 2005). Similarly, the learning-progress theory's advocacy for intermediate difficulty is reminiscent of medium-level theories of intermediate discrepancy, incongruity, novelty, or complexity.

Clearly, there is often a large overlap between the different theories of active learning, and remaining open questions concern less identifying which one of them is correct but rather which one offers the most complete account for the range of behaviours presented by individuals when they interact with information. Moreover, individual factors that define the optimal, medium-level ground for each agent, be it in terms of arousal, discrepancy, or learning-progress, remain to be identified, as well as how they vary between contexts and individuals. The question of how

individual factors influence learning strategies is central to the present thesis and will be explored from different angles, both with adults and infants. Experimental settings and questions are presented at the end of this general introduction (section 1.4).

1.2. Basic bricks for active learning

The theories mentioned in the first section of this chapter offer explanations for how individuals actively engage with information. A second step in understanding active learning is to apprehend how individuals going through this process integrate and learn from the information engaged with. Learning involves the processing and encoding of current experience, depending on relevant past experience (Radulescu et al., 2021). Thus, in order to learn from their environment, individuals rely on core cognitive functions for the selection and manipulation of pertinent past or current information. Specifically for learning, it has been proposed that these core functions are memory and attention (Radulescu et al., 2021), two central brain functions often considered as a dual tool for cognition (e.g. Chun & Turk-Browne, 2007; Mackay, 1973; Norman, 1968; Parr & Friston, 2017).

Cognition is the brain's ability to process information, which is highly related to learning, but also encompasses other related processes such as perception, without necessitating that learning should occur (American Psychological Association, 2022). There are other taxonomies for what represents a basic tool set for cognition, of which two dominant views are the executive functions and temperament frameworks. Such classifications, although slightly more general, can also help to understand what the foundations of learning processes are. In this section, the role of memory and attention in learning will be presented first, before moving on to reviewing the relevance of executive functions and temperamental factors. Finally, a general discussion of these components will summarise the key building blocks of cognition that underlie active learning.

1.2.1. Memory

Memory can be defined as "*the retention over time of experience-dependent internal representations, or of the capacity to reactivate or reconstruct such representations*" (Dudai, 2007), and a representation in memory is itself described as a "*neuronally encoded structured version of the world that could potentially guide behavior*" (Dudai, 1989). In short, memory involves both the encoding and the retrieval of information in the brain.

Tentative work to link memory to a specific brain region first identified the hippocampus as the home to the brain's memories. Yet, despite an established central role of the hippocampus (see

Andersen et al., 2006 for a review), memory actually appears to involve a more diffuse network with complex activation patterns of cells in multiple brain areas which can virtually include neurons in any brain region and are much more distributed than was first envisioned (Josselyn & Tonegawa, 2020). The information encoded and retrieved through memory is thought to be maintained in distributed networks of brain cells, termed engrams (Semon, 1904) or cell assemblies (Hebb, 1949). Such assemblies are formed and maintained by the consolidation of connections between neurons that are often activated together: this represents a way to encode the information relevant to their co-activation.

Molecular mechanisms through which such connections are formed have been studied into depth and are currently rather well-known (Josselyn & Frankland, 2018). The most prominent mechanism is long-term-potentiation: when a neuron receives a strong stimulation, it fires intensely and activates cascade reactions at its synapse and in the synaptic button of the neuron at the receiving end of the connection; lasting chemical modifications to the synapse ensue, such as a reinforced expression of receptors, which durably strengthens the connection between the pre- and post-synaptic neurons. Other mechanisms such as long-term-depression have been described that allow to extinguish unfit neuronal connections depending on neurons' activation patterns. They offer a detailed account for molecular- and cellular-level mechanisms of how engrams of connected neurons can be formed, modified, or erased to maintain and update memories.

Such memory processes are keystones for learning (e.g. Anderson, 1995; Eysenck, 1981; Schwartz & Reisberg, 1991). Indeed, memory plays a multi-faceted role in the learning process: in order to learn from their environment, individuals have to 1) select relevant features of the current experience to store into working memory and 2) organise and encode them coherently in long-term memory, updating their old representations accordingly, but also 3) retrieve past information from their long-term memory that is relevant to the current experience (Radulescu et al., 2021). Stored information that is used to enhance the processing of current information has been referred to as a prior in the computational sphere (Bernardo, 1979). Such priors are built from information retrieved from past experiences relevant to the current one, but also information accumulating from the current experience. They help to process the information from the current scene and bias the way individuals sample further information (Võ & Wolfe,

2015). Learning and memory thus appear as tightly intertwined processes, each influencing but also supporting the other through priors. Still, a distinction can be drawn in that memory represents the encoding and retrieval of information in the brain, while learning can be considered the process of integrating information to be encoded or retrieved, depending on what the learner understands of the current situation and defines as relevant (Radulescu et al., 2021).

1.2.2. Attention

Attention relates to the brain's ability to select and filter information. One of the most widely accepted views on attention is Posner and Petersen's three-system framework comprising an alerting, spatial orienting, and executive attention system (Posner & Petersen, 1990). Each system is considered a separate entity linked to different brain regions, although the three systems typically work together to achieve optimal attention in day-to-day situations. The alerting system is linked to changes in arousal with sudden information input, such as reacting to the apparition of an unexpected cue. Orienting refers to the ability to shift the focus of attention from one type of information to another, with or without associated movements of the eyes, head, or body. Executive attention is related to the resolution of conflicts between competing information with regards to one's goals. Each system has been associated with distinct brain regions and neurotransmitters: noradrenaline, arousal and vigilance regions of the brain stem and the right hemisphere for the alerting system; acetylcholine and parietal regions for the orienting system; and dopamine, midline frontal and anterior cingulate cortical regions for the executive system (Petersen & Posner, 2012).

Although distinct systems and regions were characterised, attention is also considered to function in par with sensory processes, and for this reason connections between classical attentional regions and sensory cortices are central for functional attentional processes to take place. This is attested by work with populations with impaired attention such as attention deficit hyperactivity disorder (ADHD: Fair et al., 2010; Liddle et al., 2011) and autism spectrum disorder (ASD: Hahamy et al., 2015; Ray et al., 2014). These studies have identified atypical short-range connections within sensory cortices as well as long-range fronto-sensory connections as a key marker of attentional disorders (Amso & Scerif, 2015). Attention is thus supported by the activation of a distributed network, somewhat similar to memory (Posner & Rothbart, 2007). At the neural level, the main mechanism proposed for supporting attentional processes is the synchronisation of

connected cells (Knudsen, 2007; Lakatos et al., 2008). Synchronised activity favours the transmission of the information carried by the synchronised cells over potentially competing information held in cells out of synchrony, which, in essence, provides a way to select information for attention.

The selection of information through attentional processes is crucial for learning. The learning strategies evoked earlier define individuals' goals and strategies for sampling information, which then have to be selected appropriately through attentional processes. While attention refers to the process of selecting relevant information, learning can be considered separately as a process of integration of the information once sampled, such that lasting knowledge can be acquired and memorised.

1.2.3. Executive functions

Moving on to the executive functions' framework, other basic functions that serve learning processes can be identified. They present significant overlap with the previous concepts of memory and attention but offer a different taxonomy. Executive functions refer to "*general-purpose control mechanisms that modulate the operation of various cognitive subprocesses and thereby regulate the dynamics of human cognition*" (Miyake et al., 2000). They "*facilitate new ways of behaving, and optimise one's approach to unfamiliar circumstances*" (Gilbert & Burgess, 2008). The frontal lobe, and in particular the prefrontal cortex, is considered to be the site of much of the executive functions processes (Miyake et al., 2000). However, cross-region connections between frontal and parietal regions have also been shown to be central (Hwang et al., 2010) and a network between frontal and parietal regions has been proposed to be implicated (Niendam et al., 2012). Executive functions have been consistently associated with learning outcomes, such as academic and professional achievement (e.g. Bull & Scerif, 2001; Blair & Razza, 2007; Morrison et al., 2010; Bailey, 2007). They have also been suggested to form the keystones for a higher-level function: fluid intelligence, a higher-order reasoning and problem-solving component (Diamond, 2013), which may be approached to learning.

While executive functions were first considered as stemming from the same "central executive" system (Baddeley, 1986), nowadays the most widely accepted view is that three distinct main categories of executive functions can be distinguished: working memory, inhibition and flexibility (Diamond, 2013; Miyake et al., 2000). Measuring those different functions separately is one of

the challenges faced by the field, because although they are considered as separate functions, they are also related and more often than not working in concert for efficient cognition. This leads to the "*task impurity problem*" (Burgess, 1997) of designing tasks measuring only one function, while being representative enough of that one function. This issue plays a big part in keeping the debate open on the classification of executive functions, and alternative taxonomies are still offered which consider different functions, sometimes even defined as one common unitary function (e.g. N. P. Friedman & Miyake, 2017; Doebel, 2020). We will stick to the main framework of three separate functions, which, compared to unitary accounts, is more effective in our quest to identify a basic toolset for learning process, and is more largely used in relevant literature. However, it is helpful to keep in mind that the specifics that delineate these functions remain debated.

1.2.3.1. Working memory

Working memory is the function of "*holding information in mind and manipulating it*" (Diamond, 2013). Working memory is essential for information processes because it enables to maintain and gate information that is not perceptually present anymore. This function mainly relies on regions of the dorsolateral prefrontal cortex (D'Esposito et al., 1999).

1.2.3.2. Inhibition: self-regulation and attention focusing

Inhibition involves "*being able to control one's attention, behavior, thoughts, and/or emotions to override a strong internal predisposition or external lure, and instead do what's more appropriate or needed*" (Diamond, 2013). It is a key ability for individuals which allows them not to merely react to the environment via stereotypical impulses and reflexes, but to have more flexibility in choosing an appropriate response depending on higher-level goals. This function is a hybrid of several concepts: attentional control, cognitive inhibition and self-control (Diamond, 2013). Self-control involves the inhibition of impulsive and emotional responses, which notably enables delayed gratification, the ability to remain on a task to await for a reward rather than always looking for the short-term positive outcome. Cognitive inhibition refers to the ability to resist unwanted influences from past or current information, blocking purely reactive responses which are not aligned with the individual's goals. Attentional control is the ability to maintain one's focus on relevant information, while discarding irrelevant information. Interestingly again, this function seems to have a lot of overlap with the broader concept of attention described earlier.

1.2.3.3. Flexibility: creativity and attention shifting

Cognitive flexibility represents the ability "*to change perspectives*", "*to adjust to changed demands or priorities, to admit you were wrong, and to take advantage of sudden, unexpected opportunities*" (Diamond, 2013). It is essential for faculties such as creativity, task switching, or set switching. It appears as an active effort to overcome a natural tendency for "*attentional inertia*" (Kirkham et al., 2003), the inclination to keep focusing on the same elements that used to be relevant. In that regard, flexibility appears highly related to attentional constructs as well.

1.3. Measuring active learning in humans

So far, major theories for how individuals actively learn as well as the cognitive functions that learning processes rely on have been outlined. Finally, another core aspect of research on active learning and information sampling in humans is the methods for measuring these processes and in particular working to overcome the difficulties of disentangling this complex subject. As always, different tools exist, each with their own set of strengths and weaknesses, which can be used and even combined in different paradigms to best fit the research questions. A few key measures will be presented in separate subsections, with an emphasis on their specific advantages and disadvantages, alongside a review of the main evidence linking them to learning processes in infants and adults.

1.3.1. Engagement

1.3.1.1. Time on task and drop-out time

Measuring participants' engagement with a task or information that they are presented with, is one of the simplest approaches that can be taken to study active learning. Individuals' decisions to approach information vs. avoid it are a manifestation of their strategy for actively sampling information, and can be encompassed in measures as basic as how much time an individual is willing to remain on the task at hand. Linking a participant's drop-out time to specific events throughout the experiment can be particularly informative about the context that drove their decision to move on.

One of the predictions of the learning-progress theory introduced in the section 1.1.2.3 of this chapter is that an individual's learning progress at a given moment in time is a driver for whether they will keep engaging with the same source of information or move on. In an experiment with human adults asked to freely play for a set number of trials on a gaming platform with several games of varying difficulty, Ten at al. showed that participants' likelihood to select a game over another was dependent on their learning progress (Ten et al., 2021). In their experiment, Ten et al. asked participants what information (game) they wanted to engage with, thereby passively rejecting other information.

However, asking participants when they would like to actively stop sampling from one source of information, or monitoring when they do so, is an equally informative approach that remains unexploited in human research. In artificial intelligence, well-defined policies for when to drop-out of an activity and move on to another one have been shown to be particularly important for the establishment of efficient exploratory strategies in robots (Oudeyer et al., 2005). In human Psychology though, decisions to drop-out remain an understudied topic, generally viewed as a burden for experimental design rather than a meaningful variable.

1.3.1.2. Curiosity for trivia questions

Another way to look at how participants engage with information is to measure how curious they are about a given piece of information. Recently, a new body of research in the field of curiosity made use of trivia question paradigms to instigate such curiosity states in participants (e.g. Kang et al., 2009; Baranes et al., 2015; Fastrich et al., 2018). In these studies, participants are presented with a trivia question and asked to rate how curious they feel about knowing the answer, before the answer is revealed. Work using this approach has shown that participants value the information contained in the answers, enough to trade monetary goods for it to be revealed, and that this valuation is reflected by brain activations of reward-related areas. Trivia questions appear as a powerful instrument to investigate the anticipation period before obtaining information and quantify people's willingness to work for it. However, they can also be seen as a simplistic tool in the sense that they do not offer a possibility for exploration. Indeed, the information is fully contained in the answer, and participants' choice to access it is binary: yes or no. In their day-to-day lives though, humans evolve in a complex environment in which information sources are numerous and competing, such that they have to make frequent decisions about which information to sample and how to explore their surroundings. Other paradigms that offer the possibility to sample one of several options are more appropriate for questions related to this type of behaviour. Moreover, trivia questions require a certain language level that is not suitable for all populations: infants and atypical groups cannot necessarily undergo such paradigms for example, and implicit paradigms can be more suited for them.

1.3.2. Visual exploration

Another possibility for the study of individuals' active sampling of information is to investigate their gaze patterns. Indeed, eye movements represent humans' primary means to sample

information and have long been known to be influenced by cognitive processes. Famous experiments from Yarbus (1956) showed that participants' visual exploration is methodical and directed at relevant information: when looking at a painting, people specifically fixate informative locations such as faces and objects, leaving out uninformative areas. Moreover, changes in instructions bias exploration and lead to different task-adapted fixation patterns. Indeed, eye movements are guided not only by stimuli's characteristics, but also by internal goals (Foulsham, 2015; Tatler et al., 2011). Yarbus' tracking of eye movements relied on an ingenious system with a mirror attached to subjects' eyeball. Today's corneal reflection eye-tracking methods provide less intrusive and more powerful instruments to precisely measure the location and time-course of individuals' fixations and saccades (Hayhoe & Ballard, 2005).

Eye movements recorded with eye-tracking methods have been shown to be not only dependent on stimuli and goals, but also on learning. For example, studies of eye movements coordinated with complex hand movements have shown that gaze helps the hand movements by preceding them. Participants fixate relevant locations where the hand is about to move, which informs the motor system about the upcoming movement, rather than fixating the hand itself, a less informative visual location (Land, 2006). This shows again that eye movements and cognition are tightly linked (Johansson et al., 2001). Moreover, these fixation patterns were shown to change as the participants learned to perform the hand movements: they started to skip fixations at check-points linked to specific action steps when they had learned to execute the related action, and move straight to the next point of interest (Sailer et al., 2005).

Finally, eye movements have been shown to depend on individual differences in trait curiosity. In a free-viewing paradigm of complex scenes, Risko et al. showed that participants' number of fixations and regions explored was positively correlated with their trait curiosity in a personality questionnaire (Risko et al., 2012).

All in all, eye movements thus appear as a measure of choice to track individuals' active sampling of information and learning. The non-intrusive character of current eye-tracking systems makes them particularly well-suited for infants as well as adults. Moreover, eye-movements can be recorded with a wide variety of paradigms depending on the research questions at hand. Those paradigms can be made entirely implicit and independent of language for non-verbal populations such as infants.

1.3.3. Multisensory exploration and play

One limit of eye-tracking methods is that they only consider exploration through one sense: vision. Yet, other senses contribute to how humans sample information in their day-to-day life, making natural exploration holistic and multisensorial. While accounting for all the possible ways in which humans explore is an overwhelmingly difficult task, paradigms exist that allow for more of this complexity to be investigated compared to visual exploration paradigms. Exploratory play paradigms, also termed free play, represent one important alternative. During exploratory play, infants or children are presented with toys that they are left free to explore as they like, not only visually but also through manual, buccal, olfactory, or auditory examination. Here again, accounting for all types of explorations in experimental measures of play is arduous, but the exploration itself is left open and multisensorial, regardless of the measures defined by the researchers, and thus the process is considered more naturalistic than computer-based paradigms.

The role of play in learning has been evoked by psychologists and educators for over a century (Groos, 1901; Montessori, 1912; Piaget, 1952). Schulz and Bonawitz, in an important free-play experiment, showed that pre-schoolers' exploration of toys is not only influenced by stimulus characteristics such as salience, but also by the children's knowledge and uncertainty about it (L. E. Schulz & Bonawitz, 2007). Indeed, children in their study were more likely to explore and play with toys whose functions had been presented to them in an ambiguous fashion compared to unambiguous demonstrations, and appeared to make manipulations directed at learning about unclear functions. Later studies built on these results and brought more evidence for this idea that children's exploration of toys during play is driven by learning opportunities (Cook et al., 2011; Gweon et al., 2014; Stahl & Feigenson, 2015; van Schijndel et al., 2015; Ruggeri et al., 2019; Doan et al., 2020; Siegel et al., 2021). Investigating how infants and children explore in free play settings thus appears to be another powerful tool for the study of active learning during development.

1.3.4. Brain activity

So far, possible measures for exploration and active learning have focused on how individuals gather information. This is a crucial step in the learning process but looking at how individuals

process this information is also imperative. Measures of brain activity can be particularly informative for unveiling underlying ongoing processes in the brain and shed light on how sampled information is structured and used. For example, Kang et al.'s experiment (Kang et al., 2009), evoked in section 1.1.1.3 and 1.3.1.2 of this chapter, used functional magnetic resonance imaging (fMRI) to record the brain activity while participants were presented with trivia questions. They showed that participants' valuation of information as a token worth trading money was linked to brain activity in reward areas constituting the same network that is involved with the valuation of monetary rewards indeed. Using brain imaging in this way enabled them to back their claim that information itself is valuable and rewarding to humans, something that the mere study of participants' behaviour could not prove although it hinted at it. Other imaging tools such as electroencephalography (EEG), magnetoencephalography (MEG), functional near-infrared spectroscopy (fNIRS) or even intracortical recordings of neurons in patients or animals can similarly be used to image the brain during active learning paradigms to provide information on the underlying neural processes. There are many ways in which neural activity can be linked to learning, and a full review of the literature on the topic, which concerns most of the numerous papers in brain imaging research, is beyond the scope of this thesis. Neural mechanisms thought to support the basic cognitive functions involved in learning processes were reviewed in the previous section of this chapter, and provide pointers for mechanisms particularly important, such as synaptic changes or synchronisation of neurons, which neuroimaging tools can help to expose.

1.4. Defining learning

The previous sections of this chapter laid out three main aspects of active learning that are central to its study: the principal theories determining how agents engage in active-learning behaviours, the building bricks on which these behaviours rely, and the methods through which they can be studied. These three themes form solid foundations for research on active learning in humans, in addition to which another key question now remains: what is meant by learning and which situations fit into this definition?

Learning can be defined as "*the acquisition of novel information, behaviors, or abilities after practice, observation, or other experiences, as evidenced by change in behavior, knowledge, or brain function*" (American Psychological Association, 2022). This definition highlights some key criteria for learning: 1) it involves change, 2) it endures over time, 3) it occurs through experience (Schunk, 2012). Those key points delineate what is considered learning or not but also how it can be measured. They help to exclude situations such as maturational changes, drug use or fatigue, all of which also affect the behaviour but not in an experience-driven or enduring manner. Still, distinguishing between these learning and non-learning events can be arduous as the line is often blurred. For example, temporary changes in behaviour, such as those induced by a molecule which vanish once the molecule out of the system, are not considered as learning. Yet, learning also does not have to last indefinitely as we can forget things. Moreover, learning is typically measured indirectly, via the changes that it induces. However, people have been shown to learn and acquire knowledge or skills without necessarily demonstrating them at the same time than they were learning them. Additionally, identifying the changes *"in behavior, knowledge, or brain function*" (American Psychological Association, 2022) can also be challenging. Finally, although the definition presented above presents key points that researchers and educators tend to agree constitute key points for learning, authors still concede that there is no universally accepted definition of learning (e.g., De Houwer et al., 2013, Harel & Koichu, 2010, Schunk, 2012) and different variants exist, for example questioning the validity of the experience criterion (De Houwer et al., 2013).

These different points demonstrate that learning, despite its long history of being studied, which dates back to ancient epistemology philosophers such as Plato, remains a somewhat elusive

concept (e.g., De Houwer et al., 2013, Schunk 2012). What's more, it encompasses extremely varied processes, from getting used to recognising an individual to gaining expertise on solving complex mathematical problems. However, as Schunk (2012) expressed in his book on the topic: "*commonalities often exist among different forms of learning (Bruner, 1985). Learning to read is fundamentally different from learning to play the violin, but both benefit from attention, effort, and persistence. Learning to write term papers and learning to throw the javelin may not appear to be similar, but both are promoted by goal setting, self-monitoring of progress, corrective feedback from teachers and coaches, and feelings of intrinsic motivation.*" All in all, grouping such different situations under the term *learning* appears helpful and meaningful, but it is important to keep in mind the broad variety of situations encompassed in this term, which can be approached by researchers in very different ways.

1.5. Current approach

The current PhD thesis attempts to fill in some of the gaps in the active learning literature just reviewed and to embrace some of the variety in the possible approaches, in order to build as much of a coherent and extensive picture on the topic.

In a first study presented in Chapter 2, we investigated the influence of priors on information sampling in the context of a simple object recognition task. Previous sections of this general introduction established that memory and prior experience play a fundamental role in active information sampling (sections 1.2.1, 1.2.3.1 and 1.4). This experiment sat out to evaluate the influence of priors on one of the most basic information sampling tasks that humans are constantly performing: where to look and accumulate visual information about an object in order to recognise it. We used eye-tracking technology to monitor human adults' gaze while they were looking at altered, ambiguous images of objects and manipulated stimulus informativeness as well as prior access to informative stimuli in a quantifiable manner. This allowed us to precisely evaluate participants' ability to fixate informative locations of the stimuli and recognise the objects, depending on their prior access to information.

Chapters 3 and 4 present the findings from another study with 10-month-old babies, looking at the link between infants' engagement with a stream of information and their brain activity in

response to this stimulus. Moving on to a younger population gave us the opportunity to identify basic learning mechanisms present from early on in life, when babies are still developing and learning a lot about their environment. In Chapter 3, we looked at electroencephalographic (EEG) neural correlates of how infants select and process distracting relative to relevant information. Our stimuli comprised both interesting ongoing information and distracting interruptive information, which were competing for infants' processing capacities and necessitated information selection processes. Neural synchrony has been proposed to be a major neural mechanism for attention selection (section 1.2.2), thus we hypothesised that infants' ability to select relevant information and discard distracting one would be reflected by the synchronisation of fast and slow brain rhythms, although no previous evidence existed for neural synchrony between fast and slow rhythms in infants. In Chapter 4, we aimed to link infants' behaviour and brain activity during their visit to the lab, and measures of attention and effortful control from questionnaires about how parents perceive their infants' individual traits in their daily lives.

Finally, Chapters 5 and 6 describe the results of a last study with 15-month-old infants. In Chapter 5, we tested the learning-progress hypothesis (section 1.1.2.3) in infants for the first time. Using a screen-based card-matching task, we monitored infants' gaze using eye-tracking and investigated the link between their learning progress and their willingness to remain engaged with the task. Chapter 6 linked measures from the study's four tasks together. We assessed the level of consistency in infants' exploration strategies and their relationship with individual variations in trait executive functions, and compared visual exploration measures of screen-based stimuli in similar trials from the start and end of the experiment, but also visual exploration in these trials with manual exploration during free play.

Chapter 7 contains a final discussion on these studies where the results are summarised and discussed in line with overarching themes.

Chapter 2: Priors' influence on visual information sampling in adults

2.1. Introduction

2.1.1. Human vision's efficiency for the recognition of objects

One task that humans learn to perform very fast and accurately and which they are constantly faced with is the recognition of objects in their visual environment. Humans can recognise visual objects under 150ms, as electroencephalographic (EEG) studies have shown (Thorpe et al., 1996), and can even gaze towards a detected object in less than 120ms (Kirchner & Thorpe, 2006). Object recognition in humans is not only fast, it is effective even under difficult conditions such as occlusion, variability in object size and visual angles, poor illumination or in cluttered environments (Riesenhuber & Poggio, 2000). Despite remarkable progress in the field of artificial intelligence, reaching such robust and invariant results with machines as with human vision has proven arduous, and even the latest networks remain generally less accurate than trained human observers for recognising objects in images (Shankar et al., 2020).

2.1.2. Importance of eye movements

What's more, human vision is actually much richer than artificial intelligence's simplified view of object recognition as a static analysis of images of objects: human vision is an active sampling process and involves extensive visual behaviour. Complex visual scenes like the ones that populate our daily lives are actively explored by humans through sequences of fixations during which the gaze remains on a particular location, and saccades during which the focus of gaze moves from one location to another (Findlay et al., 2003; Henderson, 2003). In fact, all animals with a developed visual system appear to rely on such saccade-fixation sequences through movements of either their eyes, their head or their body (Land, 1999). Such sequences of eye-movements are not random: when humans explore a scene, they direct their fixations to the

informative locations, such as people and objects, in a way that depends not only on the physical properties of the stimuli like local contrast or spatial frequency (Ruddock et al., 1996; Zetzsche et al., 2000), but also on the observer's goal like experimental instructions (Buswell, 1935; Yarbus, 1956; Castelhano et al., 2009) or relevant prior knowledge (Wu & Zhao, 2017).

2.1.3. The role of priors

Prior knowledge, which we will briefly call priors (Bernardo, 1979), influences observers' efficiency when they visually explore (Võ & Wolfe, 2015). Most of the work on this topic has been done in the context of visual search experiments: participants are presented with a complex stimulus, usually a scene, and are asked to find a specific object within the scene. Priors have generally been shown to aid the visual search, but they can also hinder it. For example, looking for an object within a newly learned category has been shown to take more time but produce less false alarms compared to familiar categories such as letters (Wu et al., 2013, 2016). In an influential review, Henderson (2003) defined two types of priors that guide visual search: 1) *specific* priors, acquired either over the short term, for example during an experiment, or over longer term, for example through several experiences from the observer's daily life; and 2) *generic* priors defined as "'scene-schema knowledge', *generic semantic and spatial knowledge about a particular type of scene*" which provide rules e.g. where a car is most likely to be found in an urban view.

This effect of priors does not only affect visual sampling but also conscious perception (Schwiedrzik et al., 2014). Presenting a stimulus after another stimulus can produce so-called *attractive* or *hysteresis* effects, when the perception of the second stimulus is biased towards seeing the same thing as the first stimulus, or *contrastive*, *repulsive* effects, when the perception is drawn towards perceiving something else than the first stimulus (Schwiedrzik et al., 2014; J. S. Snyder et al., 2015). Contrastive effects are well illustrated by visual illusions that occur when looking at a moving stimulus with one obvious direction, such as a waterfall: after some time, static objects near the moving objects will appear to be moving in the opposite direction, such as rocks appearing to fly near the waterfall (Adams et al., 2013). Hysteresis effects occur more often, although not exclusively, in the context of ambiguous stimuli (Brascamp et al., 2007; Fischer & Whitney, 2014), such as bi-stable Necker cubes: viewing an unambiguous cube before, e.g. face

up, will bias the perception of the ambiguous cube in the same face-up configuration as the former.

As flagged before, most of the work on the influence of priors on eye movements focused on visual search paradigms, which are known to elicit extensive sequences of fixations and saccades (Wolfe, 2020). Little is known about how priors influence eye movements in simpler tasks, such as object recognition, because humans are typically extremely skilled at recognising objects at a glance, without engaging in visual exploration through eye movements (Kirchner & Thorpe, 2006). However, when stimuli are ambiguous, observers have been shown to explore objects through extensive saccades and fixations (Kietzmann & König, 2015; Moca et al., 2011). This opens up new possibilities for the characterisation of the cognitive processes underlying object recognition: indeed, eye movements have been described as "*a window into the operation of the attentional system*" (Henderson, 2003, p. 498), and more generally of cognition (Findlay et al., 2003).

2.1.4. Investigating visual exploration for object recognition: the "Dots" method

One method that is successful in eliciting extensive visual exploration of objects is the "*Dots*" method (Moca et al., 2011; Suzuki et al., 2018), which also has the unique advantage of enabling the precise quantification of information sampled at each fixation location. In this method, stimuli are made of lattices of dots incrementally deformed to let the contours of objects appear more or less visibly. To recognise objects in dots stimuli, subjects have to actively explore, relying on Gestalts, such as proximity, grouping, and good continuation (Kubovy & Wagemans, 1995) and appear to make inferences relying on priors (Feldman, 1993, 1997, 2001).

2.1.4.1. Quantifying information content

The lattice deformation force is determined for each dot by a global visibility level *g* and by the local contour density of the original image: following a physically realistic elastic force, dots are pulled from their original position towards contour-dense regions in a controlled, quantifiable manner. Manipulating the visibility level *g* enables to quantify the overall information content of the stimuli. Moreover, this global information content can be contrasted with the local information accessed by participants: for each fixation location, the local dot displacement (LDD) in the stimulus and the local contour density (LCD) in the source image can be extracted. The contour density is used to calculate the force the displaces the dots, thus these two measures are

highly related. However, they do not fully map onto each other as it is the squared contour density that drives the displacement. Consequently, lower amounts of LCD will lead to modest LDD compared to high values of LCD's important LDD. Moreover, while the same point in a stimulus made from one given object will have varying LDD values depending on the *g* value, its LCD value will remain constant across *g* levels, hence comparing these two values can help to understand how *physical* LDD information conveys or not the *hidden* meaningful LCD information to the participants.

2.1.4.2. Previous results

The *Dots* method was introduced by Moca et al. (2011), who presented human adult participants with *Dots* images of fifty objects and seven visibility levels *g* (0-0.3), in seven blocks of set visibility. Two groups were compared: *Ascending* participants, who viewed the stimuli in order of ascending visibility *g*, and *Descending* participants, who followed the opposite order, from high to low *g*. This manipulation of visibility order was used to influence subjects' prior experience with the set of objects. They could either build strong priors from the beginning of the experiment (*Descending* group) or be unable to build priors through the first blocks (*Ascending* group). One of the main findings from this study was the evidence of a *hysteresis* effect (Schwiedrzik et al., 2014): prior experience with stimuli robustly influenced participants' ability to recognise stimuli, which was reflected in their visual exploration patterns. Indeed, all participants recognised objects more easily with increasing visibility, but *Descending* participants' recognition threshold was lower than for *Ascending* participants (see Figure 2.1, A). *Descending* participants also explored more informative locations (more local contour density, LCD) in ambiguous stimuli compared to *Ascending* participants (see Figure 2.1, B).

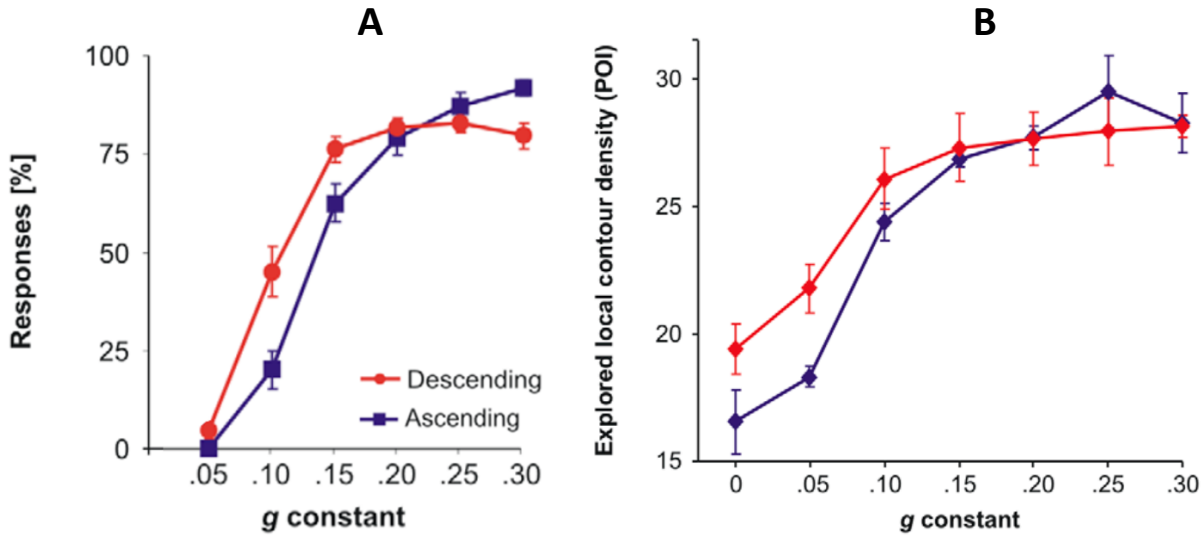


Figure 2.1: *Ascending* (blue) and *Descending* (red) participants' correct recognition of objects (**A**) and explored local contour density (LCD, **B**) as a function of visibility (*g*). Error bars represent the standard error of the mean (S.E.M.). Adapted with permission from Moca et al. (2011).

This method was adopted by another laboratory (Suzuki et al., 2018), however they did not investigate visual exploration but pupil dilation. They showed that pupil dilation could be used as a marker for implicit object recognition, using videos of *Dots* stimuli morphing from one visibility level to another.

2.1.5. Current study

2.1.5.1. Ascending, Descending and Random groups

Here, we extended Moca et al.'s paradigm with a third group of participants, for whom the stimuli appeared in a random order of visibility *g* (*Random* group). This group's presentation order granted them an intermediate access to information from the start of the experiment, in-between the *Ascending* and the *Descending* participants' access to information, allowing them to build partial priors from the start. Thus, we predicted that *Random* participants' performance would therefore lie between the *Ascending* and *Descending* participants' performance. To note, each object was presented once per block, as for the *Ascending* and *Descending* participants, and only their *g* level was randomised such that each *g* level was viewed once throughout the experiment. The data from the three groups, two of which were presented in Moca et al. (2011), were entirely

(re-)analysed, through a manual check of all the fixations that Moca et al. had automatically identified, in order to validate these results with a thorough pre-processing pipeline.

2.1.5.2. Object-specific and task-general priors

Following Henderson's typology of *specific* vs. *generic* priors in visual scene exploration (see section 2.1.3 above), we expected participants to build two types of priors for object recognition: 1) *object-specific* priors pertaining to the particular objects of the stimulus-set and that participants build as they recognise the objects; and 2) *task-general* priors related to participants' accumulated knowledge about the task and the best ways to explore *Dots* stimuli, regardless of recognition. To note, object-specific priors combine short-term priors accumulated during the task, controlled through experimental design, and long-term priors from participants past experience in their daily life, which cannot be controlled or accounted for in the context of this study. Their contribution is expected to be modest, restricted by Moca et al.'s (2011) extensive piloting of the stimulus-set, and evening out across participants.

2.1.5.3. Model: linking visibility, priors and performance

To better account for the group differences, we built a simple model (Figure 2.2) to describe the priors (middle two rows, respectively object-specific and task-general) and predicted performance (bottom) of each group, according to the current availability of information or visibility level *g* (top). We simplified the model by using 3 generic levels of visibility: low (Left; *g* = 0.00-0.05), medium (Middle; *g* = 0.10-0.20) and high (Bottom; *g* = 0.25-0.30). The model also includes a theoretical group of naïve participants (in grey) i.e., not building any priors, as a point of reference.

2.1.5.3.1. Performance

We predicted that all the groups would perform better as visibility increases, as reflected by overall increasing performance with *g* for all groups. However, we expected each group's performance relative to the other groups to depend on the strength of their priors: *Descending* participants were expected to dominate at low and intermediate *g*, while *Ascending* participants would dominate at high *g*, and *Random* participant remain at an intermediate level overall. We expected differences in performance to manifest both in terms of recognition accuracy and visual exploration, as in Moca et al. (2011).

2.1.5.3.2. Priors

The three groups of participants were expected to build task-general priors equally well along the study, mostly acquiring them at the start of the study regardless of *g* level, as reflected by symmetrical changes at low and high *g* levels. On the contrary, object-specific priors were expected to depend on *g*: they could only be built starting from medium visibility, when participants begun to recognise objects, and only to a weak extent (as reflected by *Ascending* participants' weak priors at middle *g*) while they could be fully acquired from one presentation at high visibility (as reflected by *Descending* participants' strong priors already at middle *g*). Because *Random* participants do not view stimuli in order of *g*, their prior strength is expected to be intermediate throughout *g* levels.

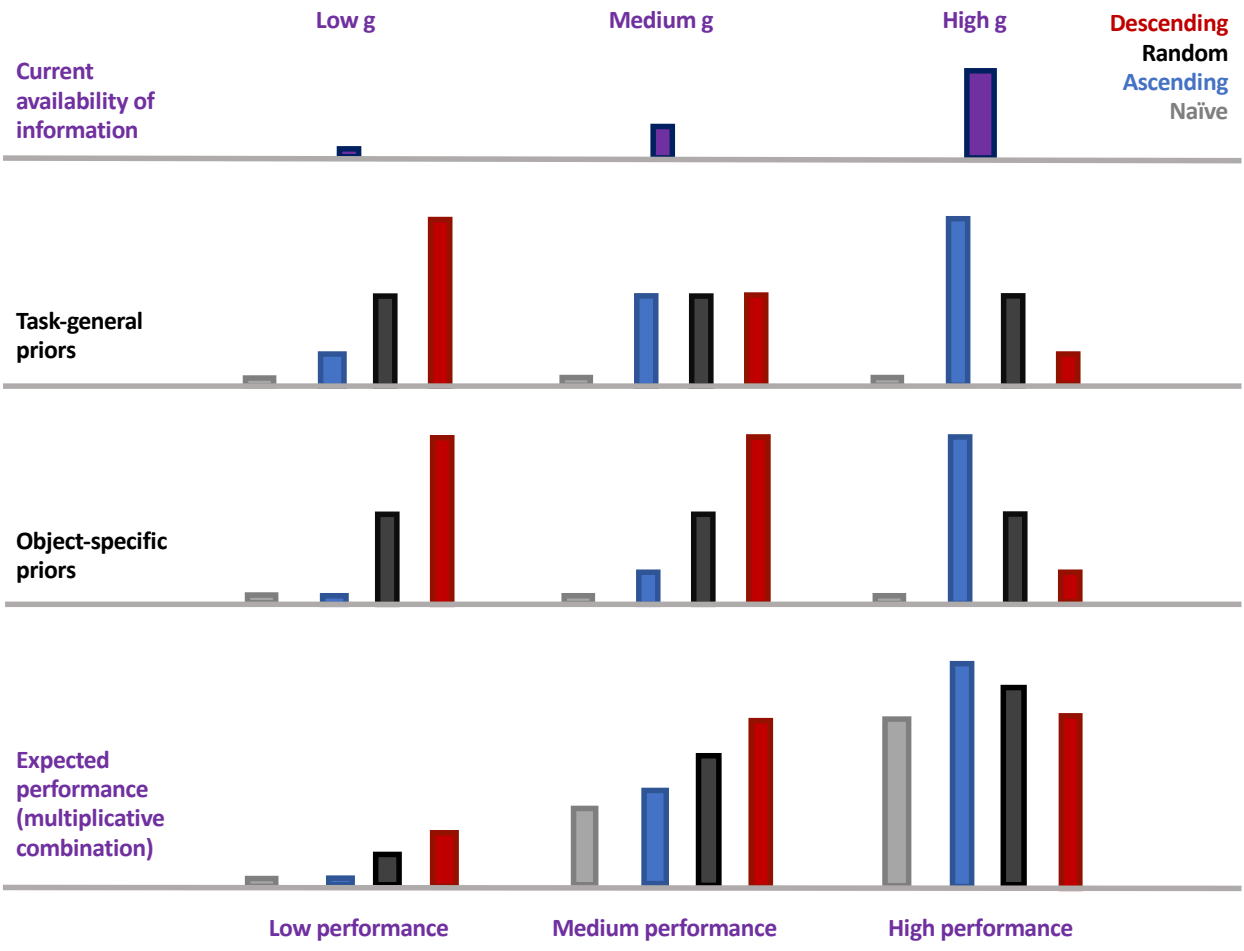


Figure 2.2: Model of participants' access to information *g* (top row), task-general (middle-top) and object-specific (middle-low) priors, as well as expected performance (bottom),

for low (left column), medium (middle column) and high *g* levels (right column). *Descending* participants are depicted in red, *Ascending* ones in blue and *Random* ones in black. Grey bars correspond to a hypothetical group of naïve participants building no priors. Higher bars signify higher values.

2.2. Methods

This study is based on work from Moca et al. (2011). They recorded all the data and analysed two of the three groups of participants (the Ascending and Descending groups, leaving the Random group's data unexplored). Here, I entirely re-analysed the data reported by Moca et al. (see methods section 2.2.5.1) and included the analysis of the yet unreported Random group.

2.2.1. Participants

18 subjects (10 females, mean age 28.3 years, S.D. 4.4 years) took part in this experiment. They either joined as volunteers or received course credits for their contribution as part of their undergraduate Psychology curriculum. They all had normal or corrected-to-normal vision. They each were assigned to one of three experimental conditions (described below and referred to as *Ascending*, *Descending*, or *Random*), resulting in three groups of N=6 subjects each. All participants gave their written consent before starting the experiment. All procedures were approved by the local ethics committee of the University of Medicine and Pharmacy "Iuliu Hatieganu" of Cluj-Napoca, Romania, under the approval No. 150/10.12.2009.

2.2.2. Stimuli

We used Moca et al.'s (2011) stimuli: they were generated via the *Dots* methods, which helps to slow down subjects' recognition of visual objects, while precisely controlling the visual information to which they are given access. Using contour information extracted from source images of objects, a deformation was applied to a 2D lattice of dots to displace dots and reveal objects' outline. The visibility of the stimuli and their information content was precisely manipulated through the use of 7 different deformation levels controlled by a gravitation constant g. It ranged from $g = 0.00$ at the lowest visibility level (no deformation) to $g = 0.30$ at the highest visibility level, with steps of $g = 0.05$. Although both the identity and the visibility of the objects varied between stimuli, this method ensured that all stimuli were made of the same local elements (dots) and shared similar physical properties across the set, facilitating the comparison of stimuli. For each of the 50 objects, 7 stimuli at each of the 7 *g-levels* were generated, resulting in a pool of 350 stimuli. Example stimuli are shown in Figure 2.3.

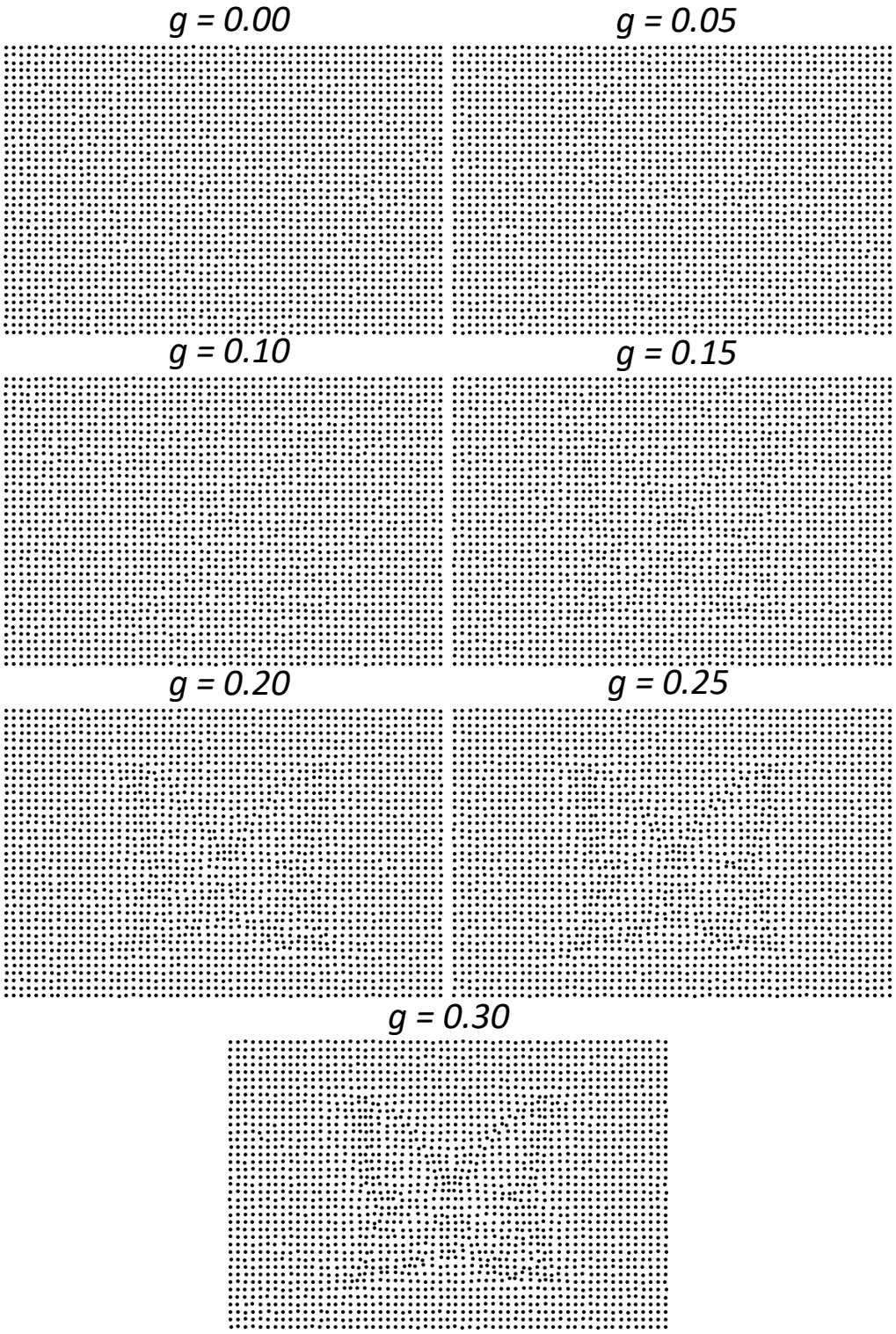


Figure 2.3: Example of *Dots* stimuli for one object at each level of visibility *g*, shown in order of increasing *g*.

2.2.3. Procedure

Participants were tested in a detection-recognition task. Stimuli were presented on a 22" Samsung SyncMaster 226BW (2ms Grey To Grey response time) at a resolution of 1,680x1,050pixels, placed at a distance of 115cm from the participant. The images were presented in the central part of the screen at a resolution of 600x400pixels. Each trial started with a red fixation cross for a random duration of 1,000 to 1,500ms. The stimulus then appeared on the screen for an indefinite duration. Participants were instructed to visually explore each stimulus for as long as they wanted to and to decide whether the dots pattern of the stimulus represented anything meaningful. After visualising the stimulus and when they wanted to, they pressed one of three buttons to indicate whether they had "*seen*" the object ("L" key: they perceived something meaningful in the stimulus and knew what it was), were "*uncertain*" what the object was ("S" key: they thought they saw something but were not sure what it was), or saw "*nothing*" in the dots grid ("A" key: they did not think anything meaningful was depicted). This was followed by a green fixation cross for 500ms, after which a message appeared asking the subject to explicitly name the object that they (thought they) had seen, in the cases when they answered "*seen*" or "*uncertain*" (guess). Their oral answers were manually recorded by an experimenter present in the room throughout the experiment. Participants finally pressed SPACE to move on to the next trial, which began after a 200ms delay. The session started with a 14-trial training block using a separate set of objects, followed by the experiment's blocks. The general design for the trials is summarised in Figure 2.4.

2.2.3.1. Presentation order – between-subjects design

Stimuli were presented in 7 blocks, each containing all 50 objects at one of the 7 visibility levels. The order of the objects within the blocks was randomised for each participant. The order of the objects' visibility throughout blocks varied between the groups to manipulate the participants' access to information. One group of participants saw the stimuli in an *Ascending* fashion: the stimuli were presented at the same visibility level within blocks, starting with a block of the lowest (no deformation) visibility level, going up to more and more visible stimuli at each block. A second group viewed the stimuli in a reversed *Descending* path, where they first saw the objects at the highest visibility level, going down in visibility at each block. Both these groups correspond to those described in Moca et al. (2011). Finally, a last group viewed the stimuli in blocks of mixed

visibility levels, in which all the objects were presented once per block, but each at a *Random* level of visibility out of the seven. This latter presentation order resulted in an access to visual information which was intermediate between the *Ascending* presentation (lowest, uninformative content first) and the *Descending* presentation (highest, most informative content first). This group was a new addition to the data presented by Moca et al. (2011).

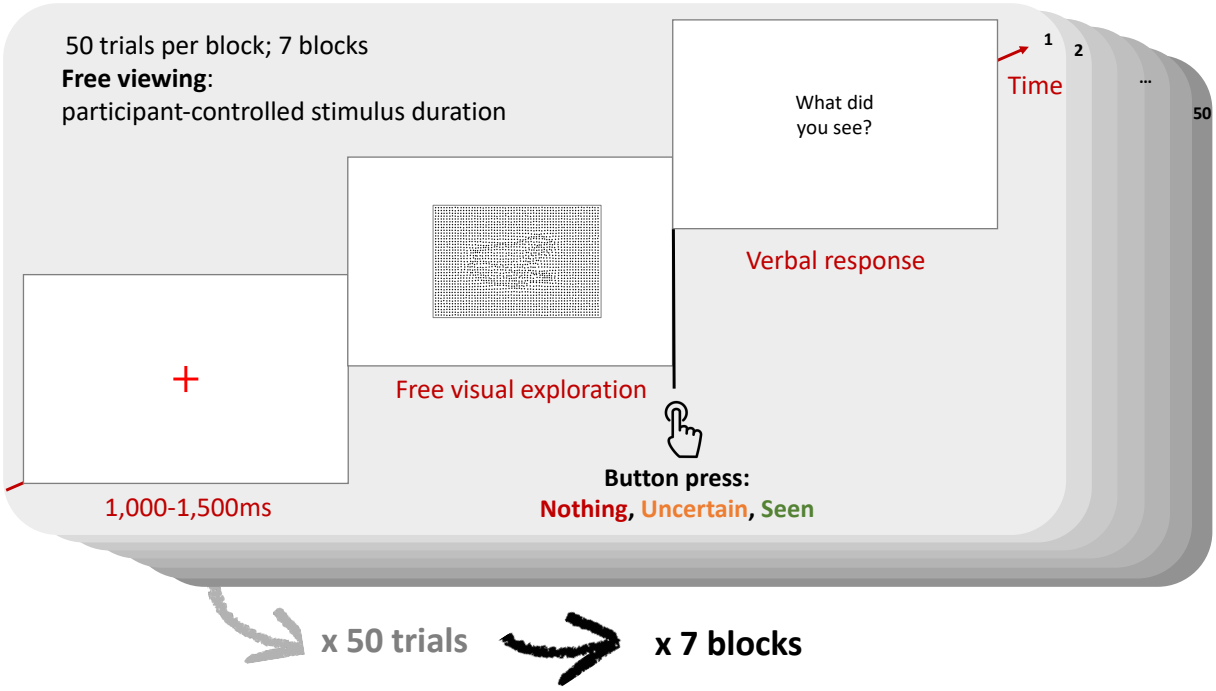


Figure 2.4: Structure of a trial; the trials were composed of a fixation cross of randomised duration in the 1,000-1,500ms interval, a free viewing exploration phase until the participant pressed one of three buttons to signify that they had viewed "Nothing", were "Uncertain" or had "Seen" an object, after what they were prompted by a sentence on the screen to verbalise the object they thought that they saw if the pressed "Uncertain" or "seen". There were seven blocks of 50 trials, and each object was shown one per block, at each of the seven *g* levels across all blocks.

2.2.4. Recordings

Participants' button-presses (Seen/Uncertain/Nothing) were recorded with precise timings, together with their verbal responses (object's name). Eye-tracking was used to monitor their gaze throughout the experiment. The eye-tracking recordings were made monocularly with an ASL EyeStart 6000 system at a rate of 50Hz. Participants' heads rested on a cheek-rest to avoid

changes in their head position during the tracking, while still enabling them to speak after each stimulus. A nine-point calibration was conducted at the start of each block, and each trials' fixation cross was used as a post-hoc calibration to correct for potential shifts in the eye position within blocks.

2.2.5. Data processing

2.2.5.1. Manual inspection of the data vs. Moca et al.'s automatic pipeline

Although two of the groups were already presented in Moca et al. (2011), all datasets from all three groups were processed anew from raw data for the present study. The same pre-processing pipeline was used, which included automatic identification of saccades and fixations, but all fixations were then manually checked for any missing, additional, or misidentified saccades. This was found to increase the quality of the data, as shown by a comparison of fixation durations between the automatic process in Moca et al.'s data and the current manually checked data. We plotted a histogram of the fixation durations, pooling all trials of all *Ascending* and *Descending* participants, for the fully automatic process (Figure 2.5A) and for the manually checked data (Figure 2.5B). This showed that the manually checked data had an overall more normal distribution: there were less of the very short and very long fixations, and the mean and median values were closer to each-other (the time bin for the mean and median values in the automatic pipeline were, respectively, 480-500ms and 900-910, which shifted to 540-560ms and 800-820ms after manual inspection). There were no substantial changes in the results between the manual and the automatic processing, thus the difference between the pre-processing pipelines will not be further discussed.

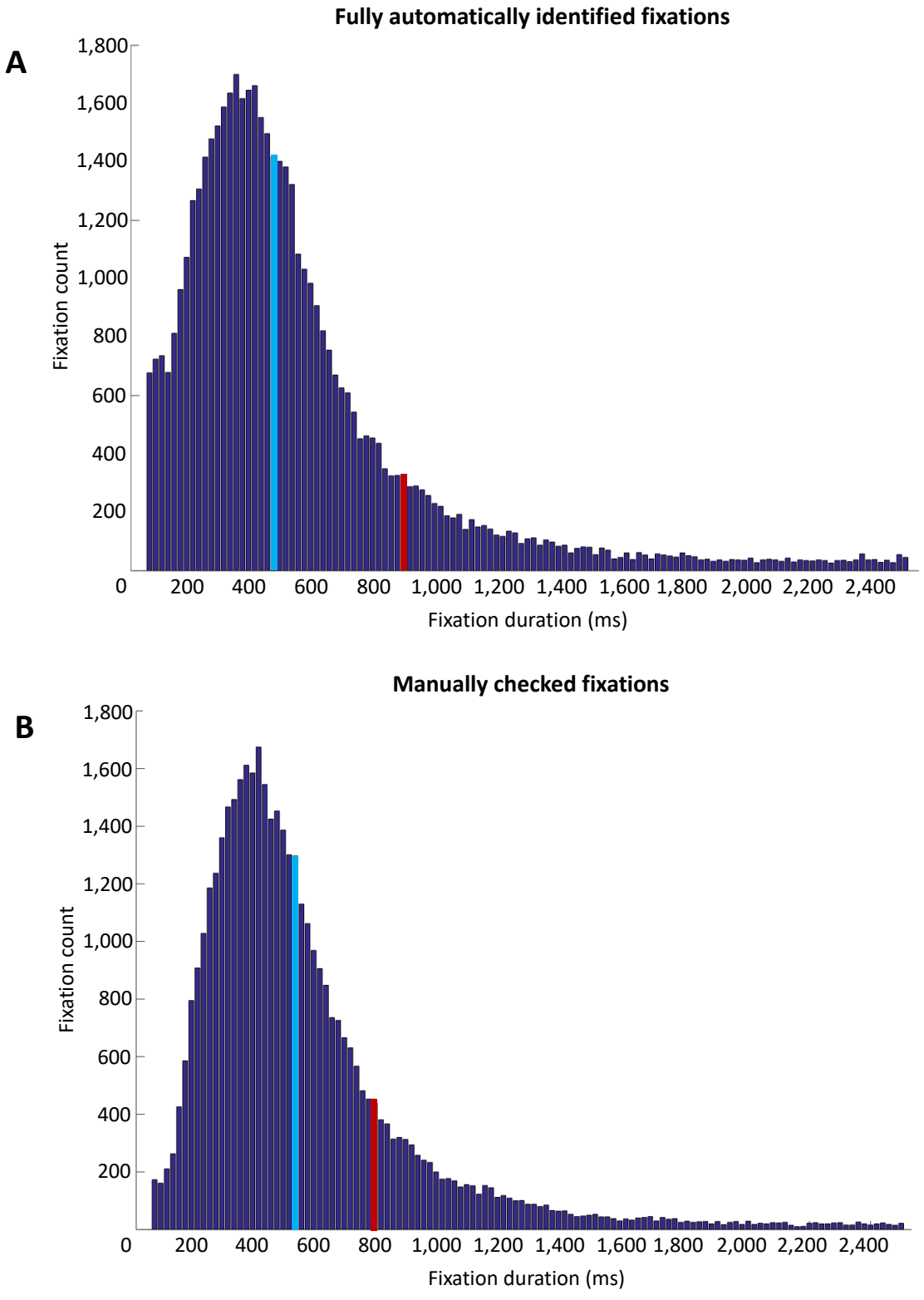


Figure 2.5: Histogram, using 20-ms time bins, of the fixation durations of all *Ascending* and *Descending* participants across all trials for the fully automatic process (**A**) and the manually checked data (**B**). The blue bin represents the bin that contains the median, and

the red bin, the mean. The tail of histogram was too long for good visualisation and was not entirely plotted here.

2.2.5.2. Pre-processing pipeline

Trials were automatically screened and any trial with over 50% data loss was discarded. Horizontal and vertical gaze location data were smoothed, and fixations were automatically detected using a simplified version of Nystrom and Holmqvist (2010) algorithm with two adaptive velocity thresholds: the velocity of the eye's position was computed and summed to obtain a horizontal-vertical composite velocity variable. Each time the velocity crossed a high threshold, a saccade was detected, whose start and end was identified using a second, lower threshold. The thresholds were computed on a trial-by-trial basis according to the trial's mean and standard deviation in the composite velocity, respectively 4 and 1.5 standard deviations above the mean. Fixations were defined as a time series between two saccades with a minimum duration of 50ms. This automatic algorithm was used to guide the parsing of the data but, as described above, a manual check of all saccades for all trials and all participants was conducted to avoid misdetections of saccades in case of local noise. The first and last fixations were discarded as they presented a significantly different profiles in standard eye-movement statistics (duration, spread) and were linked to, respectively, base-line fixations on a central cross and post-button click awaiting fixations before verbalising the answer (Henderson, 1993).

At each fixation identified, the stimulus' local information content around this point was reconstructed in two different ways. On the one hand, the information *physically* conveyed by the image was reconstructed from the stimulus image itself: taking an area of 0.5 visual degrees around the fixation, we calculated the amount of *local dots displacement (LDD)*, which directly relates to the information present at this location. On the other hand, we also calculated a more semantic*, hidden* form of local information content, using the object's source image this time: in the same area of 0.5 deg. around the fixation, we calculated the amount of *local contour density (LCD)* in the source image. Since the contour density was used to generate dots displacement in the stimulus, these two measures are highly related, however they do not fully map onto each other.

2.3. Results and specific discussion

We compared the three groups of participants, *Ascending*, *Descending* and *Random* participants, in terms of 1) behavioural responses, 2) generic eye-movement control measures of fixations, and 3) information-content measures of fixations. The last type of measure is only available because the *Dots* stimuli are generated in a highly controlled manner that allows for the extraction of information-content measures at each fixated location. To help the interpretation of these results, a discussion of each result in relation to predictions derived from our model (Figure 2.2) is included in this section, followed by a separate general discussion of the study in the next section.

We expected all groups to perform generally better as visibility increases, although with relative differences linked to their priors' strength as seen in Moca et al. (2011). The new *Random* group was expected to remain at an intermediate level between the *Ascending* and the *Descending* group across all *g* levels. We expected differences in performance to manifest across all three types of measures.

2.3.1. Behavioural results

2.3.1.1. Response type and accuracy as a function of g

First, we looked at how *g*-order groups influenced participants behavioural responses. We computed each group's average percentage of response for each *g* level in terms of response type ("Nothing", "Uncertain", "Seen"; *Figure* 2.6, top) and accuracy ("Correctly seen", "Incorrectly seen"; Figure 2.7, top) as a function of *g* level and group (*Ascending*, *Descending*, *Random*), and compared them in 3-by-7 (group*g) repeated-measure ANOVAs. Our measures were found to violate the sphericity assumption, consequently, all our ANOVAs were corrected using the Huynh-Feldt correction. Levene's tests revealed that the variances at each *g* level were also found to be unequal for most responses at most *g* levels. This is a relevant result per se showing that the priors strength influenced how heterogeneous the behaviours within the groups were. Despite this heterogeneity, we proceeded with comparing the groups in ANOVAs, which have been shown to be robust to unequal variances when the sample sizes remain equal (Weerahandi, 1995), which was our case, and to result in more conservative tests with increased likelihood for type II errors

(failing to reject the null hypothesis i.e., a false negative result) and stable type I errors (mistakenly rejecting the null hypothesis i.e., a false positive result).

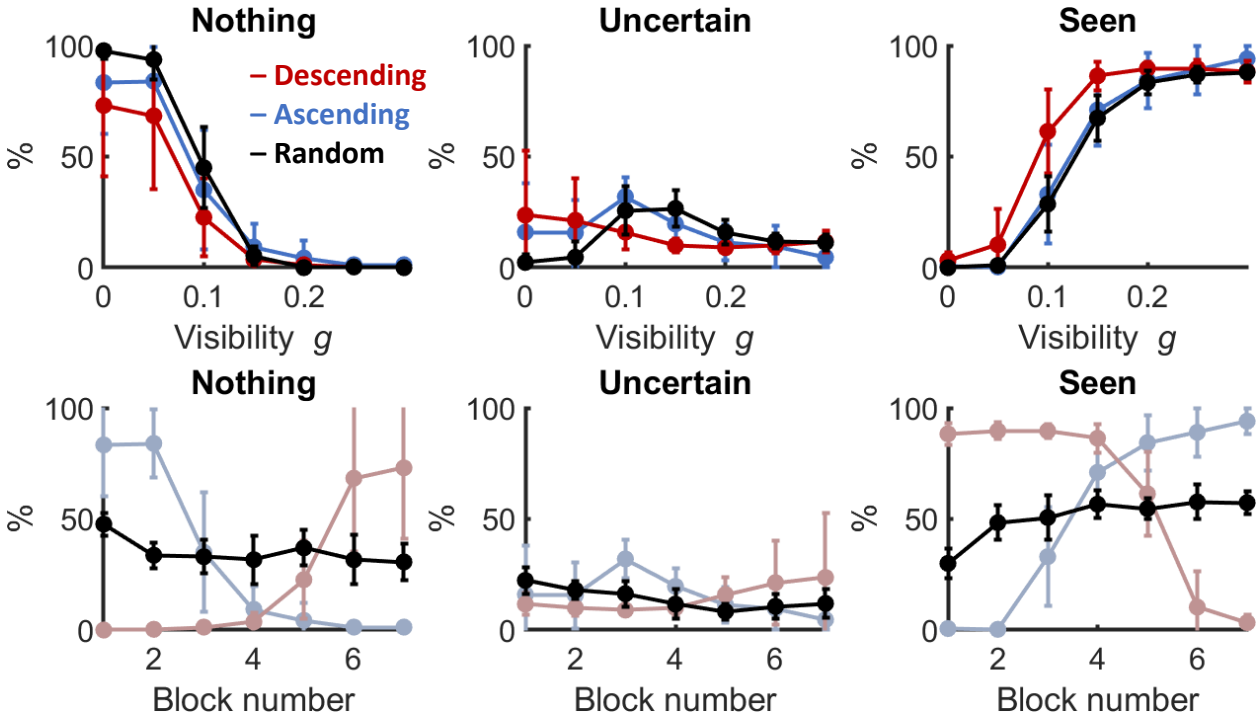


Figure 2.6: Distribution of the type of responses. The *Random* group is represented in black, the *Ascending* group in blue and the *Descending* group in red. Top: as a function of the deformation coefficient, *g*. Bottom: as a function of block order; *Ascending* and *Descending* curves are plotted with transparency as they are only included for reference and are not further analysed in text due to the impossibility to decouple learning and *g* in these groups. Error bars represent the standard deviation above and below the mean.

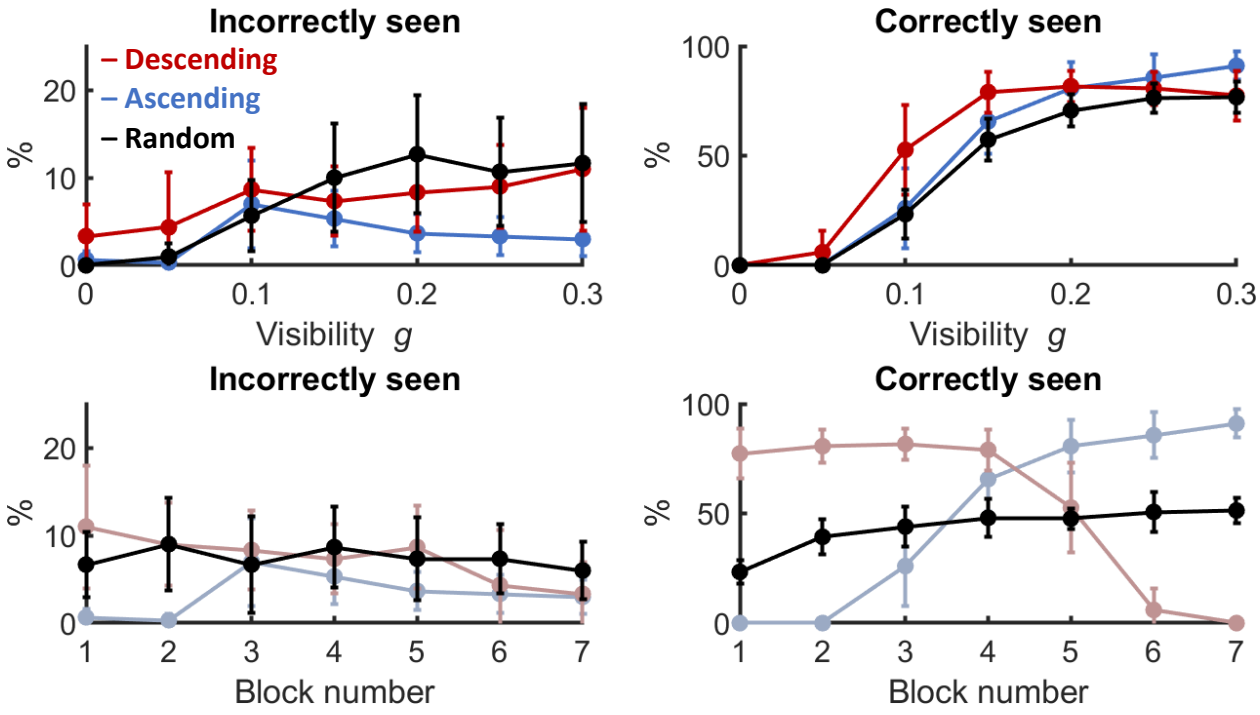


Figure 2.7: Distribution of the accuracy of Seen responses. The *Random* group is represented in black, the *Ascending* group in blue and the *Descending* group in red. Error bars represent the standard deviation above and below the mean.

Top: as a function of the deformation coefficient, *g*.

Bottom: as a function of block order; *Ascending* and *Descending* curves are plotted with transparency as they are only included for reference and are not further analysed in text due to the impossibility to decouple learning and *g* in these groups.

As Moca et al. (2011) showed, there was a strong effect of the visibility *g* on both all response type and accuracy variables, in the direction of a better recognition with increasing *g*, at the exception of the "uncertain" responses for which this effect was only trending ($F(1.729,25.932) = 3.397$, $p = 0.055$, $\eta^2 = 0.131$). This is unsurprising as these responses did not follow a linear change but a more complex pattern of general decrease with g, with a surge at middle-level, ambiguous *g* levels, which is badly characterised by standard linear statistics like the ANOVA. The groups appeared to peak at different g levels: the *Descending* participants only showed a decrease with g (peak at g = 0.00, 23.67% uncertain) while the *Ascending* group peaked at higher intermediate *g* (at g = 0.10, 32.00% uncertain). Interestingly, the *Random* group peaked at the highest level (peak at g = 0.15, 26.67% uncertain), possibly reflecting weaker priors, contrary to our hypothesis.

However, this complex effect was not clearly evidenced statistically (g*group effect trending but non-significant: $F(3.458,25.932) = 2.344$, $p = 0.089$, $\eta^2 = 0.181$), likely due to this non-linear effect being poorly described by a linear statistical test as the ANOVA.

Variable	Effect	Statistical test (F, Huynh-Feldt corrected except for group effects)	Statistical significance (p)	Effect size (η^2)
Nothing	**g**	F(2.075,31.124) = 148.66	**< 0.001**	0.828
	group	F(2,15) = 1.554	0.244	0.012
	g*group	F(4.150,31.124) = 1.462	0.236	0.016
Uncertain	**g**	F(1.729,25.932) = 3.397	**0.055**	0.131
	group	F(2,15) = 0.147	0.864	0.002
	g*group	F(3.458,25.932) = 2.344	**0.089**	0.181
Seen	**g**	F(2.861,42.921) = 389.55	**< 0.001**	0.903
	group	F(2,15) = 3.180	**0.071**	0.014
	g*group	F(5.723,42.921) = 3.642	**0.006**	0.017
Correct	**g**	F(2.759,41.384) = 389.53	**< 0.001**	0.887
	group	F(2,15) = 2.617	0.106	0.014
	g*group	F(5.518,41.384) = 5.230	**< 0.001**	0.024
Incorrect	**g**	F(2.293,34.388) = 168.22	**< 0.001**	0.820
	group	F(2,15) = 1.337	0.292	0.014
	g*group	F(4.585,34.388) = 1.259	0.304	0.012

Table 2.1: Statistical test (F), significance (p) and effect size (η^2) for each response type (nothing, uncertain, seen) and accuracy (correct, incorrect) repeated measures ANOVA test for effects of *g*, group or *g**group. Significant ($p < 0.005$) results are highlighted in grey cells and trending ($p < 0.100$) results in bold.

Moreover, there was an effect of g*group on the "Seen" and "Correct" responses (respectively, $F(5.723,42.921) = 3.642$, $p = 0.006$, $\eta^2 = 0.017$ and ($F(5.518,41.384) = 5.230$, $p < 0.001$, $\eta^2 = 0.024$), reflecting the hysteresis effect observed by Moca et al. (2011): the visibility *g* affected participants' ability to see the objects in a way that depended on their prior experience with

stimuli i.e., their group. These effects were not significant for "Nothing" and "Incorrect" responses and only trending for the "Uncertain" responses, as reported above. These results are particularly interesting when compared with our predictive model (Figure 1) of how priors and visibility influence performance. Globally, a hysteresis effect was observed as Moca et al. showed: at ambiguous *g* levels, the highest performing group was the *Descending* group who could build the strongest priors, while the *Ascending* group whose access to information allowed them to build the weakest expected priors, performed less well.

However, the newly added *Random* group did not follow our predictions. Interestingly, their "Seen" response curve mostly matched that of the *Descending* paradigm (Figure 2, top, right), but was associated with less "Correct" responses, even at medium *g* levels when the *Random* group was expected to have stronger priors and perform better (Figure 3, top, right). Indeed, when performing contrasts between the two groups at medium (0.05-0.215) and high (0.25 & 0.30) *g* levels (Table 2.2) to compare them at the generic *g* levels defined in the model, we found a significant difference between *Random* and *Descending* participants' "Correct" answers at high *g* ($t(1,26.81) = 2.216$, $p = 0.035$) but not for "Seen" responses and not for medium *g*.

Responses	***g* level**	**Mean (S.E.) (%)**	**Statistical test (t)**	**Significance (p)**
"Seen"	**Medium**	9.00 (14.6)	t(1,22.85) = 0.616	0.544
	High	8.00 (10.5)	t(1,30.52) = 0.760	0.453
"Correct"	**Medium**	21.00 (15.0)	t(1,21.01) = 1.398	0.177
	High	23.7 (10.7)	t(1,26.81) = 2.216	**0.035**

Table 2.2: Mean (Standard Error S.E.), statistical value (t) and significance (p) for the contrasts between the *Random* and the *Descending* participants' responses ("Seen" and "Correct") for medium and high *g* levels. Significant ($p < 0.005$) results are highlighted in grey cells and bold.

The difference between their "Seen" and "Correct" curves, relative to the other groups, seems to reflect a decoupling in how randomness affected participants' ability to recognise stimuli and their own perception of their abilities. These findings did not follow our predictions and seemed to reflect a poorer building of priors than expected in the *Random* group. In fact, *Random*

participants seemed to rather match the expected performance for the hypothetical *naïve* group and remained generally below both the other groups all throughout, in all behavioural measures.

2.3.1.2. Response type and accuracy as a function of block

For *Ascending* and *Descending* groups, learning and prior building are confounded with changes in *g* over time, block by block. However, in the *Random* paradigm, the two processes can be decoupled to look at learning as a function of block. This allows to investigate more into depth whether *Random* participants are in fact building priors and getting better with experience, or if they remain naïve.

We investigated participants' responses by *block* order rather than by *g* (Figures 2 and 3, bottom row). For reference, all groups are displayed, but the curves for the *Ascending* and *Descending* groups are not further discussed due to the confound between *g* and *block* for these groups. We found that subjects engaged in the *Random* paradigm performed rather consistently across blocks, except for an enhancement after block 1 with increased "Correct" and "Seen" responses and decreased "Nothing" responses, indicating punctual learning progress (Figure 2.6 and Figure 2.7, bottom row). This was reflected in post-hoc tests comparing the responses in each block through Bonferroni corrected comparisons (family of 21) revealed that the "Correct" responses in block 1 were significantly different to responses in block 3 ($t(1,102) = -3.435$, $p = 0.018$) and beyond, the "Seen" responses as well (at block 3, $t(1,102) = -3.379$, $p = 0.022$) and the "Nothing" responses in block were trending towards a significant different with block 4 ($t(1,102 = 2.983$, $p = 0.075$) and significantly different with block 7 ($t(1,102) = 3.169$, $p = 0.045$). "Incorrect" responses in block 1 appeared mildly decreasing, and were only significantly different to block 7's ($t(1,102) = 3.135$, $p = 0.047$), as well as "Uncertain" responses which seemed to also resurge slightly at the end, as they were found to be only significantly different to block 5 ($t(1,102) = 3.577$, $p = 0.011$) and trending towards a significant difference with block 6 ($t1,102) = 2.981$, $p = 0.075$) nut not block 7. All the other comparisons of later blocks with each other for all responses were non-significant ($p > 0.100$).

Had the *Random* participants continuously built priors, we would have expected a continuous increase in performance across blocks. However, performance improvement was only observed after the first block. This, again, suggests that *Random* subjects were mostly acting as naïve observers, despite their constant access to meaningful information. The increase in performance

after the first block is likely to reflect an increase in task-general priors rather than object-specific priors: after the first block, participants acquired a good knowledge of the general properties of the stimuli that they are presented with, which helped them to generally recognise objects better.

2.3.1.3. Corrected response type and accuracy for the Random group (no block 1)

To verify this, we recomputed the response curves in the *Random* paradigm, excluding the first block or the first two blocks associated to task-general prior building and qualitatively assessed the changes to the *Random* group's curves (Figure 2.8). These recomputed response curves of the *Random* group approached those predicted by our model but never reached a level intermediate between the *Ascending* and *Descending* groups as predicted. In particular, the *Random* group's "Correct" responses at medium visibility levels remained around the *Ascending* group's responses who was considered to have very weak priors at this stage and was strikingly similar to the *Descending* group's responses at the highest visibility level (g = 0.30) when this group had not formed any priors yet.

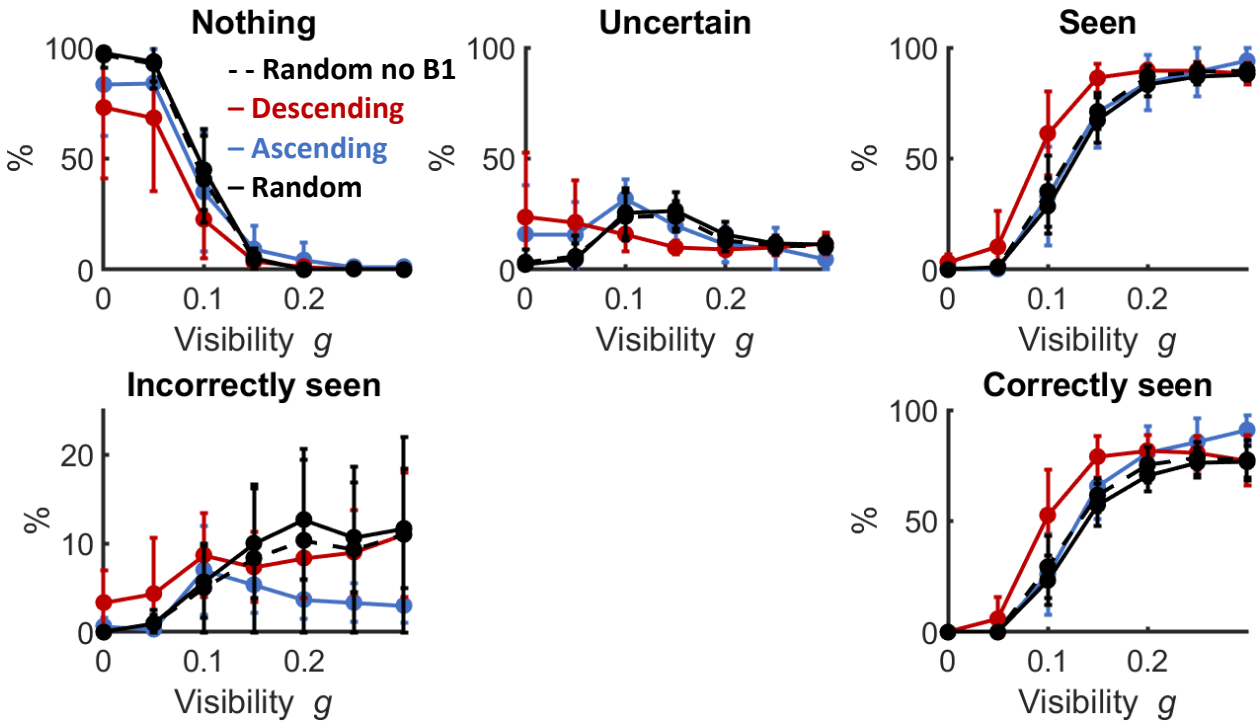


Figure 2.8: Superimposed on Figure 2.6's and Figure 2.7's curves, recomputed response curves of the *Random* group when not taking the first block into account (dashed lines). Error bars represent the standard deviation above and below the mean.

Although these observations are qualitative and not backed by quantitative statistics, they can help discussing how to interpret the *Random* group's behaviour, especially considering the striking shifts in responses in all measures towards a higher yet never intermediate performance, often closely matching the other groups at their stage of least strong priors. This is in favour of our interpretation that the *Random* participants were largely behaving as a naïve group whose behaviour was not guided by priors, contrary to what was expected given their intermediate access to informative stimuli.

2.3.2. Visual fixations: generic eye-movement control measures

Our following step was to investigate how the effects of visibility *g* and group on behaviour translated to participants' visual exploration of the stimuli. Importantly, in all paradigms, participants were free to visually explore the stimuli with no time constraint. This yielded an important number of fixations and saccades with 9.82 fixations per trial on average (S.D. = 4.90) talking all groups and *g* levels together. This particularity of the *Dots* stimuli enabled a closer investigation of eye-movement control measures and their alterations by experimental manipulation of priors and visibility. We expected that differences in performance measures would also be found in eye-movement control measures, namely fixation count, spread and duration (Figure 2.9, top row), as reported by Moca et al. (2011), and compared them in 3-by-7 (group*g) repeated-measure ANOVAs as with the response measures. Again, the measures were found to violate the sphericity assumption and all ANOVAs were corrected using the Huynh-Feldt correction. However, Levene's tests revealed no significant difference in variances at any *g* level.

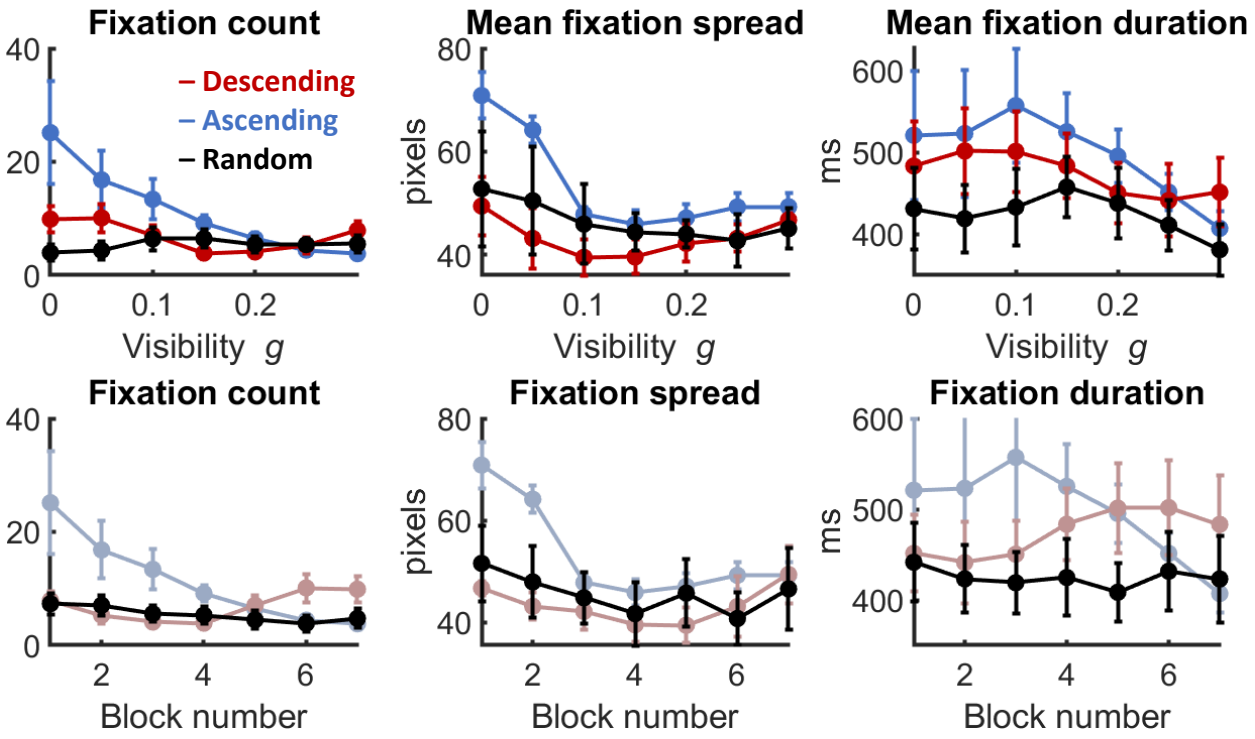


Figure 2.9: Eye-movement statistics as a function of visibility level (top) and block (bottom), by experimental group (*Random* in black, *Ascending* in blue, *Descending* in red). Left column: Fixation count per trial. Middle column: Overall spread (pixels) in fixations across the visual stimulus per trial. Right column: Duration (ms) of individual fixations. Error bars represent the standard deviation above and below the mean.

We found a significant effect of *g* on all of the eye-movement control variables (Table 2.3) which all decreased with higher visibility levels *g*. There was no significant main effect of group, although there appeared to be a trend in groups' effect on the fixation count measures ($F(2,15) = 2.961$, $p = 0.082$, $\eta^2 = 0.113$). This was accompanied by a significant interaction effect of *g**group on fixation count ($F(2.805,21.034) = 7.770$, $p = 0.006$, $\eta^2 = 0.201$). There was also a trending interaction effect on the fixation spread variable ($F(3.682,27.616) = 2.488$, $p = 0.071$, $\eta^2 = 0.071$) but not on the fixation duration variable ($F(3.842,28.813) = 0.999$, $p = 0.422$, $\eta^2 = 0.026$).

Variable	Effect	Statistical test (F: Huynh-Feldt corrected, except for group effects)	Statistical significance (p)	Effect size (η^2)
Fixation count	**g**	F(1.402,21.034) = 7.975	**0.006**	0.139
	group	F(2,15) = 2.961	**0.082**	0.113
	g*group	F(2.805,21.034) = 7.770	**0.006**	0.201
Fixation spread (pixels)	**g**	F(1.841,27.616) = 11.303	**< 0. 001**	0.161
	group	F(2,15) = 1.999	0.170	0.117
	g*group	F(3.682,27.616) = 2.488	**0.071**	0.071
Fixation duration (ms)	**g**	F(1.921,28.813) = 5.550	**0.010**	0.073
	group	F(2,15) = 0.996	0.392	0.083
	g*group	F(3.842,28.813) = 0.999	0.422	0.026

Table 2.3: Statistical test (F), significance (p) and effect size (η^2) for eye-movement control variables' (Fixation count, spread and duration) repeated measures ANOVA test for effects of *g*, group or *g**group. Significant ($p < 0.005$) results are highlighted in grey cells and trending ($p < 0.100$) results in bold.

2.3.2.1. Fixation count

The fixation count measure (Figure 2.9, left column) appeared to be the most clearly affected by the manipulation of participants' priors through the experimental groups. This measure can be seen as a reflection of participants' amount of exploration.

The *Ascending* group started with a markedly high number of fixations at low *g* which steadily decreased with increasing *g*, until reaching a level lower than all other groups at the highest *g* (contrast between the *Ascending* group and both other groups at $g = 0.00$: $t(1,48.107) = -2.172$, $p = 0.035$; and at $g = 0.30$: $t(1,48.107) = 2.684$, $p = 0.010$).

The *Descending* group's fixation count followed a more complex non-linear pattern. They started the experiment at high *g* with a high number of fixations, which was not significantly different

than both other groups (contrast between the *Descending* group and both other groups at g = 0.30: t(1,48.107) = 0.758, p = 0.452) and decreased with decreasing g, reaching the lowest value of all groups at intermediate g (contrast between the *Descending* group and both other groups at g = 0.15: t(1,48.107) = 3.158, p = 0.003) before increasing their fixation count again at low g levels, when information is scarce.

Both groups' visual behaviour followed our model's prediction (Figure 2.1) that either weak priors or scarce information would result in more intense exploration.

The added *Random* group's fixation count, on the contrary, appeared rather stable across visibility levels and a post-hoc test revealed no main effect of g in this group (F(6,96) = 0.254, p = 0.956). Our prediction that participants would compensate information scarcity by intensified exploration, which was verified by both the other groups who explored more at low g, was not verified for this group. Randomness seemed to impair *Random* participants' exploration by preventing this compensation effect.

We also looked at the evolution of fixation count by block in *Random* participants (Figure 2.9, bottom left plot), which showed a slow decrease across blocks (main effect of block: F(6, 96) = 2.815, p = 0.014, no significant Bonferroni-corrected post-hoc comparison driving the effect in a particular block), which could reflect a small decrease in motivation to explore over time for this group.

2.3.2.2. Fixation spread

Participants' fixation spread (Figure 2.9, middle column) also appeared to be modulated by their priors, or group. This measure is thought to be linked to participants' exploration efficiency: as the stimuli were placed at the centre of the images, a more central, less spread-out exploration was more optimal. Participants could acquire this knowledge regardless of the actual object presented, thus it was linked to task-general priors. Measures of information content at each fixation can be used to investigate this more precisely, but including this variable enables to test the utility of this easily accessible variable, in comparison with more precise but more design-specific information-related variables like LCD and LDD.

Ascending participants fixated areas more spread-out than both the other groups for trials of low *g-level* (contrast between the *Ascending* group and both other groups at g = 0.00: t(1,31.882) =

4.048, $p < 0.001$) before reducing their spread to a plateauing value. This was compatible with our prediction that exploration would be facilitated by stronger priors and better access to information (*g*). However, *Ascending* participants' fixation spread remained higher than both the other groups over medium and high *g* values (contrast between the *Ascending* group and both other groups at $g = 0.10\text{-}0.30$: $t(1,15.986) = 6.010$, $p < 0.001$). This last result could not be explained by our model and was unexpected. It is possible that it reflects the strength of participants' task-general priors, acquired at the start of study and possibly not updated later with increased access to information: participants appear to have remained biased to explore at spread locations, even when their object-specific priors became stronger.

The fixation spread for the *Descending* group followed a U-shaped curve, similar to what was described for the fixation count variable: compared to both other groups, they explored at more spread-out location at high *g* levels (contrast between the *Descending* group and both other groups at $g = 0.30$: $t(1,31.882) = 6.148$, $p = < 0.001$) and less spread-out for low *g* levels (same contrast at $g = 0.00$: $t(1,31.882) = 9.591$, $p < 0.001$). They explored at the lowest value of all groups and visibility levels at intermediate *g* (same contrast at $g = 0.15$: $t(1,31.882) = 6.557$, $p < 0.001$). Again, this followed our prediction that priors' strength and access to information would facilitate exploration.

Interestingly, the *Random* participants kept their fixations rather evenly spread out across *g* levels as was also previously the reported for their fixation count, with no main effect of *g* on fixation spread in this group ($F(6,96) = 1.805$, $p = 0.107$). Interestingly though, their fixations were directed at intermediately spread-out locations, which suggests that they were not completely naïve and had some task-general priors to guide their fixations at central locations, as we suggested earlier analysing their behaviour. As this was not associated with an intermediate behavioural performance, this suggests that task-general priors seem to aid the exploration but not the recognition: there seems to be a decoupling between the two processes.

We also looked at the evolution of fixation spread by block in *Random* participants (Figure 2.9, bottom middle plot), who appeared to redirect their fixation to more central locations over the first blocks, compatible with our hypothesis that this is linked to the acquisition of task-general priors (main effect of block: $F(6,96) = 3.082$, $p = 0.008$, which seemed to be mainly driven by block 1 after Bonferroni-corrected post-hoc tests, which revealed a significant different between block

1 and block 3: $t(1,102) = 3.323$, $p = 0.025$, and block 5: $t(1,102) = 3.947$, $p = 0.003$, while all other comparisons and comparisons of later blocks with each other were non-significant: $p > 0.100$).

2.3.2.3. Fixation duration

Finally, the duration of fixations was investigated (Figure 2.9, right). The groups appeared to follow different, non-linear complex changes in fixation duration with *g*, qualitatively different to the changes seen in the other two variables, which were relatively similar to each-other. This suggests that the relationship between this variable and participant's behaviour was more complex, on which we will not speculate here.

Fixation durations were significantly shorter for the *Random* participants than for the other two groups (contrast between the *Random* group and both other groups: $t(1,15) = 8.479$, $p < 0.001$), which again did not follow our prediction that this group would perform at an intermediate level. Moreover, fixation duration in the *Random* group did not seem to change throughout blocks ($F(6.96) = 0.356$, $p = 0.905$, $\eta^2 = 0.022$, see Figure 2.9, bottom-right), suggesting that this variable was not influenced by task-general priors.

2.3.2.4. Overall results for eye-movement control measures

Taken together, these results indicate that stimulus presentation order clearly impacted participants' visual sampling behaviour. In particular, randomising the stimulus presentation, although along a dimension irrelevant to the task goal, appeared to reduce fixations' duration and prevent the process of adapting exploration to the information available. Task-general priors influence on eye movements appeared to be maintained despite randomness, while task-specific priors' effect was not evidenced.

2.3.3. Visual fixations: information content of explored locations

Finally, participants fixations were analysed in terms of the local information content of each location explored. *Physical* information could be quantified using the measure of *local dot displacement* around fixations (LDD): this type of information represents information readily available to the participants. We also extracted the *hidden* information with the *local contour density* (LCD) measure around each fixation. This second type of information is related to the first but represents the *g*-unrelated information from the source image rather than *g*-related information from the stimulus. This second "LCD" type of information is ultimately the most

meaningful and directly informative. The extent to which the LCD was transferred into available LDD in the *Dots* stimuli progressively increased as the *g* parameter was increased (*g* = 0 corresponds to zero transfer, while *g* = 0.30 corresponds to high transfer). The two measures allowed us to quantify how *physical*, readily available information (LDD) vs. *hidden*, meaningful information (LCD) were sampled, as a function of visibility and presentation order.

Both information measures were analysed in two ways (Figure 2.10): computing the *average* trial value in LCD and LDD reflected participants' general ability to access information, while looking at the *total* (sum) trial value reflected participants' overall integrated information needed to reach their decision. Both the amount of total information to reach a decision and the average amount of information accessed depended on the availability of information (*g*) and participants' priors. To note, the first visibility level corresponded to no available information, thus all groups' mean or total LDD meet at the origin for g = 0. Any hidden information (mean or total LCD) accessed by the participants in this block was either due to chance or to task-general priors on where to expect the informative points to lie in the absence of physical guidance. Again, the measures were found to violate the sphericity assumption, so the groups were compared in ANOVAs corrected with the Huynh-Feldt correction. Levene's tests revealed significant differences in variances at some *g* levels, but as for the behavioural measures, we proceeded with comparing the groups in ANOVAs as the sample sizes remain equal (Weerahandi, 1995), resulting in more conservative tests with increased likelihood for type II but not type I errors.

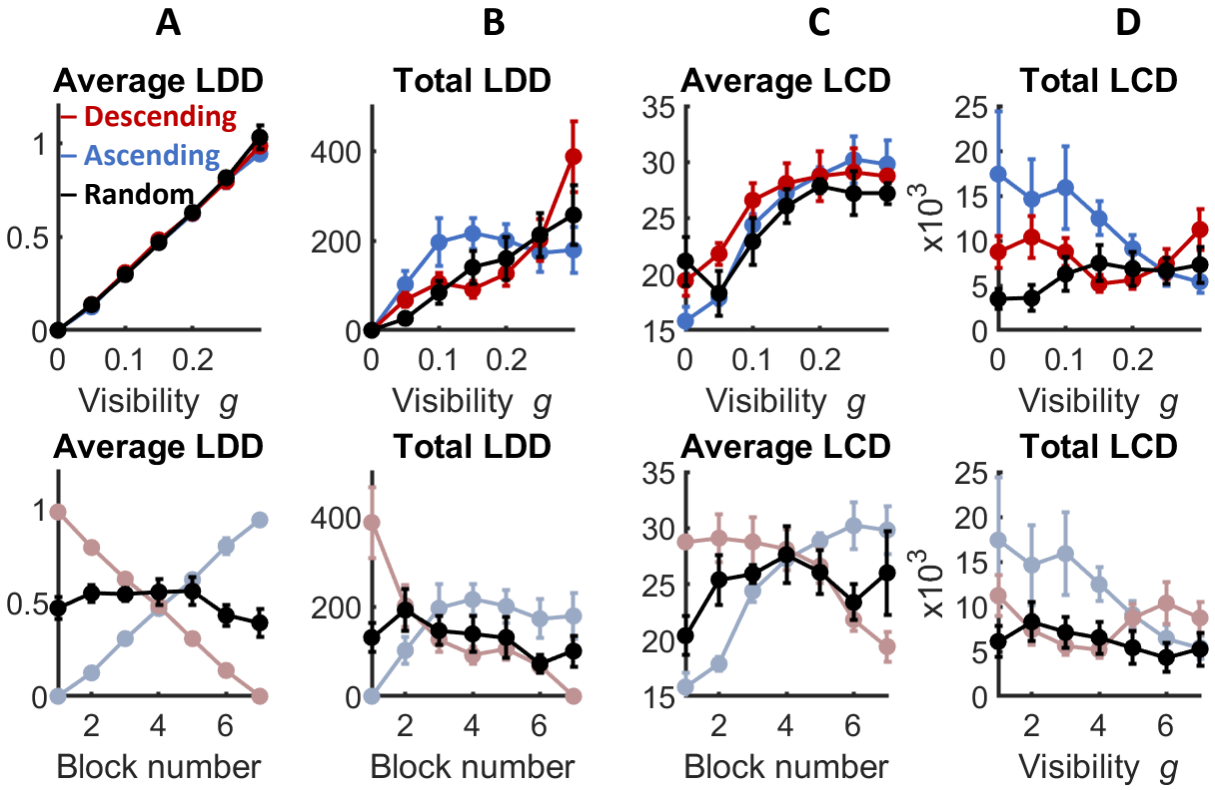


Figure 2.10: Information content of fixations as a function of visibility level (top) and block (bottom), by group (*Random* in black, *Ascending* in blue, *Descending* in red). **A.** Average local dot displacement (LDD) per fixation.

B. Total local dot displacement (LDD) as a sum of dot displacements over all fixations in the trial.

C. Average local contour density (LCD) per fixation.

D. Total local contour density (LCD) over all fixations in the trial. Error bars represent the standard deviation above and below the mean.

2.3.3.1. Physical information: LDD

We found an effect of *g* for both the average and total LDD information at fixations, as was expected from the task's design (Table 2.4): participants explored more informative areas as visibility increased. To note, images' average LDD also increased with *g*, such that this effect did not reflect better exploration.

2.3.3.1.1. Average LDD

The average LDD (Figure 2.10A, top) did not appear to be influenced by participants' group, as reflected by no significant effect of group or *g**group. This was verified by a Bayesian ANOVA, which revelead that the model of a simple *g* effect was billions of times more likely than the null hypothesis ($BF_{10} = 6.85*10^{98}$) while the model of a simple effect of group was always 10 times less likely than the null hypothesis ($BF_{10} = 0.127$). On average, subjects explored locations equally informative in terms of physical information, independent of the strength of their priors, reflecting the validity of the experimental design in increasing physical information with *g*, equally accessible to all participants.

We also looked at how the average LDD evolved by blocks in the *Random* group (Figure 2.10A, bottom). The average LDD did not appear to change over block, as reflected by no main effect of block ($F(6,96) = 1.623$, $p = 0.149$) and no significant post-hoc block-to-block comparisons.

2.3.3.1.2. Total LDD

On the contrary, for total LDD or the LDD integrated to reach a decision (Figure 2.10B), we found a strong interaction effect of g*group ($F(4.317,32.379) = 5.494$, $p = 0.001$, $\eta^2 = 0.131$).

At medium *g* level, the *Ascending* group did not appear to integrate a significantly different amount of information (contrast between the *Ascending* group and both other groups at $g = 0.15$: $t(1,42.495) = 0.261$, $p = 0.796$) while the *Descending* group integrated the least amount of information (contrast between the *Descending* group and both other groups at $g = 0.15$: $t(1,42.495) = 4.352$, $p < 0.001$). The *Random* group appeared intermediate between the groups.

At high *g* level, the *Ascending* group integrated less information than either of the other groups (contrast between the *Ascending* group and both other groups at $g = 0.30$: $t(1,42.495) = 7.644$, $p = 0.429$) while the *Descending* group did not significantly differ from the other groups (contrast between the *Descending* group and both other groups at $g = 0.30$: $t(1,42.495) = 0.799$, $p = 0.429$). Again, the *Random* group remained intermediate.

This follows what was described for the eye movement variables: the *Ascending* and the *Descending* group appeared to modulate their integration of information depending on their priors, which each followed a different non-linear curve: they integrated less information than the others when their priors were stronger. By contrast, the *Random* group's amount of

integrated physical information to reach a decision scaled linearly with *g*: their sampling behaviour appeared driven exclusively by the increase in local available information with *g*, congruent with the idea that changing priors did not guide their exploration.

Variable	Effect	Statistical test (F, Huynh-Feldt corrected except for group effects)	Statistical significance (p)	Effect size (η^2)
Average LDD	g	F(2.753,41.297) = 1088.508	**< 0.001**	0.981
	group	F(2,15) = 0.834	0.454	0.0004
	g*group	F(5.506,41.297) = 0.866	0.520	0.002
Total LDD	g	F(2.159,32.379) = 37.768	**< 0.001**	0.449
	group	F(2,15) = 0.251	0.791	0.008
	g*group	F(4.317,32.379) = 5.494	**0.001**	0.131

Table 2.4: Statistical test (F), significance (p) and effect size (η^2) for average and total LDD measures' repeated-measure ANOVA test for effects of *g*, group or *g**group. Significant ($p < 0.005$) results are highlighted in grey cells and bold font.

Looking at the *Random* group's evolution in time, there was a main effect of block ($F(6,96) = 2.488$, $p = 0.028$) which appeared to be driven by an increase in block 2 and a decrease in block 6 (Bonferroni-corrected post-hoc comparison between blocks 2 and 6: $F(1,102) = 3.119$, $p = 0.014$, all other comparisons non-significant, $p > 0.100$). Similarly to behavioural measures, total LDD seemed to be significantly improving in the second block, likely reflecting improved exploration with the acquisition of task-general priors. We also observed a decrease in the second to last block, possibly due to a drop in motivation towards the end, and a resurgence in the last block, congruent with the hypothesis of a motivational effect as participants were notified of the fact that this was their last block and likely showed a resurge in motivation.

These results suggest that the *Ascending* and *Descending* groups' integration of LDD was modulated by a combination of information availability and priors (likely a combination of task-general and object-specific priors), while the information integration of the *Random* group seemed mainly driven by information availability, with a small effect of task-general priors.

2.3.3.2. Hidden information: LCD

Next, we looked at participants' exploration in terms of *hidden* infromation, or LCD. We found similar effects than with LDD: this was expected since both measures are linked, but there were some noticeable differences discussed below.

2.3.3.2.1. Average LCD

As *g* increased, LDD increased and made LCD more accessible to participants: this was reflected by a strong effect of *g* on average LCD (Figure 2.10C; $F(4.747,71.204) = 51.184$, $p < 0.001$, $\eta^2 = 0.585$). However, the extent to which participants managed to access this hidden information at each visibility level appeared to depend on their group (which was not the case for the LDD reported earlier).

At low and medium *g* levels, the *Descending* group appeared to explore locations with higher hidden information (LCD) than the other groups (contrast between the *Descending* group and both other groups at $g = 0.00$-0.20: $t(1,17.026) = 17.931$, $p < 0.001$).

At the lowest $g = 0.00$ level, the *Random* group also appeared to explore high LCD locations (contrast between the *Random* group and the *Ascending* group at $g = 0.00$: $t(1,50.940) = -2.626$, $p = 0.011$; between the *Random* group and the *Descending* group at $g = 0.00$: $t(1,50.940) = -0.799$, $p = 0.428$). This supports the idea that *Random* subjects developed task-general priors that could guide them in the absence of information, to the same extent than the strong-prior *Descending* subjects, but unlike the naïve *Ascending* group. However, for $g > 0.00$, *Random* participants appeared to explore generally lower LCD locations than both the *Ascending* and *Descending* groups (contrast between the *Random* group and both other groups at $g = 0.05$: $t(1,50.940) = 8.853$, $p < 0.001$). This suggests that physical information, when available ($g > 0.00$) had more weight in guiding their fixations than task-general priors. There appeared to be no change in how this group accessed the average LCD with time: no Block effect ($F(6,96) = 1.741$, $p = 0.120$) and no significant Bonferroni-corrected post-hoc blocks comparison.

Finally, while the *Ascending* group started at a lower LCD level at $g = 0.00$, they accessed progressively more hidden information with increased *g* and their average LCD value exceeded the values of the other two groups at high *g* (contrast between the *Ascending* group and both

other groups at g = 0.30: t(1,50.940) = 10.815, p < 0.001; contrast between the *Random* group and both other groups at g = 0.05: t(1,50.040) = 8.853, p < 0.001).

The exploration behaviour of the *Ascending* and *Descending* groups was compatible with our model of prior-guided exploration, as they accessed more meaningful locations compared to the other groups when they were expected to have the strongest priors. However, the *Random* group appeared to build task-general priors only, as reflected by no changes in average LCD by block and a high average LCD at g = 0.00, comparable to the expert *Descending* group at this level.

2.3.3.2.2. Total LCD

Total LCD, contrary to average LCD, did not appear to be influenced by g alone (Figure 2.10D; F(2.001,30.009) = 2.174, p = 0.131, η^2 = 0.038), although there was a significant interaction effect of g*group (F(4.001,30.009) = 4.922, p = 0.004, η^2 = 0.173). This suggested that participants' amount of *hidden* information to reach a decision was influenced by g, in a manner that depended on their access to information over the study (group).

The *Ascending* group's total integrated LCD decreased with visibility (simple main effect of g for the *Ascending* group: F(6,96) = 3.814, p = 0.006), suggesting that they reached their decisions with less and less information as they built priors. The *Descending* group's total LCD however changed with g in a "s-shaped" manner which reflected interactive effects of prior strength and information availability: they needed more information to reach a decision at high g when their priors were weak, as well as at low g when the available information was low (contrast between the *Descending* group's values at g = 0.00 & 0.30, and g = 0.15: t(1,85.668) = -4.584, p < 0.001). Both groups thus followed our hypothesis that the amount of information needed to reach a decision increased when either their priors or available information were weaker.

By contrast, the *Random* group exhibited a rather flat total LCD curve at higher g which decreased for lower g. Indeed, a repeated measures ANOVA on the *Random* groups' lower g = 0.00-0.15 revealed a significant effect of g (F(3,15) = 13.711, p < 0.001) which was not found at higher g = 0.15-0.30 (F(3,15) = 0.538, p = 0.663). Thus, the *Random* group's information sampling behaviour was not influenced so much by priors but rather by the available information, which actually had the effect of reducing their need for information to reach a decision at low g, contrary to our hypothesis that participants would compensate the lack of information with increased

exploration. This effect in the opposite direction to our hypothesis was possibly due to the low motivation effects evoked earlier. There also appeared to be no time change in how much total LCD this group integrated to reach a decision, as reflected by no Block effect ($F(6,96) = 1.766$ $p = 0.114$) and no significant Bonferroni-corrected post-hoc comparison of the blocks, compatible with the idea that *Random* participants did not build object-specific priors to better guide their exploration of object-related hidden information with time.

Variable	**Effect**	**Statistical test (F, Huynh-Feldt corrected except for group effects)**	**Statistical significance (p)**	**Effect size (η^2)**
Average LCD	g	F(4.747,71.204) = 51.184	**< 0.001**	0.585
	group	F(2,15) = 0. 789	0.472	0.018
	g*group	F(9.494,71.204) = 2.236	**0.027**	0.051
Total LCD	g	F(2.001,30.009) = 2.174	**0.053**	0.038
	group	F(2,15) = 2.542	0.112	0.173
	g*group	F(4.001,30.009) = 4.922	**0.004**	0.133

Table 2.5: Statistical test (F), significance (p) and effect size (η^2) for average and total LCD measures' repeated-measure ANOVA test for effects of *g*, group or *g**group. Significant ($p < 0.005$) results are highlighted in grey cells and trending ($p < 0.100$) results in bold font.

2.4. General discussion

2.4.1. Summary of findings

2.4.1.1. General context

In this study, we used a perceptual task that offers the possibility to precisely manipulate participants' access to information along two dimensions: 1) in-the-moment available information, manipulated through modulations of stimuli's visibility level *g* in a quantifiable manner; 2) previously accessed information i.e., priors, manipulated through group differences of the block-by-block *g* order of stimulus presentation. We analysed the visual exploration and recognition of objects in *Dots* stimuli in three groups: one group viewing the stimuli in an *Ascending* order of visibility *g* (expected to be naïve at the start for low and medium *g*, and have strong priors at the end for high *g*), another in a *Descending* order (expected to be naïve at the start for high *g*, and have strong priors at medium and low *g*), and a last one in a *Random* order of *g* (expected to have intermediate priors all throughout). Groups' performance and expected priors' strength at low, medium, and high *g* levels were predicted in a detailed model (Figure 2.1) which offered a window into how available information, as well as task-general and object-specific priors, may interact to influence object recognition.

2.4.1.2. Priors' influence on objects' exploration and recognition

We showed that both types of information access influenced participants' ability to recognise objects and how they visually explored them. First, the *g* level had a significant effect on participants' visual exploration in terms of basic eye-movement control measures (fixation count, spread, and duration) but also information-content related to fixations (LCD, LDD). Secondly, we also found effects of group on participants' visual exploration of the stimuli, both in terms of eye-movement control and information-content measures. Together, these results show that: 1) manipulating the available visual information modulates visual behaviour substantially, and 2) the structure of the information presentation framework also strongly influences visual behaviour and perception threshold.

2.4.1.2.1. The case of the Ascending and Descending groups

The two groups viewing the stimuli in structured manner, the *Ascending* and the *Descending* groups, were already analysed in Moca et al. (2011). They showed that participants' prior experience with information influenced how well they performed and explored, which we took further by building a model of how priors and available information influence performance over different visibility levels. This model was able to explain changes in a lot of detail, for all the behavioural and visual variables assessed.

2.4.1.2.2. The case of the Random group

We also analysed an additional group compared to Moca et al.'s study (2011), who viewed the stimuli in a randomised fashion. Their performance appeared to match the naïve participants' performance i.e., the *Ascending* group at low *g* levels and the *Descending* group at high *g* levels, suggesting that their ability to use object-specific priors was impaired. However, their performance did improve over the first blocks of the experiment, suggesting that they acquired *task-general* knowledge at the start of the study.

Overall, these results show that randomised access to visual information, although along a dimension irrelevant to the task goal, strikingly impaired participant's ability to use object-specific priors to guide their exploration and recognition of objects. Their ability to use task-general priors however appeared to be maintained. We updated our model accordingly, adjusting Random participants' object-specific-priors and performance to match the theoretical naïve group (Figure 2.11).

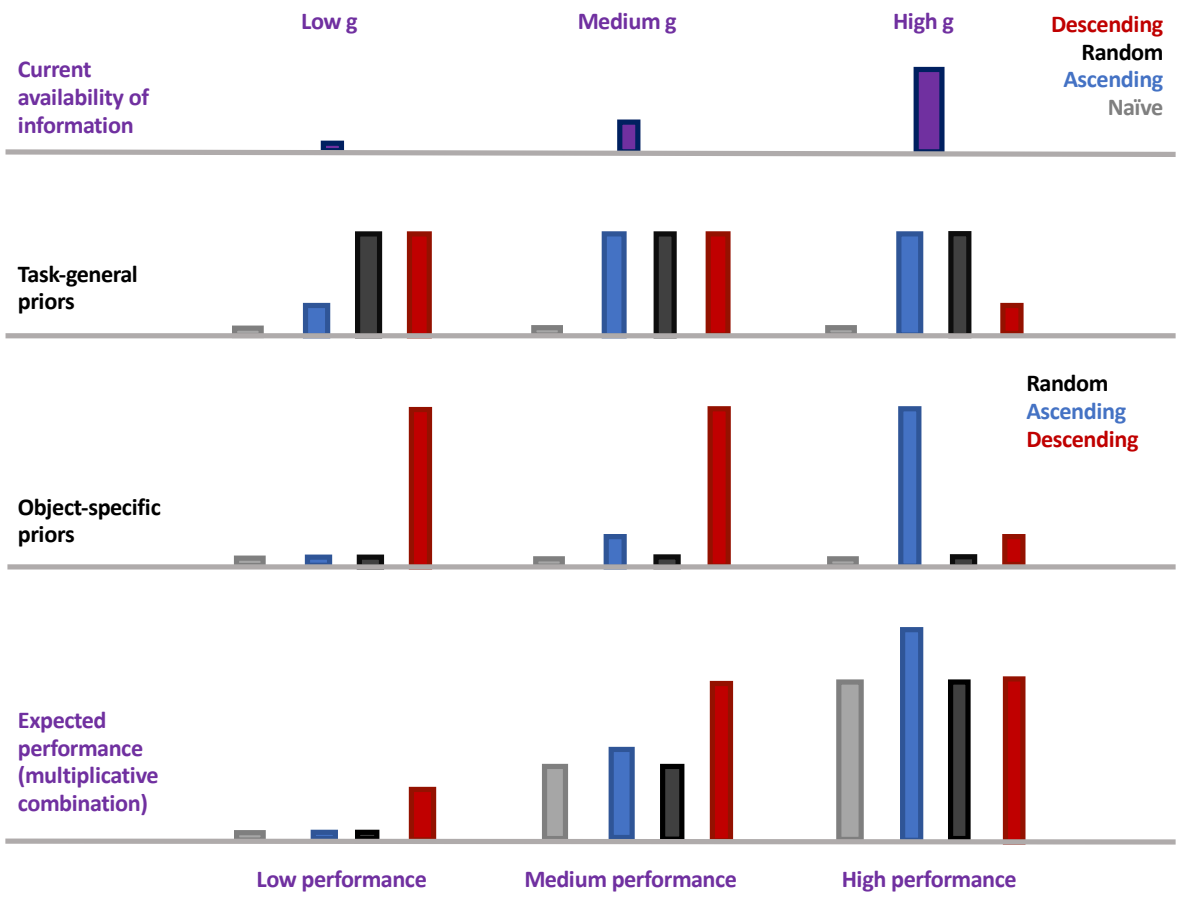


Figure 2.11: Updated model of participants' access to information *g* (top row), task-general (middle-top) and object-specific (middle-low) priors, as well as expected performance (bottom), for low (left column), medium (middle column) and high *g* levels (right column). *Descending* participants are depicted in red, *Ascending* ones in blue and *Random* ones in black. Grey bars correspond to a hypothetical group of naïve participants building no priors. Higher bars signify higher values.

2.4.2. Randomness impairs participants' ability to use specific but not general priors

The most striking result from this study concerned the *Random* group, whose behaviour did not follow our predictions. Although predicting the stimuli's *g* level did not itself carry helpful information for the recognition of the objects depicted, randomising this dimension of the stimuli across blocks appeared to have a dramatic impact on participants' ability to integrate the information that they saw, build object-specific priors, and guide their exploration accordingly. It

is not the first time that randomness is shown to impair behaviour, thus our updated model is compatible with current views on randomness. However, randomness was mostly studied in the context of visual search and had never been shown to impair even a process as basic as single object recognition.

2.4.2.1. Randomness and attention capture

One of the main experiments clearly evidencing an effect of randomness on visual performance, using randomised dimension is orthogonal to the task's goal. In a visual paradigm, Zhao et al. (2013) presented participants with shapes in different locations of the screen, whose order was predictable at some locations and random for others. This was followed by a static visual search task unrelated to the preceding shapes' presentation. They found that participants were faster to gaze at targets presented in the regions that were previously associated with predictable shapes compared to the regions where they had appeared randomly, although these events were unrelated in the sense that they did not help to predict one another. Zhao et al. suggested that this effect was due to predictable events' saliency, resulting in participants' attentional capture and facilitating processing of related events (in their case, spatially related), a view shared by other authors (Turk-Browne et al., 2009). Indeed, the detection and processing of predictable events has been shown to be both faster and more accurate than unpredictable events (e.g. Correa & Nobre, 2008; Rohenkohl et al., 2012; Bendixen, 2014; Barnes & Jones, 2000).

2.4.2.2. Randomness and computational demand

This notion was later questioned by Southwell et al.'s (2017) in a series of auditory experiments. They used two types of rapid sequences of tone pips, arranged either regularly or randomly, and concurrently (binaurally) presented these streams to participants, asking them to detect when a tone was skipped in either sequence. The authors showed that having participants track both types of sequences made the target location (in a random or a regular stream) irrelevant to performance. Rather, it was the mere presence of random information that appeared to impair participants' ability to detect missing tones in either stream: tracking a stream of each type was easier than tracking two random streams, and harder than tracking two regular ones. They concluded that neither predictable nor random information seems to capture attention more than the other, but that differences in performance between regular and random information streams rather reflect differences in cognitive load or computational demand: random

information is more distracting because it is more demanding. Indeed, processing irregularities is particularly demanding as it constantly generates prediction errors when compared to the observers' expectations, which requires a constant update of their model, while regular stimuli can be easily explained away by a predictive rule that does not need to be updated with each stimulus, and thus requires less resources.

2.4.2.3. Current study's addition to the argument

In our case, only object-specific priors were shown to be impaired, while task-general priors remained functional. It seems then that the task's randomness did not impact all types of information equally, which questions the idea of a general attentional capture of regular compared random stimuli, while remaining compatible with the hypothesis of an extra cognitive load associated with random events. We propose that participants remained able to encode information relative to their building of a general model of their environment (the task) while being unable to store detailed information about specific items of this environment (the objects). This appears as an evolutionarily effective behaviour: in a cognitively taxing context when observers are involved in the demanding task of trying to predict items escaping their expectations, encoding items' specificities in the absence of a general rule for how these items are appearing, appears secondary. Orhan and Jacobs (Orhan & Jacobs, 2014) proposed that unpredictable stimuli such as shapes that do not predict colour provoke a "model mismatch" between participants' general model of the world from their long-term priors (e.g., bananas are usually yellow) and the information that they are currently experiencing. We add that the cognitively demanding resolution of this mismatch appears to be taking precedence over other types of computations, resulting in participants' poorer performance in random contexts, linked to reduced specific but not general priors building.

2.4.3. Advantages and limits of the Dots method

The *Dots* method was used as a powerful tool for the investigation of visual exploration patterns. *Dots* stimuli are made of the same local elements, the precisely controlled displacement of which gradually introduces information about a hidden object, while keeping the overall statistics of the image constant. This allows for the visual exploration of different objects to be compared in a precise, quantitative manner. Moreover, these stimuli enable a particularly fruitful investigation of visual exploration behaviours as they force participants to explore extensively through

sequences of fixations and saccades, thereby slowing down the otherwise extremely fast process of object recognition. Another advantage of the method is that it offers quantitative measures of information at any location within the stimulus. Because the *Dots* stimuli are all made of the same local elements, *physical information* (LDD) can be quantified and compared across objects and visibility levels. In addition, *Dots* stimuli also offer the possibility to access the *hidden information* (LCD), reconstructed from the objects' source image. By increasing a single parameter's value (*g*), this information is gradually transferred into the stimuli as physically available information (LDD). The *physical* and *hidden* information content at locations fixated offers an indication about how participants make sense of the stimulus, combining the physically available information with their priors to reconstruct the hidden information.

However, these stimuli are made of altered images of objects, and thus introduce an unnaturalistic component. A validation using more naturalistic stimuli would thus strengthen the results. Moreover, *Dots* stimuli destroy high spatial frequency information i.e., fine details. While this is an advantage on the hand, since it reduces the informativeness of each location fixated and pushes participants to explore more extensively, it also limits the types of images that can be presented and recognised. For example, precise details for fine recognition are most often needed for the recognition of individual objects rather than generic categories. Additionally, images with backgrounds cannot be distinguished efficiently using *Dots* stimuli. Furthermore, information content such as colour or texture is also lost.

2.5. Conclusion

Our results indicate that priors already guide basic visual information sampling for a process as fundamental as single object recognition. We show that depending on their priors, participants do not sample the same information: from a more general standpoint, this suggests that priors can cause people to experience the world in fundamentally different ways because at a very basic level, they already sample different pieces of information, possibly coming to different conclusions when integrating it. Furthermore, the general structure through which the information was presented seemed to influence this fundamental process of guiding exploration with priors: randomness, even when introduced in a dimension orthogonal to the task's goal, destroyed this ability to guide exploration through specific but not general priors. Participants in the randomised paradigm behaved seemingly naïvely despite their intermediate access to information. These findings have important societal implications for how we structure information in everyday situations such as teaching or policy making, and stress the importance of catering for different stages of learning and presenting the information in a structured manner. This last point raises the question of how learning occurs at a different stage in life, rather than only at a different stage of acquiring experience throughout a task. Does learning over development rely on the same basic bricks than learning in fully developed adult individuals? Can young infants actively select and process information too?

Chapter 3: Neural markers of 10-month-old infants' processing of distractors

3.1. Introduction

3.1.1. Gamma oscillations as a central mechanism for visual perception and selective attention

The ability to selectively process information is an essential aptitude in a world where sources of information are numerous and everchanging. How this capacity is achieved in the brain is an extensively studied topic, for which models are emerging. One influential mechanistic account of how this happens in human adults and animal models such as rodents, is the hypothesis of communication by coherence (Fries, 2005, 2015), which posits that information exchange and information selection in the brain rely on synchronised oscillatory activity (Fries et al., 2007).

Gamma oscillations represent a particularly interesting brain rhythm for information exchange in the brain because they are localised both in time and space: they can be the vector of single units of information at the neural network level (Fries, 2009). Gamma oscillations are thought to arise from an interplay between excitatory (pyramidal cells) and inhibitory (interneurons) cells in the cortex (Börgers & Kopell, 2003) or by the interactions of interneurons with an excitatory background (Bartos et al., 2007). The properties of excitability and connectivity that characterise populations of cortical neurons bound them to fire in a rhythmic, gamma-range oscillatory manner. Indeed, given sufficient input drive, an initial firing of disorganised neurons rapidly entrains a push-pull mechanism through which the excitatory and inhibitory populations of neurons sequentially activate/deactivate each other, feeding the cyclic activation of the network and the creation of a gamma oscillations (Moca et al., 2014). For a downstream neuron receiving input from an upstream population of synchronised pyramidal cells, each cycle of the gamma

oscillation represents a time-window during which all neuronal firings are treated as one event: this is the binding-by-synchronisation hypothesis (Singer, 1999). Neurons that fire together are hereby bound into one unit of information at a given time, referred to as a cell assembly. The upstream gamma oscillation of a cell assembly can then produce the firing of downstream neurons with high efficacy because the excitatory volley is confined within a narrow synchronization window (Bruno & Sakmann, 2006). Similarly, the action potential of the downstream neurons will get included into a new cell assembly: neurons firing in the same gamma cycle will be integrated together as one unit of information by the next receiving neuron, and so on. This process forms a hierarchical, convergent system for conveying information in the brain that closely resembles what has been described at the anatomical and functional level. Notably, in the visual cortex, where the neuronal organisation has been described the most extensively in the brain, the information stream starts with the encoding of space localized information in lower visual areas, before getting more and more global and integrative in higher visual areas (Tanaka, 1996). Brain recordings from various animal species from the 80's and 90's showed that high-frequency gamma-band oscillations indeed play a fundamental role for the processing of visual information in the brain (Eckhorn et al., 1988; Gray & Singer, 1989; Engel et al., 1991; see Singer, 1999 for a review). This was also later evidenced in the human brain (Wyart & Tallon-Baudry, 2008) and more generally for sensory processing in the brain (Bauer et al., 2006; Gross et al., 2007).

To sum up, following the binding-by-coherence hypothesis, neurons that respond to the stimulation of one input will fire together in the gamma range and form a cell assembly, but won't synchronise with neurons that respond to another input, which will instead be included into a separate cell assembly. Importantly, each neuron can participate in multiple cell assemblies (Engel et al., 2001). Given a stimulus that contains several objects to encode, for example a natural landscape, the activity of several cell assemblies will encode the different objects and move up the hierarchy of the visual stream, eventually converging to the same higher-level neuron, when it will compete for being represented (Luck et al., 1997). In such cases, the gamma oscillation offers a possibility for one of the representations to win over the other one through enhanced synchronisation of the cell assembly: this both facilitates the communication of the synchronised assembly's activity, while also dismissing activity that is not aligned with it, hereby

blocking other representations. This winner-takes-all property, called biased competition through enhanced synchronisation (Reynolds et al., 1999), represents a mechanism for selective attention that relies on gamma oscillations (Fries et al., 2001).

Selective enhancements of gamma synchronisation for neurons driven by the stimulus subject of the attention focus have been described in studies of intracortical recordings in monkeys (Fries et al., 2001; Taylor et al., 2005) and rodents (Kim et al., 2016). Moreover, this has also been linked to behaviour in a study from Womeldorf et al. (2006) in monkeys, indicating that when behaviourally relevant information was processed by neurons with increased gamma-range synchronisation, reaction time decreased. This framework thus provides a comprehensive and realistic model for linking neurons firing, neural networks activity, and behavioural responses into one mechanism for implementing selective attention in the brain (Buzsáki & Wang, 2012; Fries, 2009; Jung & Carlén, 2021).

Another layer can be added to the model when considering gamma-band oscillations in relationship to slower rhythms in the brain. Just as neurons' firings occurring at the peak of the gamma cycles are amplified, gamma activity occurring at the peak of a low-frequency oscillation will also be amplified, while gamma activity at the troughs of slow oscillations will be poorly communicated (Fries, 2005; Plenz et al., 2021). Gamma modulations by slow rhythms have been described for numerous oscillations, including theta which is usually defined around 4-8Hz (Bragin et al., 1995; Buzsáki et al., 1983; Colgin et al., 2009), alpha around 8-12Hz (Cohen et al., 2008; Palva et al., 2005), slow around 1-2Hz (Hasenstaub et al., 2005; Isomura et al., 2006), ultraslow under 1Hz (Leopold et al., 2003), delta around 2-4Hz (Lakatos et al., 2005) and spindle around 12-15Hz during sleep (Peyrache et al., 2011). Thus, although a more extensive literature exists for the coupling of gamma oscillations with theta range oscillations, the coupling mechanism of gamma oscillations with slower oscillations does not seem to be specific to one range but to be ubiquitously found with many slow oscillations, depending on the context (Buzsáki & Wang, 2012). *Cell assemblies* (gamma-synchronised networks) firing within the same slow cycle are hereby bound into larger *assembly sequences* (slow wave-synchronised networks) and form higher-order units of information. Cell assemblies and assembly sequences have been referred to as the neural "letters" and "words" of the brain's language, respectively (Buzsáki, 2010). While fast gamma-band oscillations are observed over small populations of nearby neurons, slower

rhythms typically occur over larger regions and due to their slower cycles are better suited for travelling across distant sites and connecting brain regions (Steriade, 1999; von Stein & Sarnthein, 2000). Such slow rhythms are good candidates for conveying top-down information to sensory sites in order to modulate the local, high-frequency sensory information driven by bottom-up signals. Following the same metaphor, this process has been described as a "reader-initiated" mechanism for information exchange (Sirota et al., 2008), that is to say, a mechanism also useful for implementing selective attention in the brain.

3.1.2. Oscillations and selective attention machinery in the infant brain

As infants learn about the whole world around them and often without explicit guidance to externally help them filter information, being able to select which information to process is a crucial skill for them. What neural machinery is in place early on in life to support such processes is an important question that needs to be answered for improving our understanding of how the brain supports attention selection both in general and in the special case of infants. Whether the mechanism of complex interplay between fast and slow brain rhythms described above, is available to infants is a question that remains to be elucidated.

Building on the adults' literature, work in infants has shown the importance of gamma-band oscillations for visual processing in this younger population as well (Csibra et al., 2000 with 8-month-olds). Gamma oscillations have been shown to not only subserve the representation of visual objects in young infants during the first year of life, but also to be maintained during the occlusion of objects (Kaufman et al., 2003, 2005 with 6-month-olds) and to be modulated depending on whether the infants were familiar or unfamiliar with the objects represented (Pomiechowska & Gliga, 2021 with 12-month-olds). This suggests that the encoding of information through gamma activity in infants is fine-tuned to the information at hand and likely uses mechanisms of information selection such as those described in the former section. Furthermore, modulations of gamma activity in infants have been shown to be linked to perceptual learning and attention (K. A. Snyder & Keil, 2008). In a visual habituation study, Snyder & Keil (2008) brought the first evidence that gamma activity modulations relate to perceptual learning in 6-month-old infants. They showed that occipital gamma activity decreased with repeated presentations of faces but not toys, which was linked to better orienting to new faces.

Yet, whether changes in gamma activity are mediated by couplings with slower frequency rhythms in infants remains to be shown.

Converging evidence suggests that infants' development of visual attention undergoes significant changes throughout their first year of life, and that they acquire the basic blocks for a functional attention network in the second half of their first year of life (see Hendry et al., 2019 for a review). Notably, Lawson and Ruff (2004) showed that sustained focused attention, when infants actively examined objects both visually and manually, increased during the second half of the first year of life and Xie et al. (2018) showed that neural underpinnings of attention, namely changes in slow-range alpha/theta synchronisation, correlated with heart-rate defined sustained attention in 10- to 12-month-olds but not in 6 to 8-month-olds. Infants' visual processing and attentional machinery appears to come into place around 10 months of age, suggesting that 10-month-old infants might already make use of a complex oscillations interplay to support selective visual attention processes.

One way to study visual attention in infants is to look at their engagement, or distractibility, via the time that they spend on the task (Colombo & Mitchell, 1990). While infants' tendency to disengage with stimuli might seem impractical at first sight, it is a key advantage for the direct study of their intrinsic engagement. Moreover, social norms push adults to follow instructions and complete an experimental task regardless of their interest, which makes it difficult to use this measure in adults. In contrast, infants' social awareness and conformity isn't yet developed enough for this social pressure to influence them, making young populations particularly suited for the study of intrinsic distractibility and its development.

3.1.3. Current study

The 10-month age-point thus appears as the earliest developmental point when infants visual, attentional and fast oscillatory machinery is in place and functional enough to be further investigated, which motivated us to look into the mechanisms and individual differences in information selection at this age point. Here, we used a paradigm developed by Piccardi et al. (2020) to investigate individual differences in 10-month-old infants' intrinsic engagement and information selection, as well as the role of their fast gamma-range activity in this process. In this paradigm, infants are shown 10 repetitions of a 40-s video clip (an animated cartoon), randomly interrupted by 13 100ms- black-&-white checkerboards (plain stimulus). The checkers interrupt

the video and thus create a conflict between the processing of the catchy video and the plain distractors: to maintain attention to the video, infants have to inhibit the processing of these distractors.

Piccardi et al. (2020) showed that the amplitude of frontal theta oscillations before the appearance of the checkers, and occipital Event-Related Potential (ERP) P1 peak in response to the checkers, were inversely related: there was an increase in theta as the video was repeated, and a decrease in P1. This was interpreted as these measures capturing a trade-off between becoming engaged and learning about the video (theta) and decreasing attention to the checkers (P1). These measures were put in relation with infants' level of visual seeking behaviour, as reported by parents in a questionnaire. However, neither the P1 nor the theta amplitude were predictive of the visual seeking individual differences when taken separately, but only when looking at individuals' theta-P1 amplitude trade-off: given the infants' change in theta amplitude, a larger increase in P1 peak than predicted by the decrease in theta amplitude with the repetition of the video was a marker for higher visual seeking. Piccardi et al.'s results (2020) suggest that looking at brain rhythms not only on their own but also in relation to each other is important for understanding individual differences in how infants engage with information.

While Piccardi et al.'s analysis (2020) focused on understanding how attention was distributed during the task, we wanted to investigate how individual differences in neural activity translate to individual differences in engagement to the task. Given the above considerations on the role of gamma-range activity for infants' attention and engagement, we focused on this response on its own and in relation to slower rhythms. Piccardi et al.'s paradigm (2020) presents two main advantages for this investigation. 1) The researchers noted important individual variations in infants' engagement with their paradigm. While they actively worked on keeping the infants on task with external motivators, we decided to make use of the variations in distractibility exhibited by infants and let them disengage spontaneously. This enabled us to mitigate the effects of external unmonitored factors on infants' attention and investigate more directly their individual differences in intrinsic engagement with their neural correlates. 2) The high contrast and regularity of those distractors ensures that the stimuli elicit a strong sensory response in the visual brain (Leguire & Rogers, 1985). Maximising the brain responses with such stimuli brings the opportunity to unveil a great range of individual variations in the neural response. It also helps

with issues of signal-to-noise-ratio. This was crucial in our case due to the fact that infant recordings typically present high measurement error (Frank et al., 2017) and few trials, but also that high-frequency rhythms are particularly difficult to detect in extracranial signals because of the filtering properties of brain tissue, dura, skull and scalp (Moca et al., 2021).

A study of intracortical recordings (ECoG) in the visual cortex of adults implanted for medical purposes further showed that regular stripy or checkerboards (plaid) patterns elicit strong gamma oscillations in V1 (Hermes et al., 2015), which was also shown previously in intracortical recordings of monkeys' early visual areas (Lima et al., 2010), suggesting that the stimuli might similarly elicit strong gamma responses in infants as well. Finally, intracortical recordings in macaque monkeys have shown that increased gamma power in V1 were linked to enhanced coherence in neuronal firing rates (Jia et al., 2013), which is thought to lead to more efficient information processing (Singer, 1999). Thus, we use the strength of the gamma occipital response as a proxy for how strongly the participants process the stimulus.

To investigate the association between taken fast gamma responses and slow brain rhythm, we chose the visual ERP evoked by the checkers as measure of slow rhythm. This was done in the interest of noise-handling, and because Piccardi at al. (2020) demonstrated the importance of this signal for how infants engaged with the stimuli. The ERP is not a brain oscillation per se but a potential i.e., the collapse of oscillations. It presents several regular peaks roughly every hundreds of milliseconds, which makes it akin to a slow ~10Hz oscillation. The ERP is representative of the coordinated activity of large networks of neurons (Makeig et al., 2002), similarly to slow oscillations (von Stein & Sarnthein, 2000). However, while in the time-frequency (TF) power spectrum, oscillations time-locked to the stimulus are represented regardless of their phase, in the ERP, by contrast, only evoked (time-locked, phase-locked) activity contributes to the response and induced (time-locked but non-phase-locked) activity averages out. Thus, the ERP arguably contains different information about how the network oscillates in response to the stimulus compared to TF power activity. However, the fact that it is time-averaged which makes it an exclusively evoked signal, also makes it particularly robust to noise, which is once again a considerable advantage in the context of the current study. Furthermore, the ERP is known to be particularly large for the type of distractors used in this study (see Odom et al., 2010 recommandations for using them in clinical work), again helping to increase the signal-to-noise

ratio. For all these reasons, we chose the ERP as the best suited candidate for studying slow, large-network activity in a robust manner with our paradigm.

In order to measure the attentional trade-off infants face between processing the videos and the distractors, we investigated the dynamical regulation of the *fast gamma rhythms* in response to the distractors, in terms of both its *strength* (representing *local attention selection* at the cell assembly level) and its *alignment* with the slow ERP rhythm (representing more *global attention selection* at the assembly sequence level). Furthermore, to characterise potential individual variations within our general population sample, two groups of participants were compared: the group of infants who did all the experiment (*Did All*) and those who stopped before the end (*Stopped*). Infants were allocated to either group after they performed the experiment, based on their intrinsic level of engagement and attention to the stimuli. Finally, we analysed the gamma response over two different sets of trials. Firstly, to investigate the dynamics of the response over the experiment, we divided the trials into 3 stages and compared the gamma response for each stage. This came with the drawback that different infants completed different number of trials, resulting in stages of different trial lengths depending on the infants, so we downsampled the trials of the infants who *did all* to make the groups more comparable. However, there remained the issue that infants who stay until the end were exposed to more information than those who stop before. Another analysis was performed over the start of the experiment only, before participants started to drop out, such that all infants were exposed to the same number of trials for this trial set. This enabled us to compare groups over the same trials, and at the start of their exposure to the stimuli.

Questionnaires were also collected both at test (Infant/Toddler Sensory Profile: ITSP) and at a 7-month follow-up point (ITSP and Early Childhood Behavior Questionnaire: ECBQ) to provide information on infants' trait individual differences in how they engage with information outside of the lab, as observable (and reported) by parents. The relationship between the neural and the behavioural, questionnaire-based trait measures is the topic of the next chapter and will be omitted for now.

With this paradigm, we investigated both task related changes in the whole group of infants and individual variation. We expected that 1) infants will learn to selectively process the video over the distractors i.e., to block out the distractors, as reflected by a) a decrease of the gamma

response and b) a progressive de-alignment of the fast gamma and slow ERP responses when comparing start, middle and end stages of the study; and that 2) infants who will disengage attention before the end of the experiment will be less efficient at blocking out the distractors, as reflected again by a) a higher gamma response over all stages, as well as at the start of the experiment already, and b) a lesser modulation of the gamma response over time, from stage to stage.

3.2. Methods

This study is based on work from Piccardi et al. (2020): they designed the experiment, recorded data and reported their results in a publication. Here, I was interested in another question, with which the data recorded by Piccardi et al. was not always compatible. Thus, I recorded more data (referred to as phase 2 in the following section) to add to the compatible datasets from Piccardi et al. (recorded during phase 1). The analysis of this new combined dataset focused on other aspects of the data and thus the pre-processing and analysis of this data was entirely performed by me. This recording in two phases also explains some of the data loss detailed below.

3.2.1. Participants

A total of 114 full-term, healthy 10-month-old infants were tested with this paradigm, 31 of whom were included in the present analysis. Testing happened in two phases, each one presenting complications that substantially impeded on the inclusion rate: 1) infants tested in phase 1 were primarily recruited for the study of Piccardi et al. (2020) that allowed for testers to interfere with bubbles, rice cakes and/or speech to motivate the infants, which was incompatible with the aims of the current study (38/48 i.e., 79.2% incompatible sets); 2) intense use of the shared testing platform resulted in inadvertent failure of equipment for a large part of phase 2 (22/66 i.e., 33.3% failures: photodiode failure for 10 infants, video recording failure for 8 infants, file loss for 2 infants, 50Hz interference from faulty monitor for 2 infants). When these technical issues are accounted, additional drop-off rate (due to either poor signal quality – see criteria in section 3.2.4.1 – or infants not watching enough of the study) reaches 42.6% (23/54 dropped out), which is similar to other infant EEG studies, considering that there was no use of external motivators to keep the infants on task.

In phase 1, 10 datasets were compatible with the study (no intervention of the experimenter to keep the baby on task), 1 of which was excluded for poor EEG signal quality and 2 for technical issues (video failure and photodiode failure), while 7 infants were included. In phase 2, 44 datasets were usable, 4 of which were excluded because of technical issues (incorrect net placement for 2 infants, inadvertent door opening for 1 infant and pacifier given during the experiment for 1 infant), while 13 infants did not tolerate the EEG net well and did not participate long enough to be included (less than 2 video repetitions), and 3 infants were excluded for poor

EEG signal quality due to excessive movement artefacts, resulting in 24 infants included from the second batch of testing, and 31 infants included in total (14 males and 17 females, mean age 10 months and 4.0 days, S.D. 9.7 days, min. 9 months and 11.3 days, max. 10 months and 18.8 days).

The target number of usable dataset for this study was 35, such that medium to large effects could be detected. However, this number was not reached as the COVID pandemic hit the UK towards the end of the testing session (see COVID impact statement on page 4 of this thesis). Still, the number of recordings obtained before the laboratory was shut down was deemed close enough to the target for it not to affect the viability of the study. A sensitivity power analysis was conducted with using the software G*Power (version 3.1, https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower) to verify this claim. Standard default values were used since literature using the same measures and population as in the present study was lacking to inform this analysis more precisely. Values were sat as such: the significance threshold was α = 0.05, the power 1-β = .80, among-measure-correlation 0.5 and there was no sphericity correction (ε = 1). The analysis was conducted for a 2-group by 3-measurements repeated measured ANOVA design, both for the achieved sample size N = 31 and for the targeted sample size N = 35.

The analysis showed that the effect sizes to which the study was sensitive were indeed very similar between the two groups. For the achieved number of participants, the analysis was sensitive to medium (f > 0.25) within-subjects and interaction effects: f = 0.23, F = 3.16 (for N = 35 participants: f = 0.22, F = 3.14) and large (f > 0.40) between-subjects effects: f = 0.42, F = 4.18 (for N = 35 participants: f = 0.40, F = 4.14). The values for the achieved ad targeted sample sizes were very close and not qualitatively different, confirming that the COVID-related changes in sample size did not affect the viability of the study.

All infants were recruited via a volunteer database at the Centre for Brain and Cognitive Development at Birkbeck, University of London. They were born full-term (37-42 gestation weeks and over 2.5kg at birth), were typically developing and had no history of neurodevelopmental disorders in the nuclear family, as reported by the parents. Parents gave their written consent before any data was collected and the procedures were approved by the Ethics Committee of the Psychological Sciences Department at Birkbeck, University of London (approval no. 171805 and 181902).

3.2.2. Stimuli

We used the same stimulus set as in Piccardi at al. (2020, see Figure 3.1). Stimuli were made of two elements: 1) an ongoing 40-s video clip from the Walt Disney cartoon Fantasia, which repeated for 10 times, 2) static, 100-ms black and white checkerboards, presented 13 times per video repetition (12 times, 35s video in repetitions 5 and 10) and 128 times in total. For each video repetition, the inter-stimulus interval (ISI) between each checkerboard was randomly jittered in the window 2-4s. All participants viewed the stimuli at the same timings within each video repetitions, which ensured that they all were exposed to the same information. At each checkerboard presentation, the ongoing video clip paused and resumed immediately after. The video clips were presented with the cartoon's original music and at the centre of the screen in a 22.5cm × 12.5cm area (21° × 12°). The checkerboards were also presented centrally in a 30cm × 30cm area (28° × 28°), with an average luminance of 1.56 cd/m^2 for the black patches and 2228 cd/m^2 for the white patches.

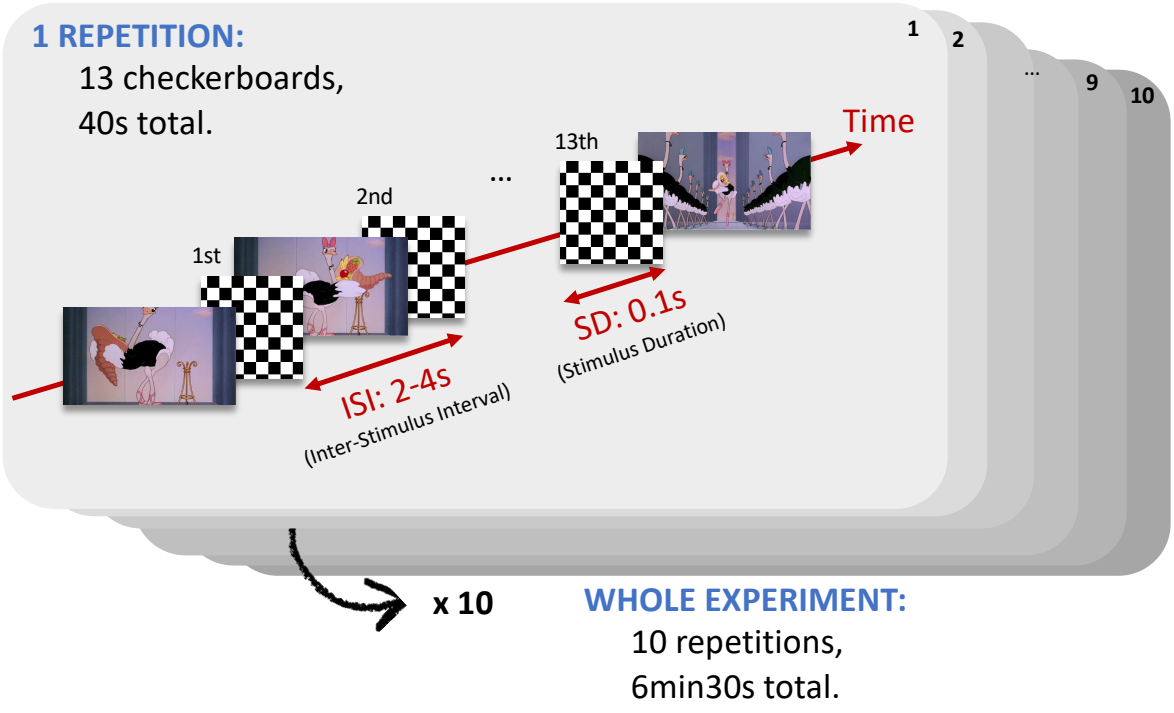


Figure 3.1: experimental design. The whole experiment consisted of 10 repetitions of a video clip (6min30s total). One repetition lasted for 40s and was composed of 13 100-ms black-and-white checkerboard (stimulus) presented every 2 to 4s and interrupting a background video clip.

3.2.3. Procedure

Families were greeted in a playroom with toys, where infants could play with an experimenter while another experimenter explained the study and procedure to the parent, answered their questions and collected their informed written consent. Families were then taken to the testing booth. The experiment took place in a darkened, electrically insulated booth at the Centre for Brain and Cognitive Development's Babylab in central London, UK (Birkbeck, University of London). Infants were sat on their parent's lap, approximately 60cm away from a TV screen (27 inches: 59.77cm × 33.62cm) and wore a wet 128-channel HydroCel Geodesic Sensor Net, connected to a NetAmps 400 amplifier (Electrical Geodesics, Inc., USA), online referenced to the vertex of the net (Cz). Electrical signals and stimulation events were recorded continuously at a sampling rate of 500 Hz (testing phase 1) or 1,000 Hz (testing phase 2) on a computer using NetStation (Electrical Geodesics, Inc., USA). Stimulus presentation was controlled on a different computer using PsychToolBox for Matlab (http://psychtoolbox.org/) and both computers were connected serially. The infant's behaviour was also recorded at 25 Hz using a camera placed just below the screen and framed around their face. The recording experimenter monitored infant's behaviour throughout the experiment using this video. Another experimenter stayed behind the parent and infant in the darkened booth during the experiment for safety but did not interact with them. The experiment lasted for 6min30sec in total but could be interrupted by the experimenter outside the booth if the infant was showing prolonged disinterest: 10s or more of continuous crying, or 30s or more of continuous disinterest, which always included interrupted cries and could include one or several of the following behaviours: not looking, turning away, looking to parent for help, moving a lot, pulling net off. Parents also had the possibility to stop when they wanted to, although the case did not occur. When two parents were present, the second parent stayed outside with the experimenter monitoring the recordings.

Once the experiment was complete, families were taken back to the Babylab's playroom where parents filled in the ISTP questionnaire. They were also offered a Babylab diploma and Babylab T-shirt for their time and were offered a reimbursement of their travel expenses, but received no monetary payment for their visit. Parents were then contacted by e-mail 6 months after their visit to the Babylab for an online follow-up to complete questionnaires (ISTP again and ECBQ). Questionnaire-related data at either time point will be discussed in the next chapter.

3.2.4. Data pre-processing

3.2.4.1. EEG pre-processing

Each individual's EEG data was first screened for artefacts using NetStation software (Electrical Geodesics, Inc., USA) as described below. A bandpass filter between 1 and 220Hz and a notch filter at 50Hz were applied for the purpose of this screening only. Two epochs were created around each checkerboard stimulus onset: a pre-stimulus epoch (-1,000-0ms) and a post-stimulus epoch (0-1,000ms). Automatic artefact rejection was applied on pre- and post-stimulus epochs separately using NetStation's built-in tool with a sliding window of 80ms. For each epoch, channels were marked as bad if they contained variations larger than 2,000uV. Epochs were rejected if more than 25 channels were marked as bad for this trial. Finally, channels were rejected if marked as bad for over 33% of the epochs. Flags for rejected channels and trials were exported to a CSV file to be used in the next steps. 3 participants were excluded at this point because over 20% of their electrodes were bad, and 1 because of excessive movement artefacts.

Raw data was loaded into Matlab, converted to a universal format using EEGlab's toolbox (https://sccn.ucsd.edu/eeglab/index.php) and read into the software EEGProcessor (in-house software developed at the Transylvanian Institute of Neuroscience: https://tins.ro/) for pre-processing. Datasets were cropped 10s around the start and end of the experiment and electrodes that were rejected in NetStation were interpolated using the nearest neighbours. The nearest neighbours were manually selected in order to obtain a symmetrical distribution around the electrode to interpolate, where more than 1 clean electrode was available in the vicinity: 1-6 electrodes were chosen, with an average set size of 3.17 electrodes, S.D. = 1.44 electrodes. Bad electrodes from the face and the outermost ring of the net were rejected permanently because an insufficient number of clean neighbours were available to interpolate them in most cases. Data was then re-referenced to the average of all non-permanently rejected electrodes, and low-pass filtered with a sharp order 5 Butterworth filter at 200Hz. This filter is used to avoid aliasing, which is particularly important before downsampling the data. Thus, a sharp i.e., high-order filter was chosen. Phase 2 data collected at 1,000Hz was then downsampled at 500Hz to match phase 1 data's sampling rate. All datasets were then bandpass filtered with an order 3 Butterworth filter at 1-200Hz and notch filtered for electrical artefacts at 49-51Hz, 99-101Hz and 149-151Hz with an

order 3 Butterworth band-stop filter. This time shallow, low-order filters were chosen to minimise the temporal smearing and ringing effects of sharper filters.

Finally, the individual datasets were loaded into Matlab for further cleaning of artefacts using the FieldTrip toolbox (https://www.fieldtriptoolbox.org/). First, FieldTrip's visual rejection tool was used for manual artefact rejection. This late and more fine-tuned rejection is complementary to the first step of automatic rejection that was formerly applied in NetStation for removing very noisy data, which was needed early on to avoid contamination of clean data during pre-processing. The variance, minimum, maximum, range, kurtosis and z-value of all trials on the one hand and all electrodes on the other hand were plotted as clouds and visually inspected for outlier trials, which were rejected. Finally, FieldTrip's tool for Independent Component Analysis (ICA) was also used to clean the data from ocular artefacts. ICA traces and topographies were visually screened to identify components corresponding to eye blinks or eye movements in both shape and scalp location. Those were taken out of the data when projecting the components back into sensor space. 1-3 ocular components were identified for 28 out of 31 individuals (90.3%, 2 individuals with no component identified), with a mean number of components of 1.55 (standard deviation: 0.85).

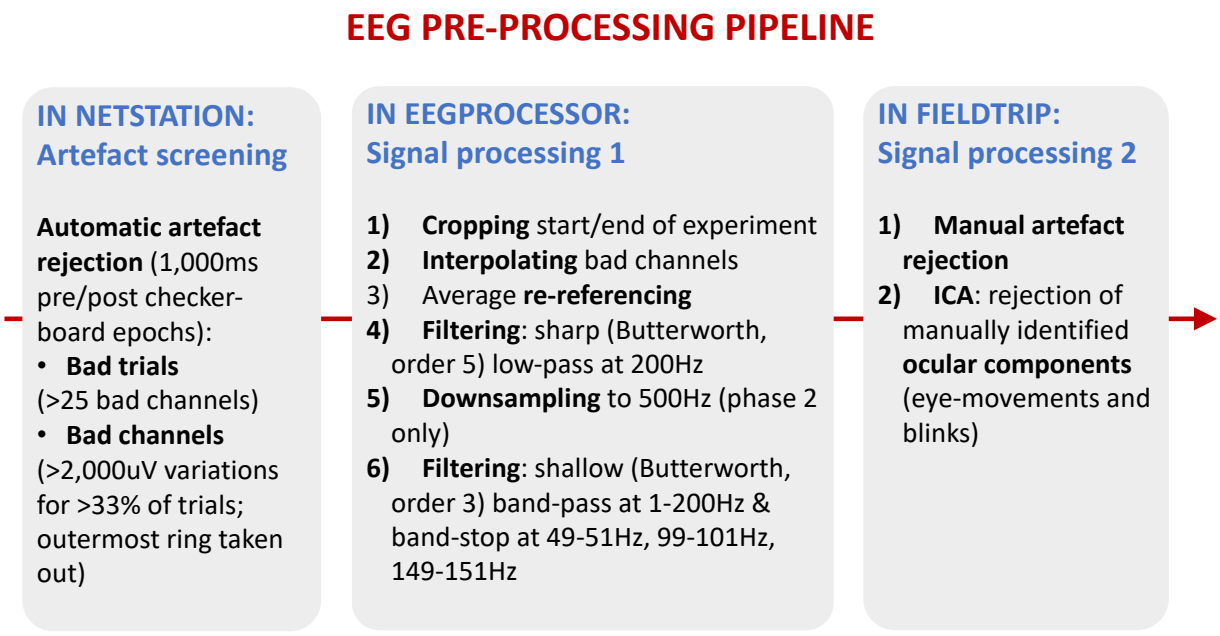


Figure 3.2: Summary of the EEG pre-processing pipeline. The pre-processing was performed in three main steps, each in a different software: 1) NetStation software was

used first for artefact (trial and channel) screening; 2) EEGProcessor was used for signal processing (cropping, interpolating, anti-aliasing, downsampling, filtering); 3) finally, FieldTrip was used for the last signal processing steps (manual artefact rejection and ocular component removal using ICA).

3.2.4.2. Video coding

The video recordings of infants (25 Hz) included a picture-in-picture image of the stimuli on screen at the bottom left of each frame, as shown in Figure 3.3. Each video was manually coded offline, frame-by-frame, using the software ELAN (https://archive.mpi.nl/tla/elan). The coder marked the start and end of infants looks towards the screen, which served to reject trials when baby did not look. Additionally, the coder also marked the start and end of the experiment. The coder also marked the start of the first 5 and last 5 checkerboards' presentation, which was used to re-align the video coded data with the EEG signal. For a randomly picked 16% of the videos (5 infants), coding of the start and end of the experiment as well as the looks start and end was also performed by a second coder to insure the impartiality of the coding. To assess the consistency of the two coders' marking for both of the looks start and end, and of the experiment's start and end, two separate intraclass correlation coefficients (ICCs: McGraw & Wong, 1996) were computed using Arash Salarian's ICC function for Matlab (Salarian, 2022). The coders were found to be highly reliable for both looks ($r = 0.9993$, $p < 0.0001$) and experimental timings ($r = 1.0000$, $p < 0.0001$). There were 2 instances out of a total of 47 looks events when either one coder did not code an event that the other did (and 45 events that they both coded; 4.26% disagreement).

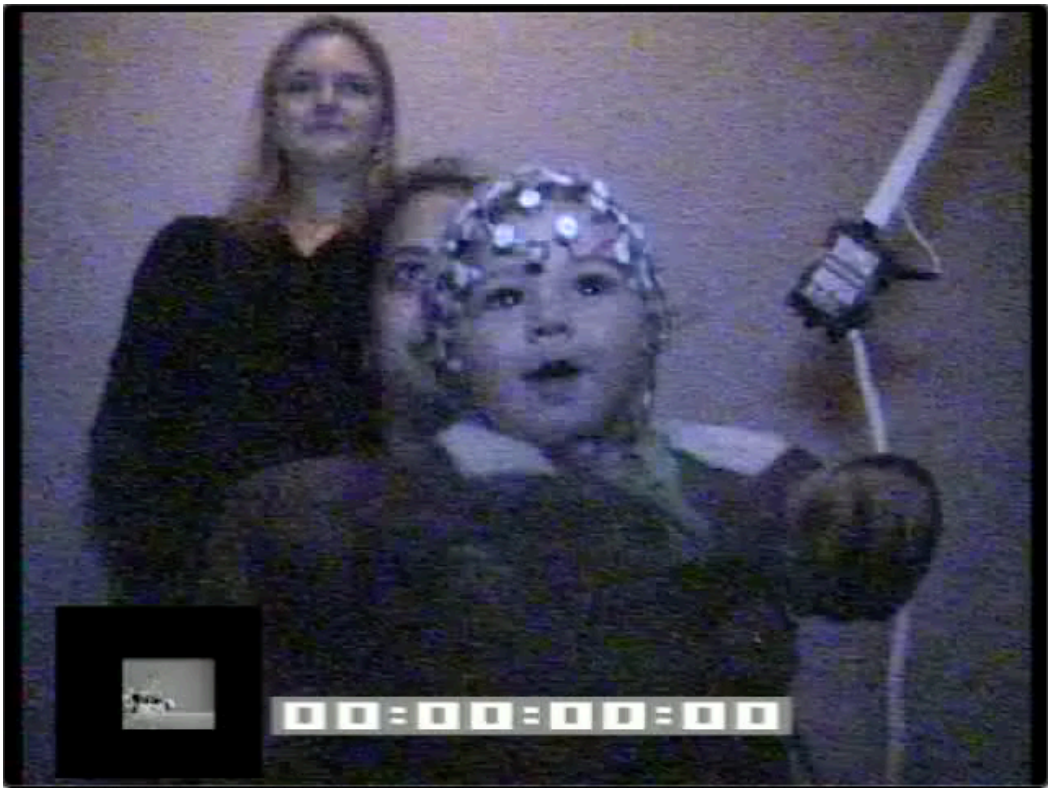


Figure 3.3: Example screenshot of a participant's video recording. A picture-in-picture representation of what the infant is watching on the TV screen is visible at the bottom left of the recording. The timestamp at the bottom of the image is irrelevant. The infant can be seen wearing the EEG net, sitting on their parent's lap in a darkened booth while an experimenter silently stands behind them.

3.2.4.3. Group division

Infants were divided into 2 groups, based on the duration for which they engaged with the experiment (obtained using the EEG triggers for the start and end of the experiment): infants who did all the experiment (*Did all* group), and infants who stopped before the end (*Stopped* group). Gender and age descriptive variables for the whole sample, as well as the two groups of infants, are given in Table 3.1.

Group	N total (F)	Mean age at test (S.D.)	Min. age	Max. age	Mean time-on-task (S.D.)	Min. time	Max. time
Whole sample	31 (17F)	10m, 4.0d (9.7d)	9m, 11d	10m, 19d	4min56s (1min43s)	1min58s	6min30s
Did all	15 (11F)	10m, 4.2d (8.2d)	9m, 11d	10m, 19d	6min30s (0.08s)	6min30s	6min30s
Stopped	16 (6F)	10m, 3.9d (11.4d)	9m, 19d	10m, 16d	3min28s (1min1s)	1min58s	5min14s

Table 3.1: Age and gender descriptive variables for the whole sample and for each group of participants; F stands for female and S.D. for standard deviation.

We found no significant difference between the group who *did all* and the group who *stopped* in terms of the infants' age when they were tested: t(1,29) = -0.095, p =0.925, Cohen's d = -0.034. Thus, we did not further investigate the effect of age in this study. Contrastingly, we compared contingency tables for gender in both groups using a chi-square test, which showed that the proportion of males and females in each group was significantly different: $\chi^2(1,29) = 4.014$, p = 0.045. 10 males and 6 females *stopped* before the end, while 4 males and 11 females *did all* the experiment. The effect of gender was thus further explored in the upcoming analyses, by performing gender splits in addition to group splits.

3.2.4.4. Trial division

Trials were defined as 2s segments centred around the onset of the checkerboards. This trial length enables enough time post-stimulus to observe the arising and the extinction of the visual response to the stimulus, as well as a stable pre-stimulus baseline.

3.2.4.4.1. Definition of the experiment's segments

In order to study the temporal dynamics of participants' responses along the experiment, trials were grouped according to two possible time frames:

1) the experiment's *objective* time course: we used the first 2 of the total 10 blocks, ***video repetitions 1-2*** (26 trials), to isolate trials completed by all, regardless of group, (this was

the biggest segment analysable since the first dropouts occurred in the next video repetition);

2) the participant's *subjective* time course: we used the number of completed trials for each given individual and divided it by 3, in order to obtain 3 ***stages*** of equal length representing the start/middle/end stages of the individual's subjective experience with the study.

The former trial division enabled us to study how participants respond in a time-lapse that is the same regardless of group. This allowed to ask the question of whether individual difference in later disengagement with the task can be predicted from individual differences in how the distractors are processed from the very beginning of the task. The latter trial division enabled us to study changes in participants' responses as their intrinsic engagement with the content of the experiment evolves.

3.2.4.4.2. Downsampling of the trials

For both trial division approaches, trials when the baby was not looking to the screen were then discarded using the video coding data, as well as trials which were flagged as bad during automatic or manual rejection. Infants contributed different amounts of trials per stage or group. To ensure comparable signal-to-noise-ratio, trials were randomly downsampled per individual such that for each stage (or for video repetitions 1-2), the average number of included trials for each group of participants (those who *stopped* before the end and those who *did all*) were matched. There was a maximum of 26 potential trials available for the vid. rep. 1-2 analysis before cleaning and downsampling, and a maximum of 43 trials per stage (for infants who *did all*). Trial number mean and standard deviations before and after downsampling, for each group and each trial division, are shown in Table 3.2.

	Mean (S.D.) number of trials: before downsampling/after downsampling			
Group/Segment	Stage 1	Stage 2	Stage 3	Video rep. 1-2
Stopped	15.06 (6.34) / -	14.25 (7.23) / -	10.75 (5.46) / -	16.69 (3.98) / -
Did all	31.40 (8.82) / 15.07 (0.26)	32.33 (6.75) / 14.33 (0.45)	33.47 (8.86) / 10.93 (0.26)	18.40 (6.29) / 16.67 (4.75)

Table 3.2: Mean (Standard deviation) number of trials before and after downsampling for each group, each segmentation of the experiment (stages 1 to 3 and video repetitions 1-2).

3.2.5. EEG data extraction

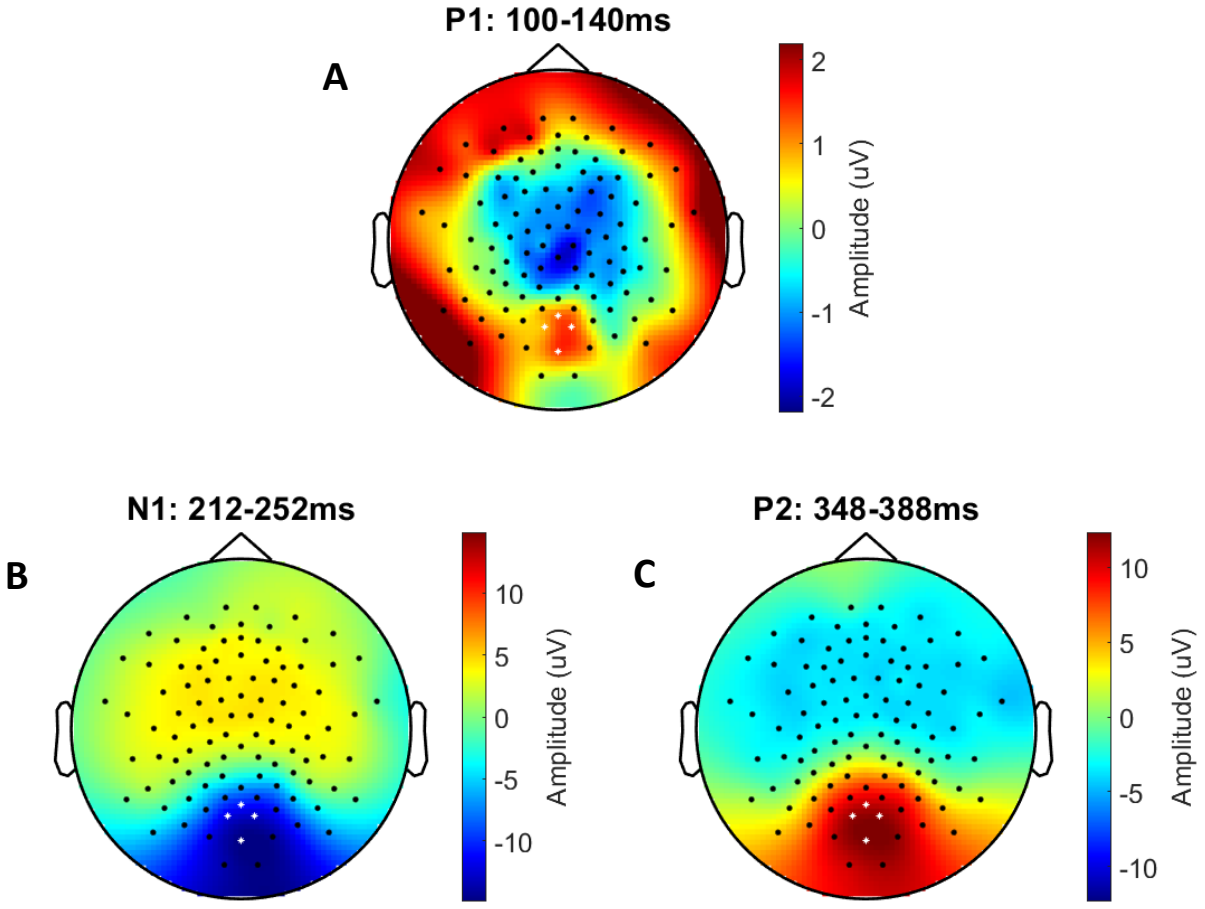


Figure 3.4: Grand average topographies for the 3 main ERP peaks: P1 (**A:** 100-140ms), N1 (**B:** 212- 252ms) and P2 (**C:** 348-388ms). Black dots are drawn at each electrode position, the set of electrodes of interest is represented by white stars.

Data were extracted from the pre-processed and downsampled sets of trials for each segment of the experiment separately (stages 1-3 or video repetitions 1-2). We used 4 occipital electrodes, which sit on the primary visual cortex areas (see Figure 3.4 for a visualisation of the scalp locations): 71 (left Oz-POz), 72 (POz), 75 (Oz) and 76 (right Oz-POz). This area was the most relevant for picking up signals from the visual cortex in order to study infants' processing of our visual stimuli (Gliga & Dehaene-Lambertz, 2005; Piccardi et al., 2020). Because of the low spatial resolution of EEG, it is important to keep in mind that these electrodes are likely to reflect the visual cortex' activity combined with the activity of other regions. We chose several electrodes to account for individual differences in head shape or net placement but kept a relatively small set for the activity to be as specific as possible, which facilitates the study of cross-rhythms couplings. For the purpose of illustration, grand-average topographies of the 3 main ERP peaks were obtained in FieldTrip using the pre-processed, downsampled sets of trials for all stages at once, all participants at once. Topographies were drawn from a 40ms windows centred on the grand-average peak timing (see following section about the ERP extraction) and revealed a focal activity around the set of electrodes of interest for each peak (Figure 3.4). The P1 peak appeared very focal and small in amplitude, possibly due to important individual variations in the cortical folding of the primary visual areas.

3.2.5.1. ERP data extraction

Event Related Potentials (ERPs) were obtained by applying an additional lowpass filter at 30Hz in Fieldtrip, averaging included trials (i.e., trials not rejected during artefact detection nor downsampling) and smoothing the curve using a 7-point moving average. In order to obtain the grand-average topographies presented above, the peak timings of the grand-average ERP were obtained using a min-max methods: the timing of the positive peaks were identified as the maximum in the 200-ms time window 0-200ms for P1 (identified at 120ms) and 200-400ms for P2 (identified at 368ms), and the timing of the N1 peak was identified as the minimum in the time window 100-300ms (identified at 232ms). Those timings served to centre the topographies of the 40ms segments around each peak shown earlier.

3.2.5.2. TF gamma power data extraction

Time-frequency (TF) spectrograms (40-120Hz) of the post-stimulus response to the checkerboards (0-1,000ms) were obtained in the software EEGProcessor by extracting power

data for included trials as follows: 1) we used a Fourier window of 1,000ms to transform the data into frequency space, 2) a Blackman sub-window of 300ms extent with a 5ms step, 3) each trial was de-skewed using a log transform to mitigate the 1/f effect of power in the data, 4) and finally baselined to the 1,000ms pre-stimulus window (during which babies were watching the cartoon) using a pseudo Z-score method (Ciuparu & Mureşan, 2016). Using the log transform before baselining and averaging the data is equivalent to taking the geometric mean of the data which, according to the minimum cross-entropy principle, is the most informative single representation of multiple representations (trials) taken together (Loughlin et al., 1992).

Both ERP and TF data were extracted per individual and per stage (individual level) but also for each stage, all participants of one group taken together (group level), and finally for all stages, all participants taken together, regardless of group (grand-average level). The grand-average data was only used as a qualitative reference to identify a time-frequency window of interest in which the gamma activity was found. The window identified was subsequently used for the analysis of the group level and individual level TF data. The window of interest from the grand-average data was visually identified between 100-700ms and 50-100Hz (see Figure 3.5). Note that this window was identified on the grand average-sample, whole-experiment data. This way, any difference between stage or group was irrelevant to the process of identifying the window of interest, which prevented circular inference while ensuring a data-driven analysis of the data.

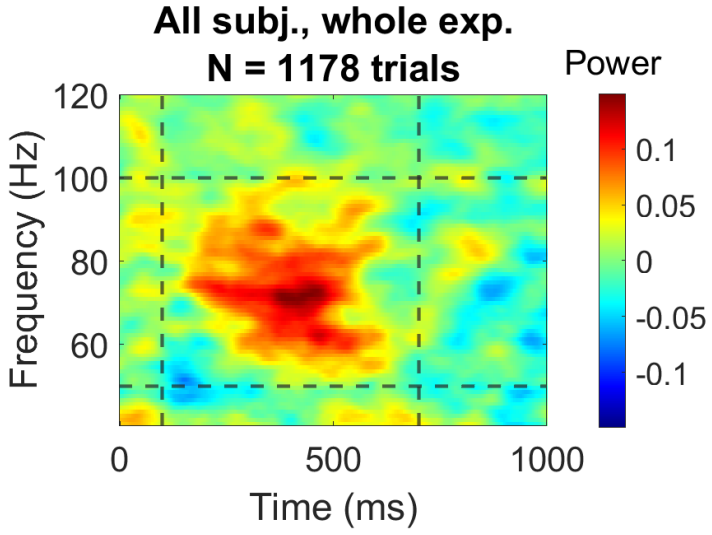


Figure 3.5: TF (time and frequency resolved) spectrogram extracted on the whole sample of participants for the activity in the post-stimulus (0-1,000ms) time window and high-

frequency (40-120Hz) frequency window. Dashed lines represent the time-frequency window of interest in which most the gamma activity was found to occur, namely between 100-700ms and 50-100Hz.

At the individual level, the data was reduced to obtain one value per individual and segment of the experiment, such that multiple comparisons could be readily drawn using repeated-measure Analysis of Variance (ANOVA) designs. To this aim, the maximum value within the frequency window of interest (50-100Hz) was taken for each time point. Taking the maximum reduced the dimensionality of the data from 3 (frequency*time*power) to 2 (time*power), while ensuring to preserve most of the information about the gamma rhythm. Indeed, gamma rhythms have been shown to be mostly located in frequency but to be subject to drifts in frequency across time, which are well represented by taking the maximum power value in a band of interest (Dăbâcan & Mureşan, 2017). The time-resolved maximum gamma response was further reduced by taking its average over the time window of interest (100-700ms). The average best represents the whole trial across time since the gamma response is thought to unfold over time rather being confined to one local time point. This yielded a single value to represent the relative modulation of gamma by the distractors, relative to the ongoing videos.

3.2.5.3. Scaled Correlation analysis (SCA) between gamma and ERP data

Finally, we examined the relationship between types of responses: individuals' ERPs (a global, network-wide response) and their time-resolved max. gamma response (a fast, local activity). Scaled Correlation Analysis (SCA) was used to investigate the ERP-gamma relationship (Nikolić et al., 2012). SCA measures neuronal synchrony between two time series, without assuming periodicity (as opposed to spectral techniques such as coherence estimates). This property makes it particularly robust to temporal offsets, jitters, changes in signals' temporal relationship but also to the absence or the presence of multiple regular rhythms characterising the co-occurrence of the two signals analysed, which helps to increase the signal-to-noise-ratio. Furthermore, because gamma activity is fast and occurs in short-time bursts, fast modulations are of particular interest here, but they are likely to be riding slower, larger co-variations that mask the faster processes of interest. SCA is a particularly well-suited method for the study of fast co-variations between signals because it uses a *scaling parameter* to evaluate the correlation between signals on small time windows within the analysis window: making use of restricted sampling properties, this

prevents slow co-variation components from contributing to the final coefficient obtained and avoids masking fast scale modulations (see Figure 3.6). The final correlation coefficient obtained is the average of the Pearson correlation coefficients obtained for each sliding, non-overlapping scale segment of the analysis window. Although we did not make use of this property in the current study, SCA interestingly yields more than 1 correlation value between the signals: signals are shifted left and right relative to each other, and correlations values are obtained for various time-lag value within a defined correlation window. This is useful to unveil potential regularity in how fast coupling processes occur over time.

Here we used a sliding scale segment of 50ms, resulting in the removal of components under 20Hz due to restricted sampling, with a correlation window of 300ms, over an analysis window of -200 to +1,000ms around stimulus onset (this corresponds to our 100-700ms window of interest, with a 300ms padding before to enable the computation of correlations up to +/-300ms lags).

Because the ERP is a potential measure while the gamma response is a power measure, neuronal firing will be reflected by activity of opposite signs: when neurons fire, this creates a *negative* extracellular *potential* (Jurjuţ et al., 2019; Plenz et al., 2021), which has a *positive power* representation because power is a squared value, bound to be positive. Thus, **alignment** between these two signals will be represented by a **negative correlation** value, while **de-aligned** signals will be **positively correlated**.

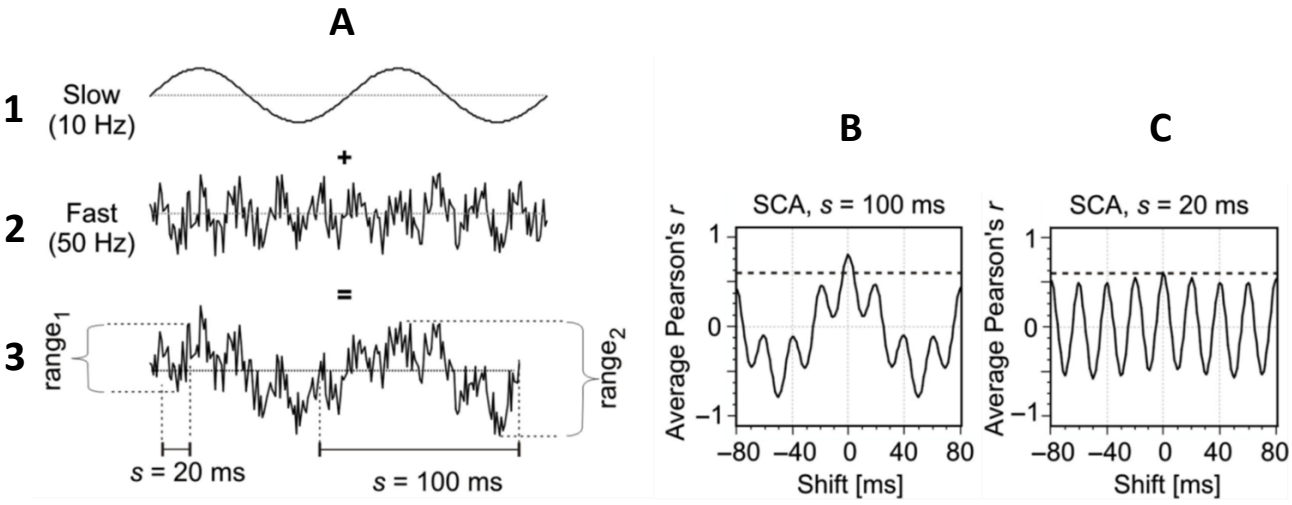


Figure 3.6: computation of the SCA, adapted with permission from Nikolić et al. (2012). **A.** Simulated signals for demonstration purposes. The main signal (**A.3**) is composed of two

continuous signals: a regular, slow component (**A.1**) and a noisy, fast component (**A.2**). **B.** Result of the SCA when using a scale segment (s) as large as than the period of the slow component: fast co-variations are partly masked by the larger slow co-variation. **C.** When the SCA is computed using a scale segment (s) reduced to the period of the fast component, the contribution of the slow component is removed, and the correlation properties of the fast component are revealed much more accurately. The horizontal dashed line indicates the correlation inserted into the fast components, which should be recovered by the analysis.

3.3. Results

We looked at participants' gamma activity to test our hypotheses that 1) participants learned to block the distractors over the course of the experiment i.e., their gamma activity in response to the distractors decreased stage-by-stage; 2) participants who *did all* the experiment were more efficient in ignoring the distractors i.e., their gamma activity in response to the distractors was a) generally lower and b) less modulated over the experiment.

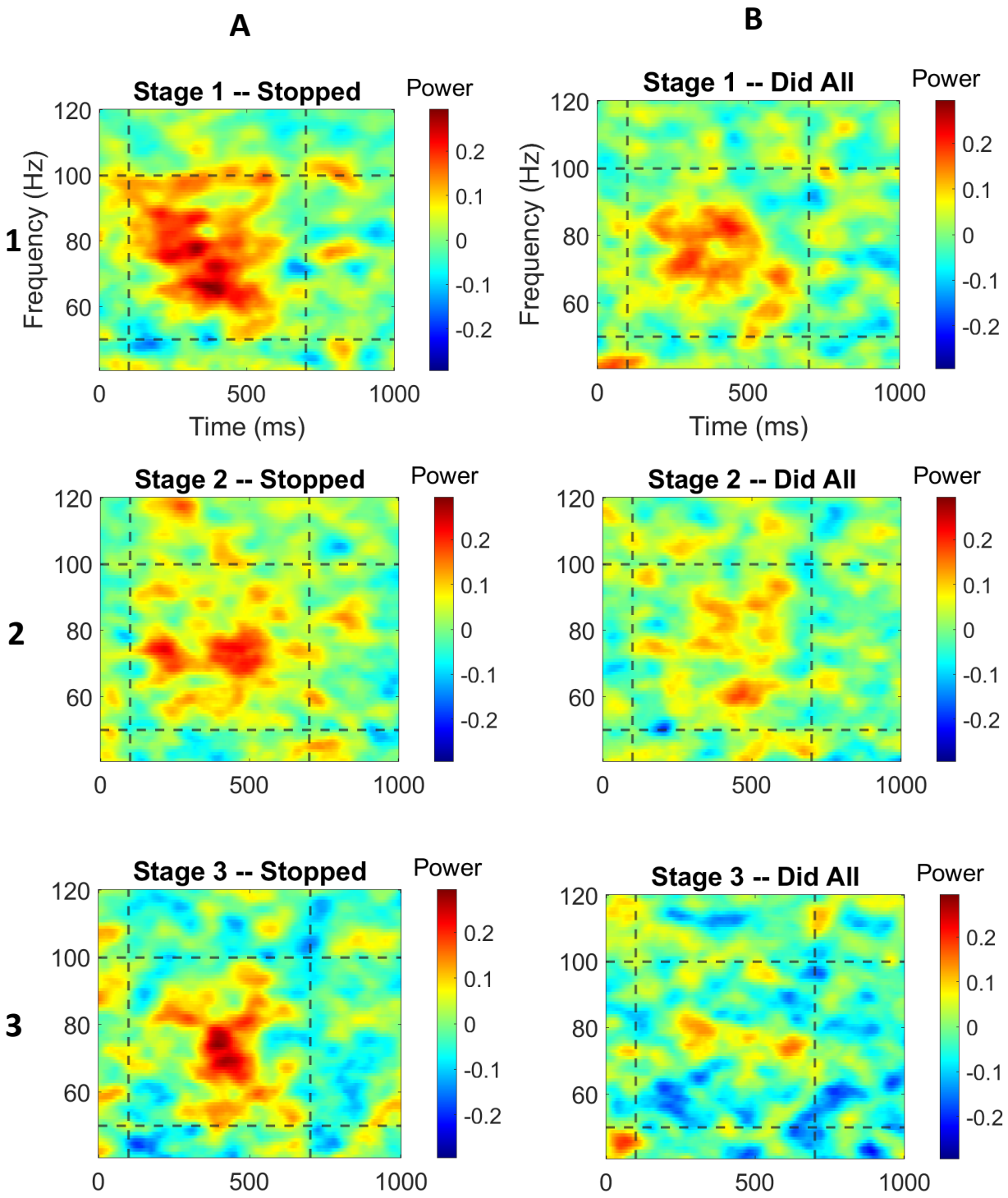
3.3.1. Group-level gamma response to the distractors

We first analysed gamma activity extracted at the *group level* (all participants' included trials taken together, section 3.3.1) to identify *consistent group dynamics* during the experiment, before investigating *individual dynamics* in a further analysis to come (section 3.3.2)

Group-level dynamics can be somewhat ambiguous as they can result from two different types of individual-level dynamics. On the one hand, a decrease in the individual responses will be reflected by a decrease in the group activity. On the other hand, individual responses of similar strength, but with more individual variations in the time and space location of the participants' responses will also produce a decrease in the group response, this time due to a decrease in the *consistency* of the group-level activity (i.e., a more individualised response) and not a decrease of the response *strength*. This can be disentangled by looking at the group-level data in relationship to the individual-level data, to be described in the next section 3.3.2.

3.3.1.1. Groups TF (time and frequency resolved) gamma activity per stage

In order to investigate changes over the course of the experiment, we analysed groups' brain response over 3 stages (Figure 3.7). Stages corresponded to individuals' total number of completed trials divided by 3 and randomly downsampled such that the groups' average trial number per stage were matched (see Methods section 3.2.4.4). Note that the use of a log transform before baselining and averaging the data ensured that only features present across multiple trials (i.e., consistent within the group) would be prominent in the group spectrograms.


A
B
1
2
3
Stage 1 -- Stopped
Stage 1 -- Did All
Stage 2 -- Stopped
Stage 2 -- Did All
Stage 3 -- Stopped
Stage 3 -- Did All
Power
Frequency (Hz)
Time (ms)
120
100
80
60
0
500
1000
0.2
0.1
0
-0.1
-0.2

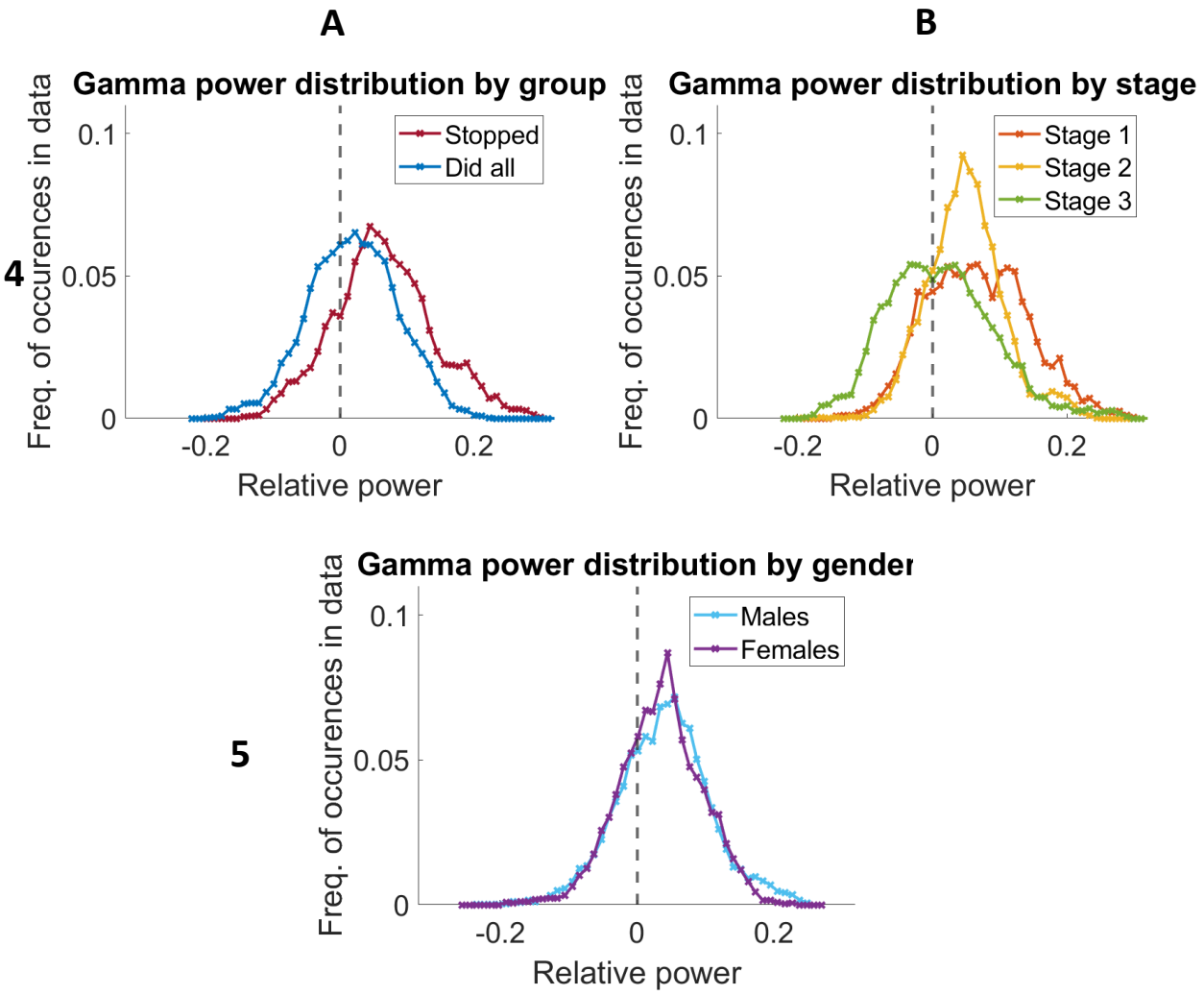


Figure 3.7: **1-3.** TF (time and frequency resolved) spectrogram extracted at the group level for each stage separately, showing the gamma activity in response to the checkerboards relative to the pre-stimulus activity in response to the video. Each stage is shown on a different line (Stages 1, 2, 3) and each group in a different column (A: *Stopped*, B: D*id all*). Dashed lines represent the time-frequency window of interest in which the statistics were computed (100-700ms and 50-100Hz). **4-5.** Distributions of the grand-average power values within the time-frequency window of interest by group (**4.A**) or), stage (**4.B**) or gender (**5**).

The group-level gamma response to the checkerboards (Figure 3.7) appeared to both 1) decrease stage by stage (1 vs. 2 vs. 3 and 4.B), and 2) remain larger for the group who *stopped* throughout experiment (A vs. B and 4.A). To test the statistical significance of this observation, the distribution of power values in the time-frequency window of interest (see Figure 3.7, 3) were

obtained and compared (Grün, 2009; Grün et al., 2002) pairwise using Kolmogorov–Smirnov tests, Bonferroni corrected for 5 multiple comparisons (see Table 3.3 and Table 3.4). There was a significant and small main effect of stage: on average, $D = 0.249$, $p < 0.0001$, Cohen's $d = 0.491$ with mean (standard deviation) power values of 0.067 (0.074) at stage 1, 0.049 (0.055) at stage 2, and 0.013 (0.055) at stage 3. This stage-by-stage decrease in participants response to the distractors, relative to their activity in response to the videos, was in support of our hypothesis that participants increasingly blocked the distractors in favour of the videos as the experiment advanced.

Further, the analysis revealed a significant and medium effect of group ($D = 0.245$, $p < 0.0001$, Cohen's $d = 0.615$), with mean (standard deviation) power values of 0.065 (0.074) for participants who *stopped* and 0.020 (0.065) for participants who *did all*. There was also a significant difference between genders, but its effect-size was very small (Cohen's $d < 0.2$), and it was thus considered anecdotal ($D = 0.056$, $p < 0.0001$, Cohen's $d = 0.083$), with mean (standard deviation) power values of 0.039 (0.066) for the males and 0.034 (0.059) for the females.

Descriptive value	**Stage 1**	**Stage 2**	**Stage 3**
Mean	0.067	0.049	0.013
Standard deviation	0.074	0.055	0.078
	Stage 1 vs. 2	**Stage 2 vs. 3**	**Stage 1 vs. 3**
Cohen's d	0.277	0.519	0.676
D-statistic	0.190	0.280	0.276
BF corrected p-value	<.0001	<.0001	<.0001

Table 3.3: descriptive mean, standard deviation, and statistical values for the group-level gamma activity per stage.

Descriptive value	***Stopped***	***Did all***
Mean	0.065	0.020
Standard deviation	0.074	0.065
Cohen's d	0.615	
D-statistic	0.245	
BF corrected p-value	< 0.001	

Table 3.4: descriptive mean, standard deviation, and statistical values for the group-level gamma activity per group.

This is in support of our prediction that the group who *stopped* overall processed the distractors more i.e., blocked them less, than the group who *did all*. However, both group-level effects could be explained by individual-level responses to the distractors in the group who *did all* being either 1) smaller or 2) less consistent, compared to individuals who *stopped* before the end of the experiment. This was investigated in section 3.3.2 below.

3.3.1.2. Groups TF (time and frequency resolved) gamma activity at video repetitions 1-2

The stage-by-stage analysis of the group level TF data enabled us to compare data from participants despite their different trial numbers and test the hypothesis that participants' processing of the distractors over the course of the experiment differed depending on whether or not they engaged until the end of the experiment. However, this analysis design came with the drawback that participants from each group were not exposed to the same amount of information over time: infants who *did* all the experiment were exposed to more information than those who *stopped*, which could simply give them a better chance to adapt compared to those who stayed for less long. In order to investigate whether the groups of participants processed distractors differently from the onset, we compared group level TF datasets obtained from the first 2 video repetitions of the experiment. The video repetitions 1 and 2 represented the portion of the experiment that was completed by all participants, regardless of their group (see Figure 3.8).

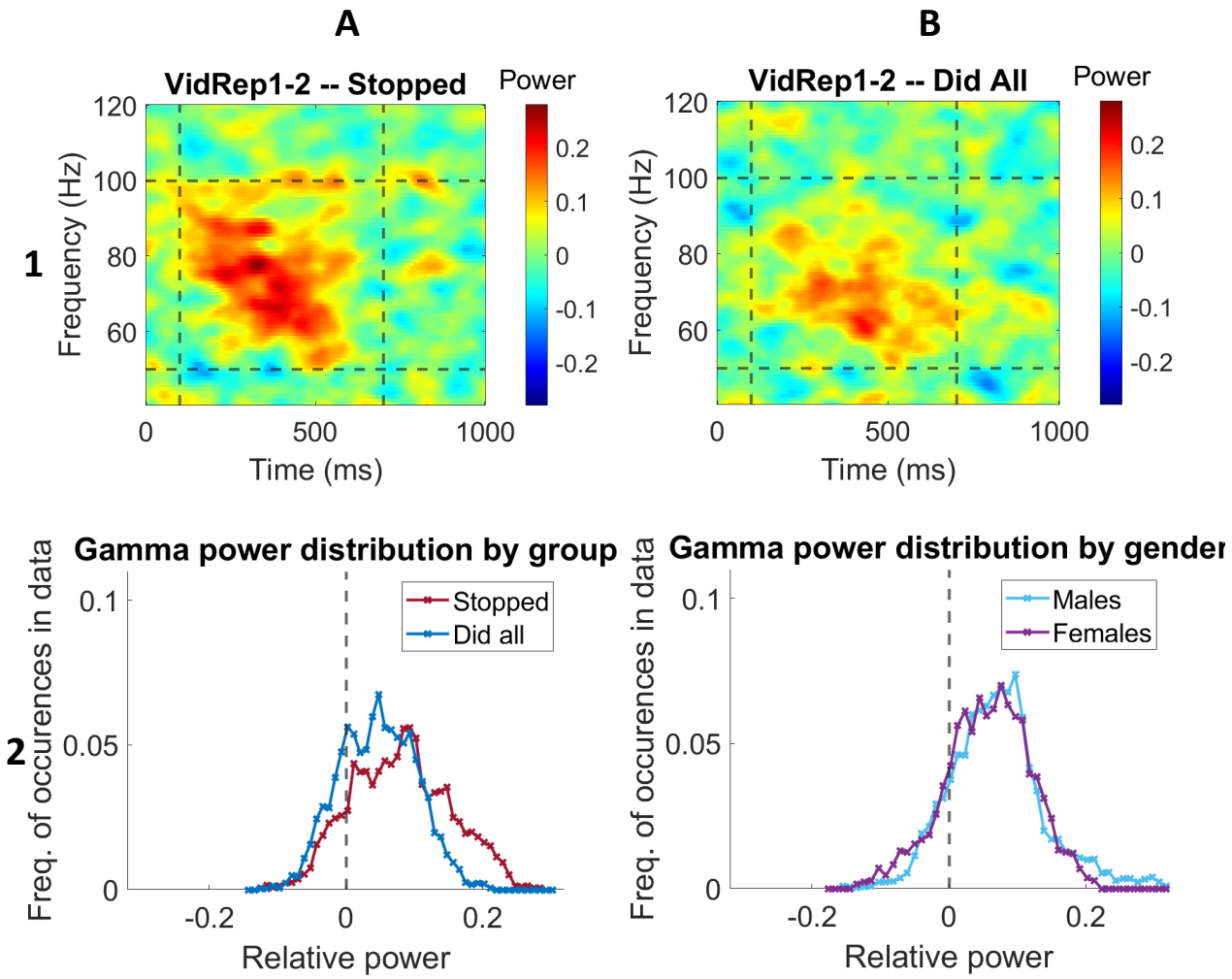


Figure 3.8: **1.** TF (time and frequency resolved) group-level spectrograms for the first 2 video repetitions, in response to the checkerboards and relative to the pre-stimulus video response baseline. Each group is shown in a different column (A: *Stopped*, B: *Did all*). Dashed lines represent the time-frequency window of interest in which the statistics were computed (100-700ms and 50-100Hz). **2.** Distributions of the power values within the time-frequency window of interest by group. (**2.A**) or gender (**2.B**).

Again, during video repetitions 1-2, the group-level gamma response to the checkerboards appeared significantly higher for the group who *stopped* than the group who *did all*, with a medium-sized effect (D = 0.210, p < 0.0001, Cohen's d = 0.570), with mean (standard deviation) power values of 0.080 (0.071) for the group who *stopped* and 0.047 (0.055) for the group who *did all*. There was also a significant difference between genders, of small effect-size (D = 0.064, p <

0.0001, Cohen's d = 0.216), with mean (standard deviation) power values of 0.071 (0.068) for the males and 0.057 (0.062) for the females.

This result confirmed our hypothesis that, compared to the participants who *did all*, participants who *stopped* seemed to show enhanced processing of the distractors relative to the video. This difference was significant from the start of the experiment, for video repetitions during which all participants contributed data, which excluded the possibility that this difference can be explained by differences in time or amount of information exposed to, but rather pointed to intrinsic differences between those two groups of participants. Still, our group-level results could be explained by individual-level responses to the distractors either 1) smaller or 2) less consistent in the group who *did all* as laid out earlier, which will be explored in the next section.

3.3.2. Individuals' gamma activity

To disentangle between different possible individual-level explanations of our group-level results, our next step was to analyse gamma activity extracted at the individual level. Similarly, the use of a log transform before baselining and averaging the data ensured that only features consistent across multiple trials would receive prominence in the spectrograms. Thus, only data *consistent within the individual* was maximally represented in the analysis, as opposed to data *consistent within the group* in the former analysis. This facilitated our analysis considering the low, down-sampled trial numbers (10.75 to 15.06 per participant, on average) because it gave low weight to noisy input vs. consistent input, even when the amplitude of the noise was high.

To facilitate multiple comparisons across groups and stages, the dimensionality (3-D) of the spectrograms was reduced: we took the maximum value over frequency per time point and the mean over time, within the time-frequency window of interest (100-700ms, 50-100Hz), yielding a single value per stage and individual, suitable for multiple comparisons through ANOVAs. Furthermore, the two groups compared were by definition exposed to different amounts of trials, which could in turn influence the analyses. Thus, the trials of the group who *did all* were downsampled to match those who *stopped* on average, and the analyses were performed using individuals' number of included trials (Ntrials) as a covariate for ANOVAs, or as a conditioning variable for partial correlations. This way, we both mitigated and accounted for the potential role of the number of trials variable (Ntrials) when assessing group or stage-by-stage differences.

3.3.2.1. *Individual's' mean gamma response stage by stage*

First, we looked at individuals' gamma response to the distractors, relative to the baseline (which corresponded to the video pre-checker presentation) at each stage of the experiment and compared the groups and stages in a 2*3 (group*stage) repeated measures ANOVA (Figure 3.9), using Ntrials as a covariate.

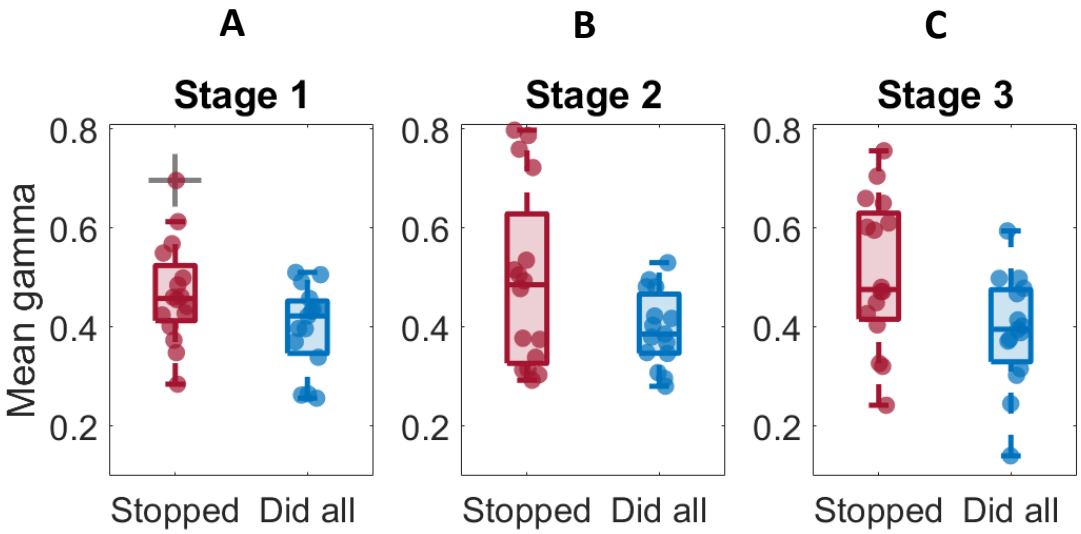


Figure 3.9: baseline corrected mean gamma power in response to the checkerboards extracted at the individual level for each stage separately. Each stage is shown in a different column (Stages 1, 2, 3, respectively, A, B, C) and each group in a different colour (red: *Stopped*, blue: *Did all*). The middle line represents the data's median while boxes represent the 25% and 75% percentiles, and whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range) which are represented as grey crosses when present in the data. Note that outliers were still included in the statistical analysis of the data.

We found a significant ($F(1,28) = 10.266$, $p = 0.003$) and large ($\eta^2 = 0.127$) group effect, with mean (standard deviation) power values of 0.491 (0.122) for the group who *stopped* and 0.396 (0.068) for the group who *did all*. There was no significant effect of gender, with ($F(1,28) = 0.565$, $p = 0.459$, $\eta^2 = 0.009$) or without ($F(1,29) = 1.682$, $p = 0.205$, $\eta^2 = 0.037$) using the Ntrials covariate.

Further to this, we looked at this relationship using the continuous measure of time-on-task (rather than the dichotomous grouping) and correlated the individuals mean gamma response

over the whole experiment with the time that they spent on the task (see Figure 3.10). The partial correlation controlling for covariations in NTrials revealed a highly significant, large and negative relationship (Pearson's R = -0.586, p < 0.001). Taken together, these results brought strong evidence in support of our prediction that the individuals who *stopped* before the end of the experiment overall processed the distractors more, or blocked them less, than the group who *did all*. However, if infants attenuated the strength of their response to the checkerboards as the experiment went on, such as what was described at the group level, this effect could be absent in infants who *stopped* simply because they went through less trials than those who *did all* and thus had less time to attenuate their response.

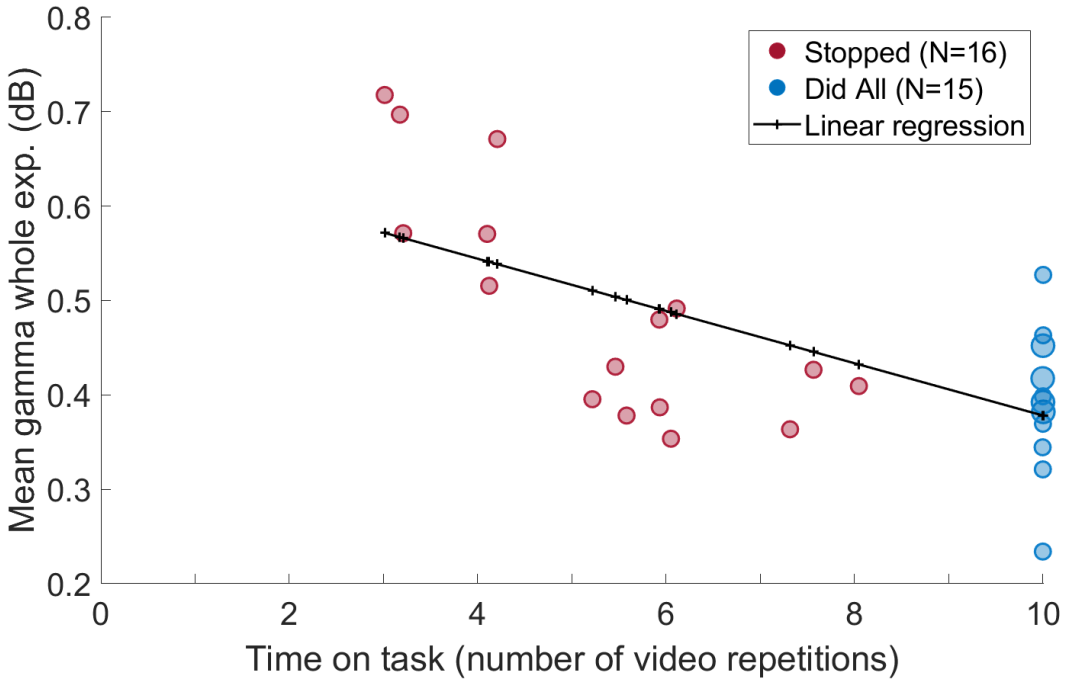


Figure 3.10: individuals mean gamma response over the whole experiment as a function of their time-on-task (number of video repetitions); marker diameter is proportional to the number of individual values represented by the same point. The linear regression between both values given individuals data points is represented as a black line on which crosses are drawn to figure the points' projection on the regression line.

Next, we looked at the stage-by-stage changes. The ANOVA analysis with Ntrials as a covariate revealed a significant effect of stage ($F(2, 56) = 3.621$, $p = 0.033$, $\eta^2 = 0.034$) and no interaction effect of stage*group ($F(2, 56) = 0.607$, $p = 0.549$, $\eta^2 = 0.006$). Means and standard deviation

values are reported in Table 3.5. This result followed our prediction, based on the group data and Snyder and Keil's habituation study (2008). It suggested that individuals' gamma responses became less strong with time.

Group	Descriptive value	Stage 1	Stage 2	Stage 3
Stopped	Mean	0.468	0.494	0.511
	Standard deviation	0.102	0.182	0.149
Did all	Mean	0.399	0.397	0.393
	Standard deviation	0.085	0.076	0.112

Table 3.5: descriptive mean and standard deviation values for the mean gamma activity per stage and group.

To further investigate the stage-by-stage changes, especially given the significant difference in groups' variance at stage 2 evoked earlier, we computed the difference in mean gamma activity between stages 2 & 1, and 3 & 2 for each individual (see Figure 3.11). The data was analysed in a 2*2 (group*stage difference) repeated measures ANOVA, using NTrials as a covariate.

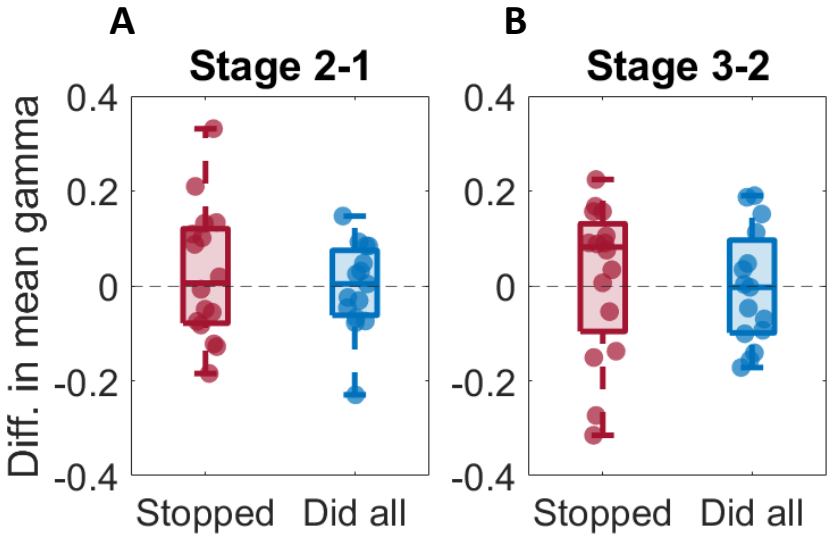


Figure 3.11: stage-by-stage difference in gamma activity extracted at the individual level for each stage separately, showing the mean gamma power in response to the checkerboards, relative to the pre-stimulus response to the video. Each stage difference

is shown in a different column (Stages 2-1 and 3-2, respectively, A and B) and each group in a different colour (red: *Stopped*, blue: *Did all*). The middle line represents the data's median while boxes represent the 25% and 75% percentiles, and whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range, none found for this data).

This analysis revealed no significant effect of group ($F(1, 28) = 1.170$, $p = 0.289$, $\eta^2 = 0.008$) but a significant stage difference effect ($F(1, 28) = 7.322$, $p = 0.011$, $\eta^2 = 0.137$) and no stage difference*group effect ($F(1, 28) = 0.006$, $p = 0.940$, $\eta^2 < 0.001$). The values for each stage and group were also statistically compared to 0 through t-tests and were all found not to significantly differ from zero (Table 3.6). These results altogether converged to confirm that individuals' mean gamma activity did appear to significantly vary between stages. Finally, there was no significant effect of gender ($F(1,28) = 0.054$, $p = 0.817$, $\eta^2 < 0.001$) or stage diff.*gender (($F(1,28) = 0.880$, $p = 0.356$, $\eta^2 = 0.016$) when using Ntrials as covariate, or not ($F(1,29) = 0.036$, $p = 0.851$, $\eta^2 < 0.001$; $F(1,29) = 0.068$, $p = 0.796$, $\eta^2 = 0.002$).

Group	**Stage difference**	**Mean (S.D.)**	**Statistical test (t)**	**Significance (p)**	**Effect size (Cohen's d)**
Stopped	**2-1**	0.026 (0.139)	t(1,15) = 0.750	0.465	0.188
	3-2	0.016 (0.160)	t(1,15) = 0.412	0.686	0.103
Did all	**2-1**	-0.002 (0.093)	t(1,14) = -0.096	0.925	-0.025
	3-2	-0.004 (0.122)	t(1,14) = -0.122	0.905	-0.031

Table 3.6: Statistical values for the comparison with zero of the mean difference in gamma of each group, at each stage difference (uncorrected values).

3.3.2.2. Individuals mean gamma activity at video repetitions 1-2

Similarly to the stage-by-stage analysis of the group-level TF gamma activity, the stage-by-stage analysis of the individuals' mean gamma activity was susceptible to be influenced by variations in

how many stimuli participants had been exposed to. In order to investigate whether individuals processed distractors differently at the neural level, regardless of the number of stimuli seen, we compared individuals' mean gamma activity during the first 2 video repetitions only out of the 10 repetitions that constituted the whole experiment (see Figure 3.12).

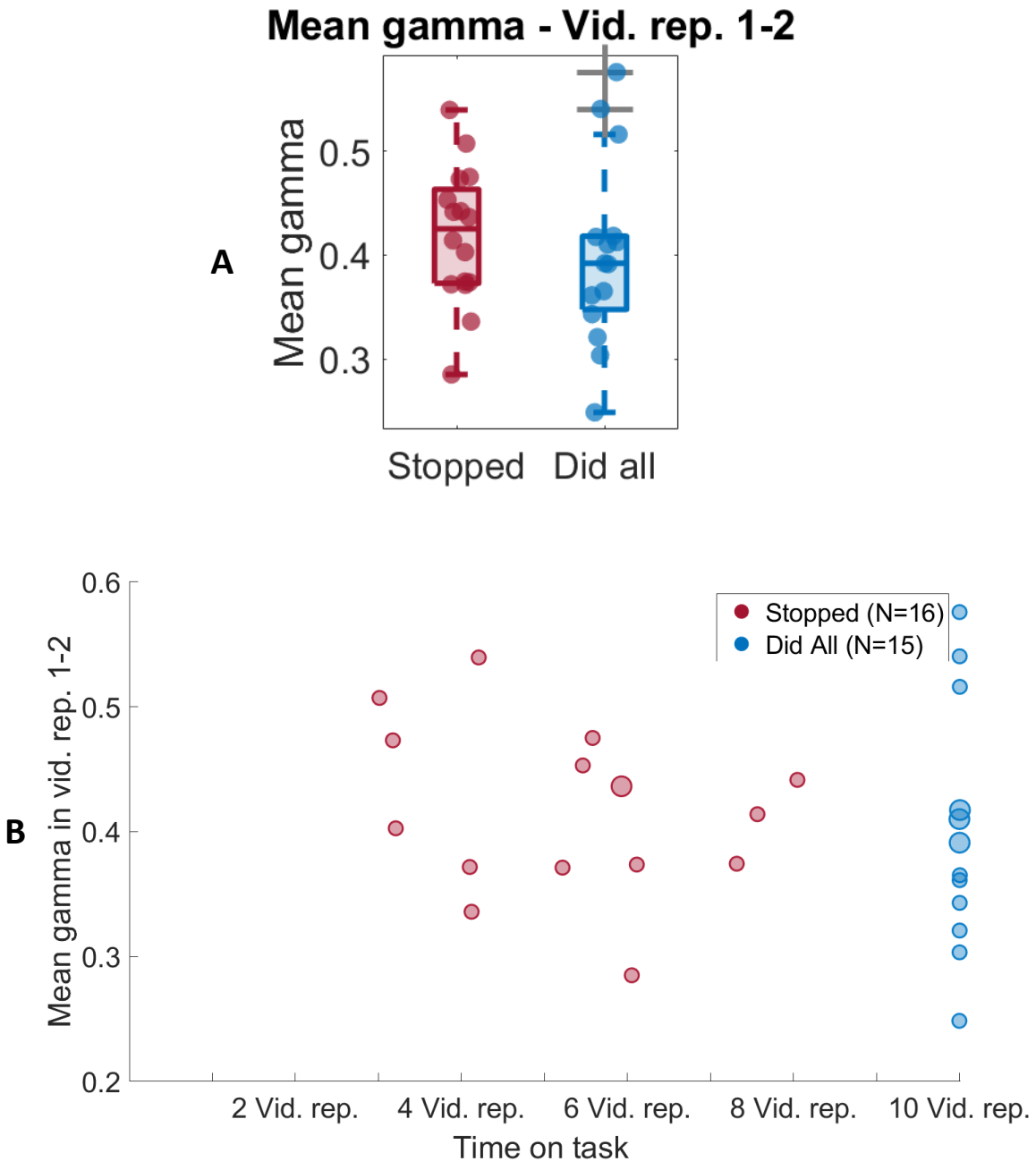


Figure 3.12: **A.** mean gamma activity in response to the checkerboards, relative to the pre-stimulus response to the video and extracted at the individual level for video repetitions 1-2. Each group is depicted in a different colour (red: *Stopped*, blue: *Did all*). The middle

line represents the data's median while boxes represent the 25% and 75% percentiles, and whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range), which are represented as grey crosses when present in the data. Note that outliers were still included in the statistical analysis of the data. **B.** Individuals' mean gamma response over the whole experiment as a function of their time-on-task (number of video repetitions); marker diameter is proportional to the number of individual values represented by the same point.

An ANOVA using Ntrials as a covariate revealed no significant effect of group ($F(1,26) = 0.801$, $p = 0.378$, $\eta^2 = 0.014$) nor gender ($F(1,26) = 0.045$, $p = 0.833$, $\eta^2 < 0.001$). To note, this effect was not evidenced even when removing the three outliers identified in the group who *did all* ($F(1,25) = 2.586$, $p = 0.126$, $\eta^2 = 0.075$). Furthermore, there was no significant partial correlation (controlling for NTrials) between individuals' mean gamma at video repetitions 1-2 and the time-on-task variable (Pearson's $r = -0.187$, $p = 0.322$). This was surprising, especially since we did find a group difference when looking stage-by-stage as well as in the group's time-frequency data presented in the former section. It is possible that we lacked statistical power to evidence these effects here.

To this point, the analyses have shown that how much time participants spent on the task was inversely related to how strongly they responded to the distractors. Indeed, the longer the individuals stayed, the lower their overall gamma response to the distractors relative to the videos was all along the experiment. This was in support of our hypothesis that participants who *did all* the experiment were more efficient in ignoring the distractors and generally responded less to the distractors. We also found evidence for a learning effect of blocking the distractors over the experiment. Group-level as well as individual-level gamma responses to the distractors relative to the video significantly varied across the start, middle, and end stages of the experiment. The next analysis further investigated this looking at a potential driver for the changes in the response throughout the experiment.

3.3.3. Gamma responses in relation to the ERP

Former analyses investigating the *strength* of individuals' response revealed evidence that it contributed to a learning (stage-by-stage) effect. We thus decided to investigate an aspect of the individual gamma responses that could contribute to such changes in the response strength: the

timing of the individuals' fast gamma activity. Rather than looking at it in absolute terms, we investigated the timing of the gamma activity in relation to other activity in the brain. Indeed, how gamma activity is transmitted in neural networks is highly dependent on *when* it happens *in relation to other slower rhythms* (see introduction section 3.1.1; Fries, 2005; Plenz et al., 2021). We investigated how the individual-level gamma (local, fast, low-level sensory information) and ERP (wide, slow, higher-level information) activity were correlated, using Scaled Correlation Analysis (SCA, see Methods section 3.2.5.3 for more details). SCA is a powerful tool to look at how two noisy rhythms relate to each other because it does not expect periodicity in the relationship between the signals, and is robust to shifts in phase across trials (Nikolić et al., 2012). Because the ERP is a potential, increased neuronal activity translates to a negative voltage deflection in the ERP, as opposed to a positive power increase: a negative correlation between ERP and gamma activity would thus reflect aligned rhythms, while a positive correlation would reflect de-aligned rhythms.

3.3.3.1. Scaled correlation between gamma and ERP activity across stages

SCA at lag 0 of gamma and ERP values for groups and stages (Figure 3.13) were compared using a 2*3 (group*stage) repeated measures ANOVA with Ntrials as a covariate.

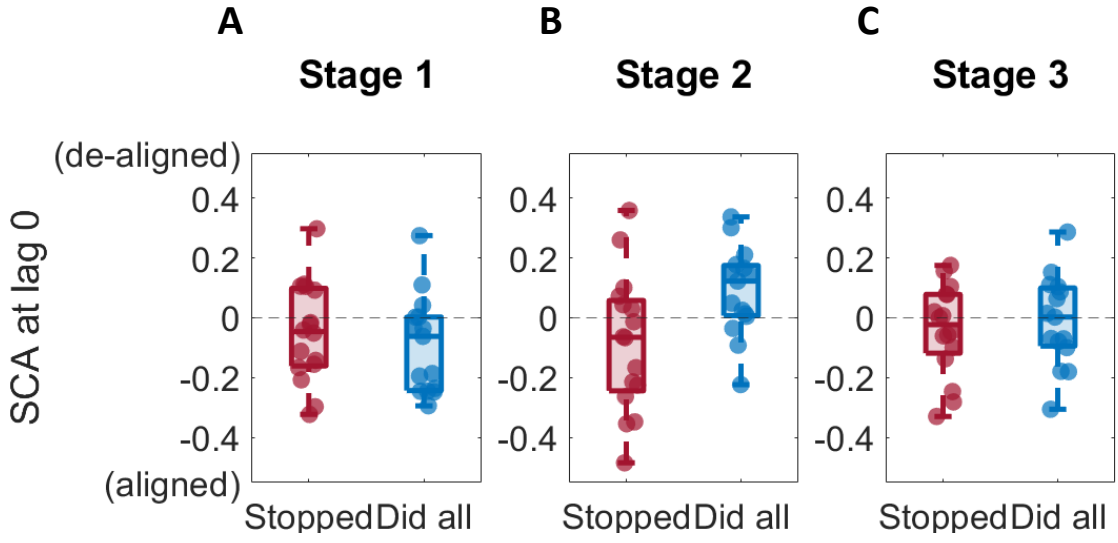


Figure 3.13: Scaled Correlation Analysis (SCA) of the ERP activity in relation to the gamma activity, for each stage (1, 2, 3, respectively A, B, C) and each group (*Stopped*: red, *Did all*: blue). The middle line represents the data's median while boxes represent the 25% and

75% percentiles, and whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range, none found for this data).

The analysis revealed a *trending* ($F(1,28) = 2.950$, $p = 0.097$), medium-sized ($\eta^2 = 0.095$) effect of group with mean (standard deviation) values of -0.004 (0.168) for the group who *stopped* and -0.059 (0.182) for the group who *did all*. The SCA values were not found to be significantly different from 0 for the group who *did all* ($t(1,29) = 0.651$, $p = 0.520$) nor for the group who *stopped* ($t(1,29) = -0.275$, $p = 0.786$). There was no main effect of gender ($F(1,28) = 0.1956$, $p = 0.662$, $\eta^2 = 0.002$) or stage*gender ($F(2,56) = 1.267$, $p = 0.289$, $\eta^2 = 0.028$).

Furthermore, the ANOVA revealed no significant effect of stage ($F(2,56) = 2.093$, $p = 0.133$, $\eta^2 = 0.045$) but a *trending* medium interaction effect of stage*group ($F(2,56) = 2.894$, $p = 0.064$, $\eta^2 = 0.062$), with means and standard deviation values per group and stage described in Table 3.7. The scaled-correlation values per group and stage were also compared to zero with t-tests (uncorrected and Bonferroni-corrected values for three comparisons). No comparison was significantly different from zero after correction (Table 3.7). However, in the group who *did all*, the values at stage 1 and 2 appeared significantly different from zero before correction, with relatively big medium effect sizes (respectively, $t(1,14) = -2.135$, p uncorr. = 0.051, p corr. = 0.153, Cohen's d = -0.551; and $t(1,14) = 2.361$, p uncorr. = 0.033, p corr. = 0.099, Cohen's d = 0.601). It is possible that the Bonferroni correction was too stringent for these effects to survive it in our sample.

Group	**Stage**	**Mean (S.D.)**	**Statistical test (t)**	**Significance (uncorr. p)**	**Significance (BF-corr. p)**	**Effect size (Cohen's d)**
Stopped	**1**	-0.051 (0.166)	$t(1,15) = -1.233$	0.236	0.708	-0.308
	2	-0.083 (0.229)	$t(1,15) = -1.456$	0.166	0.498	-0.364
	3	-0.040 (0.150)	$t(1,15) = -1.060$	0.306	0.918	-0.265

Did all	**1**	-0.090 (0.163)	t(1,14) = -2.135	**0.051**	0.153	-0.551
	2	-0.090 (0.148)	t(1,14) = 2.361	**0.033**	0.099	0.601
	3	-0.011 (0.152)	t(1,14) =-0.274	0.788	1.000	-0.071

Table 3.7: Statistical values for the comparison with zero of the gamma and ERP SCA value of each group, at each stage difference (uncorrected values and Bonferroni-corrected values for three multiple comparisons). Results below the significance threshold ($p < 0.005$) are highlighted in grey cells and results below the trending threshold ($p < 0.005$) in bold font .

These results provided weak evidence that the group who *did all* might be modulating the alignment of their gamma activity (local, fast, low-level sensory information) in relation to their ERP activity (large, slow, higher-level information) in response to the distractors, especially in the first 2 stages. Because we were interested in how infants *modulated* their alignment (SCA) over time rather than their overall score over the whole experiment, group averages were a poor representation of what could be happening at the individual level. This brought the need for an analysis of individual *change* (the stage-to-stage difference) in ERP-gamma SCA.

3.3.3.2. Stage-by-stage changes in scaled correlation between gamma and ERP activity

To further investigate modulations in gamma-ERP alignment over time within each individual, we computed the stage-by-stage difference in SCA values for each individual (see Figure 3.14) and compared the groups in a 2*2 (stage diff.*group) repeated measures ANOVA with Ntrials as a covariate.

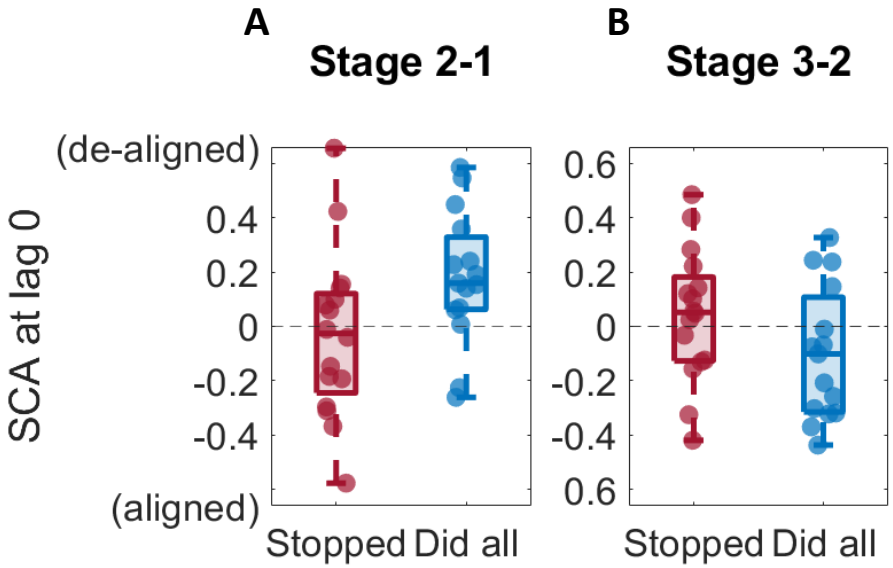


Figure 3.14: Stage-by-stage difference in the Scaled Correlation Analysis (SCA) of the ERP activity in relation to the gamma activity, for each stage (2-1 and 3-2, respectively A and B), and each group (*Stopped*: red, *Did all*: blue). The middle line represents the data's median while boxes represent the 25% and 75% percentiles, and whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range, none found for this data).

The analysis revealed no effect of group ($F(1,28) = 0.724$, $p = 0.402$, $\eta^2 = 0.021$) or stage diff. ($F(1,28) = 0.021$, $p = 0.885$, $\eta^2 < 0.001$) but a significant, medium-sized interaction effect of stage diff.*group ($F(1,28) = 4.398$, $p = 0.045$, $\eta^2 = 0.110$). Comparing the stage-by-stage changes in SCA to 0, we found no significant change in the group who *stopped* at either stage difference (see Table 3.8), and a significant change in SCA between stages 1-2 ($t(1,14) = 2.870$, p BF-corr. = 0.012) but not between stages 2-3 (see Table 3.8) in the group who *stopped*. This was confirmed our former suggestion that only the group who *did all* appeared to be modulating the *alignment* of their gamma response to the checkerboard over the first two stages: they moved from a more aligned state in stage 1 to a more de-aligned state in stage 2 and finally an in-between state in stage 3.

Group	Stage difference	Mean (S.D.)	Statistical test (t)	Significance (uncorr. p)	Significance (BF-corr. p)	Effect size (Cohen's d)
Stopped	**2-1**	-0.032 (0.306)	t(1,15) = -0.419	0.681	1.000	-0.105
	3-2	0.043 (0.243)	t(1,15) = 0.717	0.485	0.970	0.179
Did all	**2-1**	**0.180 (0.243)**	**t(1,14) = 2.870**	**0.012**	**0.024**	**0.741**
	3-2	-0.101 (0.246)	t(1,14) = -1.592	0.134	0.268	-0.411

Table 3.8: Statistical values for the comparison with zero of the gamma and ERP SCA per stage difference (with and without Bonferroni-correction for two multiple comparisons). Significant ($p < 0.005$) results are highlighted in grey cells and bold font.

Finally, a control ANOVA using NTrials as a covariate revealed no effect of gender ($F(1,28)$ 2.445, $p = 0.129$, $\eta^2 = 0.013$) or stage*gender (($F(1,28) = 0.588$, $p = 0.449$, $\eta^2 = 0.016$).

3.3.3.3. Scaled correlation between gamma & ERP activity at video repetitions 1-2

Finally, we also compared the SCA values of both groups (see Figure 3.15) during the first two video repetitions of the experiment in an Analysis of Covariance (ANOVA) using Ntrials as a covariate and found no significant effect of group ($F(1,26) < 0.001$, $p = 0.986$, $\eta^2 < 0.001$). The values were also compared to zero and did not significantly differ from zero for either group (*Stopped*: $t(1,28) = -0.657$, $p = 0.516$; *Did All*: $t(1,28) = -0.611$, $p = 0.546$). Furthermore, we found no significant correlation (computed partially, conditioned on Ntrials) between the SCA value at video repetitions 1-2 and the time-on-task variable (Pearson's $R = 0.070$, $p = 0.710$). There was also no effect of gender as tested in an ANOVA using Ntrials as a covariate ($F(1,26) = 0.016$, $p = 0.901$, $\eta^2 < 0.001$).

These results suggested that both groups started at a similar point in terms of gamma-ERP SCA, when they hadn't yet modulated the alignment of their fast and slow responses. This result supported the idea that the SCA was related to a *learning effect*: it was not present at the start of the experiment when infants hadn't yet had a chance to learn about the stimuli, and was only exhibited later on, in the group who *did all* the experiment.

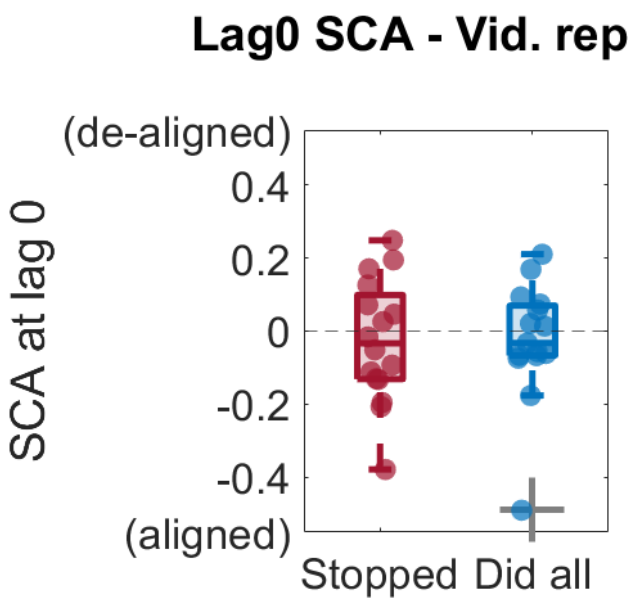


Figure 3.15: Scaled Correlation Analysis (SCA) of the ERP activity in relation to the gamma activity, at video repetitions 1-2 for each group (*Stopped*: red, *Did all*: blue). The middle line represents the data's median while boxes represent the 25% and 75% percentiles, and whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range), which are represented as grey crosses when present in the data. Note that outliers were still included in the statistical analysis of the data, but contributed to a non-normal distribution of values in the group who *did all* (W(29) = 0.868, p = 0.31); analyses were thus reported using a Mann-Whitney test.

Group	**Mean (S.D.)**	**Statistical test (t)**	**Significance (p)**	**Effect size (Cohen's d)**
Stopped	-0.027 (0.168)	t(1,15) = -0.641	0.531	-0.160
Did all	-0.026 (0.164)	t(1,14) = -0.608	0.553	-0.157

Table 3.9: Statistical values for the comparison with zero of the gamma and ERP SCA value of each group, in the first two video repetitions.

3.4. Discussion

3.4.1. Summary of findings

3.4.1.1. Relationship between individuals' engagement and strength of gamma activity to distractors

We showed that individual differences in infants' gamma activity in response to distractors relative to an ongoing video reflected individual differences in infants' behaviour in terms of how long they engaged with the experiment overall. Infants who *did all* the experiment overall responded less to the distractors relative to the video, as shown by a lower gamma response to the distractors compared those who *stopped* before the end. This was seen both at the group and the individual level, and for every stage of the experiment. This result is in support of our prediction that infants who maintain attention for longer also block the distractors more i.e., exhibit a less *strong* brain response.

3.4.1.2. Relationship between stimulus repetitions and strength in gamma activity to distractors

We also found that infants seemed to modulate their response to the distracting information relative to the ongoing catchy information throughout the experiment, as shown by a stage-by-stage decrease of the group-level and of the individual-level gamma response to the distractors relative to the video. This was in support of our hypothesis that infants would learn to block out the perception of the distractors to focus on the catchy videos, reflected by a decreased response to the distractors over time.

3.4.1.3. Relationship between stimulus repetitions and individuals' relative alignment of gamma activity to distractors in infants who engaged for longer

Finally, looking further into what the variability of the stage-by-stage groups' response reflected for each individual, we showed that infants who *did all* the experiment seemed to modulate the *alignment* of their gamma response to the distractors relative to slow rhythms. This is an explanation for the fact that their responses averaged out at the group level. However, there was no evidence that the group who *stopped* did. Indeed, gamma-ERP alignment for the individuals

who *stopped* did not significantly differ from 0 for any of the 3 stages. On the contrary, in the group who *did all*, we evidenced a shift from more gamma-ERP alignment (negative SCA values) to more gamma-ERP de-alignment (positive SCA values), followed by a drop to zero values in the last stage, possibly due to a fatigue effect. This result suggests that participants who *did all* the experiment were able to block the distractors using modulations of their local, fast gamma activity in relation to their wide-network slow ERP responses. They started in a state where information about the distractors can be optimally transmitted throughout the network and moved on to a state in which the transmission is impeded. Individuals who *stopped* before the end, however, appeared to maintain a middle ground, neither aligned nor de-aligned state, suggesting that although their responses became more individualised with time (as shown by the stage-by-stage decrease in the group response), they did not make use of the slower rhythms timings to modulate the transmission of their fast gamma activity. More research is needed to understand what the decrease in cross-subjects consistency in the gamma response means in the group who *stopped*. Future work looking at the trial-by-trial changes in response alignment could shed light on the process, possibly unveiling changes that can explain the variability in the group who *stopped* too.

3.4.2. Validating the paradigm's use for the study of distractibility in infants

Our results validate the usefulness of Piccardi et al.'s (2020) paradigm with distractors for the study of 10-month-old infants' individual differences in intrinsic engagement and sensory information processing and selection, as well as their neural correlates in the gamma range. Our results show that letting infants disengage spontaneously from our task yielded relevant information on individual differences in their processing of information, and more specifically of distractors vs. catchy stimuli. Interestingly, the time-on-task variable appeared to carry useful information about individual differences both when used as a binary measure to split the infants into groups who *stopped* vs. *did all* the experiment, and when it was used as a continuous measure in correlation analyses. This makes the paradigm versatile and adequate for various experimental designs, which is particularly helpful for the study of individual differences in a notoriously noisy population as infants, a field which remains relatively untouched and where few options are available. It is worth noting though that we use a repetitive paradigm in which the same "catchy" (video) stimulus repeats over for 10 times. This is helpful to limit effects of surprise

and to keep the interest levels to an intermediate level at which individual differences in engagement can be observed. However, it also reduces the catchiness of the stimulus with time.

3.4.3. The difficulty of studying noise-sensitive gamma rhythms in infants

3.4.3.1. Study's ways to address issues with studying high-frequency oscillations in infants and limitations

High frequency rhythms are particularly subject to noise, which might explain why, in a population such as infants for which noise handling is already notoriously difficult, research on the topic has been limited. In this study, we had to work with low trial numbers and important sources of individual variations unrelated to the question at hand, potentially leading to high noise. These were both inter-individual differences such as head-shape, sulcation, or net placement variations (due to babies moving when placing it), but also intra-individual variations such as babies' movements throughout the experiment, touches to net, or possible parental interference (even if discouraged in the guidelines). Yet, it is worth noting that the skull of infants is not yet fully formed and rigid at 10 months: their softer and thinner skulls let more of the brain's electrical activity travel to the electrodes, thus the recorded EEG signals are higher compared to adults. Additionally, we used highly contrasted and quasi-regular stimuli known to elicit a large visual response (black-&-white checkerboards, see Leguire & Rogers, 1985), which also helped us to increase the signal-to-noise-ratio. Finally, we used an active window as a baseline that consisted of segments when the babies were watching the catchy videos, and not resting, as is commonly used in EEG research. Although this did not help with signal-to-noise-ratio issues, this choice of baseline was helpful to make inferences on how participants traded off between processing the plain distractors (stimulus) vs. the catchy videos (baseline). The power activity that we analysed reflected their response to the checkerboards *relative to the video* and thus informed us on how much the checkerboards were *distracting infants away from the video*. Ultimately, despite the signal-to-noise-ratio problems brought by this solution, it was particularly informative compared to studying each response separately and helped our investigation.

3.4.3.2. Study's way to address issues with studying fast and slow couplings in infants and limitations

In our analyses, we used a proxy for a wide-network slow oscillation: the Event Related Potential (ERP). While the ERP was a viable and more robust candidate which enabled to answer these questions for the first time, it would be helpful in the future to move on to the study of infants' gamma responses in relation to more standardly defined oscillations, such as alpha- or theta-range oscillations. Indeed, the mechanisms linking the generation of those slow oscillations by networks of neurons and their effects on gamma-synchronised neurons are better described and provide a stronger theoretical framework for these processes (Cohen, 2014 p.54). Moreover, both alpha (see Foxe & Snyder, 2011 for a review) and theta (e.g. Landau et al., 2015) slow rhythms have been specifically linked to the implementation of attentional mechanisms.

3.4.4. Implications for theories of attention and sensory processing in infants

Individual differences in brain activity, namely fast oscillations' *strength*, and their *alignment* with a slow ERP rhythm, were linked to individual differences in behaviour, namely infants' engagement as measured by the time-on-task variable. These results add to the scarce literature on gamma oscillations and their cognitive correlates in infants. It is particularly important to bridge the gap between the infant and the adult literature on how fast oscillations support information processing, as more work in adults but also in other species increasingly point to the central role of these fast rhythms for information processing in the brain (Buzsáki & Wang, 2012; Fries, 2009; Uhlhaas et al., 2010). Moreover, although couplings between fast and slow oscillations are central to theories of information processing in the brain, to our knowledge, they were never yet evidenced to exist in the infant brain. For the first time, we showed that 10-month-old infants can adapt how information is processed in their brain through a complex interplay between fast and slow rhythms, namely between gamma activity and the ERP. This result has powerful implications for our understanding of how the brain supports information selection both in general and in the special case of infants. We reveal the existence of a complex machinery for adapting oscillations alignment early on in life, which appears to sustain infants' selective processing and selection of visual stimuli and relates to infants' engagement. Future research is needed to investigate whether this is available earlier in development and when this ability arises.

Parents' reports of infants' sensory processing and temperament (including attention scores) that were collected for this study through questionnaires, will be useful to take this further. Looking at parents report i.e., behavioural characteristics observable in infants' daily life, will help to validate the utility of this paradigm for capturing attentional and sensory aspects of infants' distractibility beyond the lab walls. This will be helpful to link brain, behaviour in the lab, and everyday behaviour and will be further explored in the next chapter of this thesis.

3.5. Conclusion

All in all, we showed that the brain of 10-month-old infants adapts its processing of information through a complex interplay between fast and slow rhythms. This has never been shown before in infants, although it is a central ability in theories of information processing in the adult brain. This finding thus has important implications for models of information processing in the brain, helping to define the core mechanisms involved by evidencing their availability in the developing brain. We also showed that individual differences in brain activity (fast gamma oscillations strength and alignment with a slow ERP rhythm) were linked to individual differences in behaviour (time-on-task). The next chapter will continue investigating these individual variations in relation to infants' trait sensory processing and temperament, reported by parents in two questionnaires.

Chapter 4: Linking neural markers of 10-month-old infants' distractibility with their trait sensory and attentional abilities

4.1. Introduction

In the former chapter, we showed that modulations in infants' fast gamma responses are important for their processing of visual sensory information. Notably, we showed that individual differences in infants' brain response to distractors were linked to individual differences in their overall distractibility, reflected behaviourally by the time that they spent on the task. Importantly, group differences in the strength of the response were present at any stage of the task, pointing to individual differences not being a result of differential exposure to stimulation but to a general, possibly trait difference in how infants respond. Moreover, chapter 3 highlighted the importance of the brain's gamma rhythms for sensory processing and information selection through attentional mechanisms (Buzsáki & Wang, 2012; Fries, 2009). In the current chapter, we take these results further and ask whether the neural and behavioural effects observed during the lab visit linked to individual differences in infants' daily-life behaviour, outside of the laboratory. More specifically, we investigate the role of infants' trait sensory and attentional processes for the processing of distracting information in our paradigm. On the one hand, the SCA analysis suggested an active process of information selection, through which infants inhibit the distractors, in a way that related to how much they engaged with the task. On the other hand, the general strength of the gamma response suggested different sensibilities to the stimuli, linked to a more or less active processing. We expected these to map on to different questionnaires traits, respectively attentional and inhibitory scores from the Early Childhood Behavior

Questionnaire (ECBQ) Effortful Control factor, and Low Threshold sensory processing scores from the Infant/Toddler Sensory Profile (ITSP).

4.1.1. Measuring trait sensory processing in infants

One of the most widely used methods for measuring individual differences in sensory processing in infants is the Infant/Toddler Sensory Profile (ITSP) questionnaire (Dunn, 1997; Dunn & Daniels, 2002). This questionnaire was developed around a theory of sensory processing in infanthood and childhood proposed by Dunn (1997), which aims to bridge concepts of neuroscience and behavioural psychology. Dunn's model poses that infants' processing of information can be explained along two continua (Figure 4.1): 1) a *neurological threshold* continuum that reflects individual differences in the amount of sensory stimulation needed to spark a response based on the individual's neural characteristics, and at which extremes infants are either sensitised (Low threshold: tend to respond a lot) or habituated (High threshold: tend not to respond much); and 2) a *behavioural* continuum that reflects individual differences in infants' self-regulation strategies in response to stimuli, at which extremes infants either respond passively (in accordance with the neural threshold) or actively (to counteract the effects of their neural threshold). Four possible types of responses to sensory stimuli are proposed according to the model, at the four quadrants of the combination of these two continua: Poor Registration (high threshold, passive behaviour), Sensation Seeking (high threshold, active behaviour), Sensory Sensitivity (low threshold, passive behaviour) and Sensation Avoiding (low threshold, passive behaviour).

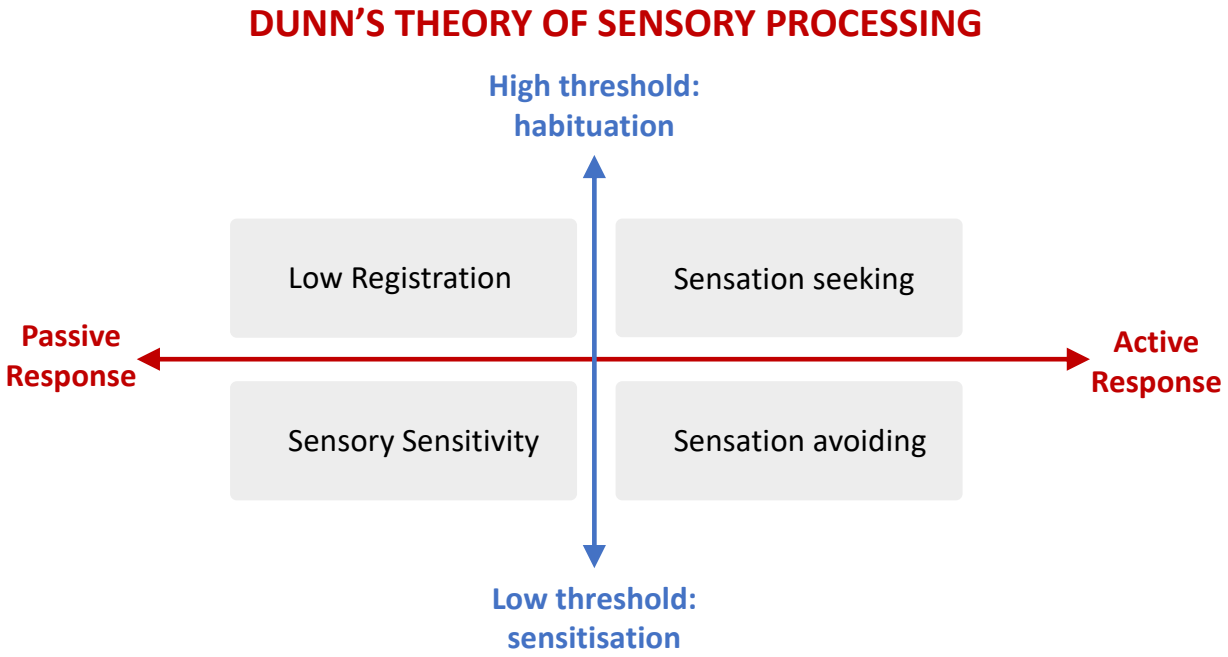


Figure 4.1: Dunn's theory of sensory processing, based on Dunn (1997); four types of sensory responses are shown in four quadrants, according to two continua: a neural threshold continuum in blue (vertical) and a behavioural response continuum in red (horizontal).

This model was tested in adults (Brown et al., 2001) and children aged 3- to 9-year-old (McIntosh et al., 1999) with a markedly stronger score in one of the model's four quadrants. The researchers measured skin conductance in response to stimuli across five sensory modalities (olfactory, auditory, visual, tactile, and vestibular). They assessed both the amplitude of the responses and the number of trials until habituation (absence of skin response), as proxies for the individuals' neural and behavioural thresholds, respectively. The studies showed that individuals with sensory processing patterns markedly stronger for the Low Threshold half of the model (Sensory Sensitivity and Sensation Avoiding profiles) had skin responses of higher amplitude than those in the High Threshold half (Low Registration and Sensation Seeking profiles). Additionally, they found that individuals in the Passive end of the behavioural continuum responded in accordance with their hypothesised threshold: individuals with Low Registration profiles habituated more quickly than those with Sensory Sensitivity profiles. On the contrary, individuals at the Active (counteracting) end of the continuum responded in contradiction with their hypothesised threshold: individuals with Sensation Seeking profiles habituated less quickly than those with

Sensation Avoiding profiles. All in all, these results support the predictions of Dunn's model and suggest that Dunn's distinction between the four-quadrant is relevant in individuals down to 3 years of age.

Importantly, while Dunn's model is designed to help practitioners and researchers to identify inter-individual differences such as those investigated in Brown (2001) and McIntosh' (1999) studies, Dunn emphasises that intra-individual variability is also expected, for example depending on context e.g. tiredness at different times of the day, or variability between sensory systems e.g. visual vs. auditory stimulation.

Dunn's model provides a very useful framework through which to study individual differences in sensory processing, that are at the core of this study. Dunn's attempt to link neuroscience and behavioural concepts is highly relevant to the questions at hand in our study, however it still lacks reference to an important and influential concept both for behavioural and neural theories of cognition: the concept of attention as a cognitive skill. In chapter 3, we reviewed evidence suggesting that modulations in gamma oscillations, especially in relation to slower rhythms, are particularly well suited to underlie attentional mechanisms in the brain for an adaptive processing of sensory information (Buzsáki & Wang, 2012; Fries, 2009).

4.1.2. Measuring attention and temperament traits in infants

Another widely used tool to study individual differences in development is the Early Childhood Behavior Questionnaire (ECBQ), which measures individual differences in children's temperament (Putnam et al., 2006), including attention. Temperament refers to an individual's personality traits that guide their reactions to stimuli. It involves individual differences in emotional, motor and sensory reactivity, as well as processes of self-regulation of one's reactivity (Rothbart & Derryberry, 1981). In children, temperament has been shown to be best described by three main factors (Ahadi et al., 1993; Rothbart et al., 2001): Negative Affectivity (fear, anger/frustration, discomfort, sadness), Surgency/Extraversion (activity level, sociability, impulsivity and enjoyment of high-intensity pleasure) and Effortful Control (inhibitory control, attention, perceptual sensitivity and low-intensity pleasure). While Negative Affectivity and Surgency/Extraversion are considered as reactivity factors, involving mostly automatic responses, Effortful Control is thought to support processes of self-regulation, including attentional processes. In fact, Effortful Control is the last factor to arise during the second half of the first

year of life into toddlerhood and is thought to rely on the later maturation of the brain's executive attention network (Rothbart et al., 1994, 2007).

The executive attention network is composed of the anterior cingulate gyrus and lateral prefrontal areas of the brain (Fan et al., 2003, 2005). It has been shown to activate during the detection and resolution of conflicts (Botvinick et al., 2001), inhibition and orienting (Fuentes, 2004) or conscious perception (Posner & Raichle, 1994). The executive attention network seems to be involved in processes of information selection that span further than classical attentional processes, and involve intertwined processes of executive control and self-regulation, more generally (Posner & Rothbart, 1998; Rueda et al., 2005). Moreover, this network has been shown to be functionally connected to sensory areas of the brain. In a study combining functional Magnetic Resonance Imaging (fMRI) and Event Related Potentials (ERPs), Crottaz-Herbette & Menon (2006) showed that there was a strong functional connection between the cingulate cortex and sensory cortices. This connection was specific to the task at hand: when attending to visual stimuli, the connection was found with primary visual areas, whereas in auditory trials, the connection was found in the auditory cortex. They suggest that a dynamic communication between the cingulate and the primary sensory cortices is at the core of attentional control processes.

All in all, self-regulation processes measured by the Effortful Control factor and relying on the executive attention network are thought to play a central role in attentional processes. They enable infants to pick their sources of information, control the information to take in and form optimal strategies to regulate their reactions (Rothbart et al., 2007). Behavioural and neural theories of the executive attention network provide a strong ground for us to study the link between the brain responses recorded in the visual cortex during our experiment, and trait measures of Effortful Control.

4.1.3. Current study and hypotheses

In the current analysis, we looked at questionnaire-based trait measures of infants' sensory processing (ITSP at 10 and 17 months) and effortful control (ECBQ at 17 months), in relationship to experimental measures for their neural (gamma strength and alignment) and behavioural (time-on-task) responses to our stimuli. Experimental measures were described in the former chapter. This chapter expands on it with trait measures.

We expected infants who *did all* the experiment to respond more moderately to sensory stimulation as compared to infants who *stopped*, and thus to score higher in all or some of the four ITSP quadrants. Moreover, we expected them to also score higher in the visual processing scale specifically, since the stimuli were presented in the visual modality, while general processing scores were expected to be similar in both groups as a control. We did not hypothesise to see an evolution of these traits between test and follow-up but rather expected stable effects between the two points.

We also expected infants who *did all* to score higher in the Effortful Control Factor compared to those who *stopped*, and in particular in the attention-related sub-scales of this factor (Attentional Focusing, Attentional Shifting) as well as the other executive function scale (Inhibitory Control) but not the other 2 scales of this factor (Low Intensity Pleasure and Cuddliness).

4.2. Methods

Detailed descriptions of the study's methods were provided in chapter 3. Here, only additional details specific to the present analyses are provided.

4.2.1. Participants

Participants described in the previous chapter were included in the present analyses: 114 infants were tested and 31 infants were included in the analyses (see Table 4.1 here for descriptive variables, and Methods section 3.2.1 in chapter 3 for more details). Participants were split into two groups: those who *did all* the experiment (N = 15) and those who *stopped* before the end (N = 16). This chapter also includes analyses of data collected online 7 months posterior to the lab visit. 9 families did not wish to contribute data at this point. 22 infants were included in the follow-up analysis (see Table 4.1 for descriptive variables) and were similarly divided into those who formerly *did all* (N = 10) and those who *stopped* (N = 12).

As in the former chapter, a sensitivity power analysis was conducted with the software G*Power (version 3.1, https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower). It was conducted for exact correlations, using a two-tailed bivariate normal model with a significance threshold $\alpha = 0.05$ and power $1-\beta = .80$. This analysis revealed a sensitivity for correlation values above a critical R = +/- 0.36 for N = 31 subjects (data collected at test) and R = +/- 0.56 for N = 22 subjects (data collected at follow up). In practice, this meant that the analysis of the data at test was partly sensitive to small effects (R > 0.3) and fully sensitive to medium (R > 0.5) and large (R > 0.7) effects for the test-date data, while the follow-up data analysis was not sensitive to small effects, partly sensitive to medium effects, and fully sensitive to large ones.

Group	N total at test (N females)	Mean age at test (S.D.)	Min. - max. age at test	N total at follow-up (N females)	Mean age at follow-up (S.D.)	Min. - max. age at follow-up
Whole sample	31 (17)	10m, 4.0d (9.7d)	9m, 11d - 10m, 19d	22 (12)	16m, 27.4d (24.2d)	15m, 19d - 18m, 21d
Did All	15 (11)	10m, 4.23d (8.20d)	9m, 1 d - 10m, 19d	10 (8)	17m, 5.72d (31.99d)	15m, 19d - 18m, 21d
Stopped	16 (6)	10m, 3.90d (11.39d)	9m, 19d - 10m, 16d	12 (4)	16m, 20.17d (12.31d)	16m 1d - 17m, 11d

Table 4.1: Age (mean and standard deviation, S.D.) and gender variables for the whole sample and for each group of participants.

As described in chapter 3 (Methods section 3.2.1), there was no significant age difference between the two groups at test (t(1,29) = -0.095, p =0.925, Cohen's d = -0.034). Similarly, at follow-up, although there was significantly more variance in age for the group who *did all* compared to the group who *stopped* (Levene's test of unequal variances: F(1,20) = 6.293, p = 0.021), there was no significant age difference after correcting for the variance gap (t Welch(1,11.502) = -1.433, p =0.179, Cohen's d = -0.633). Additionally, we compared the time passed between the two data collection dates in both groups, for which again more variance was present in the group who *did all* (Levene's test: F(1,20) = 9.016, p = 0.007) while no group difference was found (t Welch(1,11.296) = -1.612, p =0.135, Cohen's d = -0.713). All in all, this suggests that the two groups were comparable in age at any point of the data collection. Thus, we did not further investigate the effect of age in this study.

Additionally, the groups' gender distribution were significantly different at test ($\chi^2(1,31) = 4.014$, p = 0.045) as reported in Chapter 3 (10 males and 6 females *stopped*, 4 males and 11 females *did all*), as well as at follow-up ($\chi^2(1,22) = 4.791$, p = 0.029; 8 males and 4 females *stopped*, 2 males and 8 females *did all*). Thus, we examined the effect of gender in the upcoming analyses.

Finally, the follow-up group was tested for the effects seen in chapter 3, to ensure that the conclusions drawn in the last chapter were still applicable to the reduced follow-up group. Results to these tests can be seen in Appendix 1.

4.2.2. Procedure

The infants and their families went through the testing procedure based on the work of Piccardi et al. (2020) which was described in the former chapter (section 3.2.3). After the testing was completed, parents were asked to fill in the Infant/Toddler Sensory Profile (ITSP) questionnaire for 7-36 months at the lab. Families were also contacted by e-mail 6 months after their visit to the Babylab for an online follow-up. If they expressed their willingness to take part after up to two email contacts, they were sent secure links to complete the ITSP again as well as the Early Child Behavior Questionnaire (ECBQ). At either data collection point, the questionnaire could be completed by any of the parents. 2 fathers filled in the questionnaires at test (1 infant *did all*, 1 infant *stopped*), 29 mothers at test and 22 mothers at follow-up, yielding no group differences in the gender of the parent who filled in at either data collection points.

4.2.3. Questionnaires

4.2.3.1. Infant/Toddler Sensory Profile (ITSP)

The Infant/Toddler Sensory Profile (ITSP; see Dunn, 1997; Dunn & Daniels, 2002) is a standardised parent-report questionnaire for measuring infants' sensory processing abilities in their daily life. It exists in a shorter 36-item version for infants from birth to 6 months, and in a 48-item version for infants from 7 to 36 months, which was the version administered at both data collection points for this study. Parents respond to questions about how often they observe their infant produce a response to sensory experiences according to a 5-point scale: "Almost always"/ "Frequently"/"Occasionally"/"Seldom"/ "Almost never" (see questionnaire and scoring method in Appendix 2). Each question is assigned to one of four "quadrants" i.e., a class of responsiveness ("Low registration": lack or low awareness of sensations, "Sensation Seeking": enjoyment and interest in increasing sensations, "Sensory Sensitivity": distress and distraction from sensations and "Sensation Avoiding": controlling or limiting the amount and type of sensations), and one of six "sensory systems" ("General Processing", "Auditory Processing", "Visual Processing", "Tactile Processing", "Vestibular Processing" and "Oral Sensory Processing"). Consequently, the

questionnaire yields six sensory systems scores and four quadrant scores. There is also one "combined quadrant" score ("Low Threshold"), which is the combination of "Sensory Sensitivity" and "Sensation Avoiding" scores and which has been shown to be "relevant to some aspects of poor sensory processing, particularly for children who are fussy or who require a great deal of structure" (*ITSP Technical Report*, 2005). Lower scores indicate a higher frequency in the response measured by the scale, e.g. a lower Sensory Avoiding score reflects more avoidance. Questions included e.g. "My child's behaviour deteriorates when the schedule changes" (General Processing, Sensory Sensitivity) or "My child enjoys looking at shiny objects" (Visual Processing, Sensation Seeking). The full questionnaire and its scoring method can be consulted in Appendix 2. We were interested in any of the quadrant scores, particularly those contributing to the Low Threshold scores, as well as Visual Processing and General Processing items.

4.2.3.2. Early Child Behavior Questionnaire (ECBQ)

The ECBQ (Putnam et al., 2006) is a parent-report standardised questionnaire that assesses 18- to 36-month-old infants' temperament according to 18 scales (see questionnaire and scoresheet with all the scales in Appendix 3 and Appendix 4). The original material comprises 201 items, but a short version has also been developed which assesses the same scales with 107 items (Putnam et al., 2014; Putnam & Rothbart, 2006). This was the version administered for the current study. The items were answered according to a 7-point scale (1: "never", 2: "very rarely", 3: "less than half the time", 4: "about half the time", 5: "more than half the time", 6: "almost always", 7: "always"). Higher scores indicate a more pronounced trait. The 18 scales have been shown to reliably cluster according to 3 non-overlapping factors: "Negative Affectivity" (8 scales), "Surgency Extraversion" (5 scales) and "Effortful Control" (5 scales). Scales related to infants' executive functions ("Attention Focusing", "Attention Shifting", "Inhibitory Control") are part of the "Effortful Control" factor, together with two other scales ("Cuddliness" and "Low Intensity Pleasure"). Thus, we were interested in the Effortful Control factor and its attention-related subscales in particular. The questionnaire was originally designed for children aged 3 to 7 years. The ECBQ was subsequently designed for 18- to 36-month-old populations, together with the Infant Behavior Questionnaire (IBQ) for 3- to 12-motnh-old infants. Although the earliest age-point IBQ (3-12 months) is appropriate for the current sample at the date of the lab visit already, that version of the questionnaire does not comprise measures of executive functions yet. Thus,

we did not administer the IBQ but chose to contact parents after their visit to the lab in order to obtain the scores of interest at the earliest age point available, post-visit.

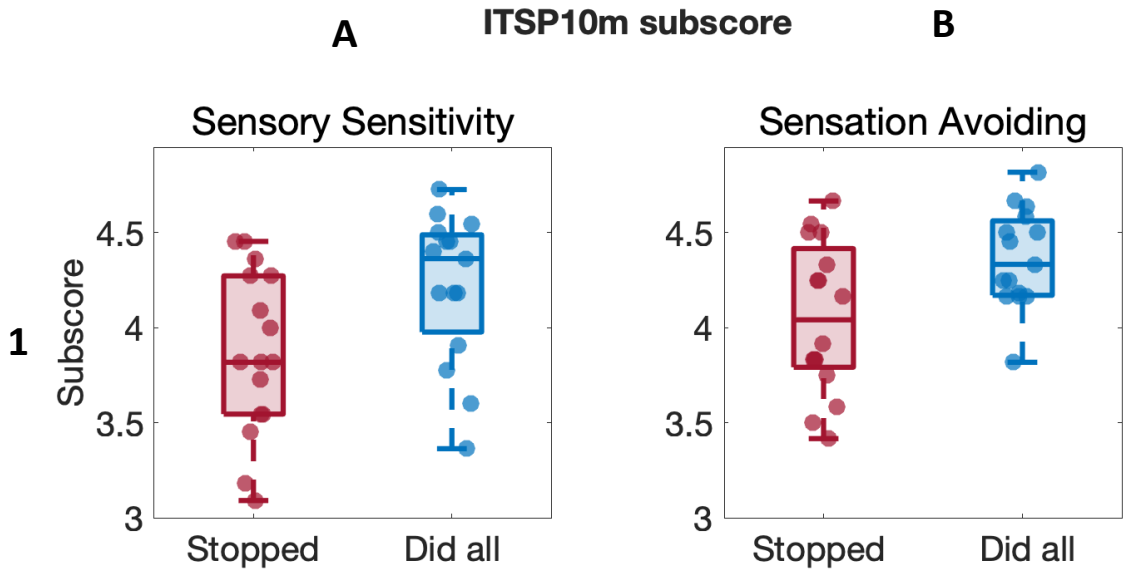
4.3. Results

4.3.1. Relationship between behaviour in the lab (time-on-task) and in daily life (scores)

In our first analysis, we looked at a comparison of the groups' scores in the two parent-reported questionnaires: the Infant/Toddler Sensory Profile (ITSP) which measured infants processing of sensory information at both age points (the in-lab visit at 10 months, and the follow-up at 17 months), and the Early Childhood Behavior Questionnaire (ECBQ) which measured infants' temperament at the later age point.

4.3.1.1. Time-on-task and sensory processing at test (ITSP at 10 months)

We expected infants who *did all* the experiment to show higher scores in their sensory responses in all or some of the four ITSP quadrants, reflecting generally less extreme reactions to sensory stimulation, both milder positive reactions such as seeking, and milder negative ones such as avoiding. Moreover, we expected that the infants who *did all* would also score higher in the visual processing scale specifically, since the stimuli were presented in the visual modality. Finally, we looked at general processing as a control to investigate potential differences stemming further than the modality at hand.



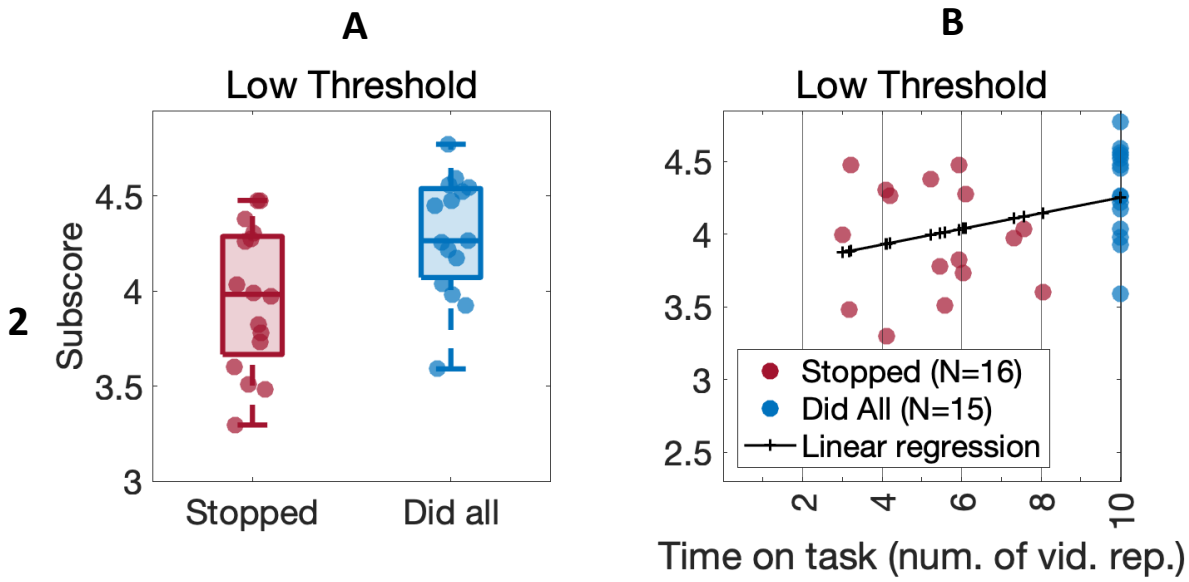


Figure 4.2: ITSP sub-scores at 10 months. The first line represents the sub-scores: Sensory Sensitivity (**1.A**) and Sensation Avoiding (**1.B**), and the second line represents the composite score Low Threshold (**2 A & B**). Each group is shown in a different colour: red: *Stopped*, blue: *Did All*.

1.A, 1.B, 2.A: The boxes represent the range between the 25th and 75th percentiles, the middle line signifies the data's median, and the whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range, none found for this data); dots depict all the individuals' sub-score values.

2.B: Individuals Low Threshold scores as a function of the time spent on the task (number of video repetitions); marker diameter is proportional to the number of individual values represented by the same point. The linear regression between both values given individuals data points is depicted as a black line on which crosses are drawn to figure the points' projection on the regression line.

Independent samples two-tailed t-tests revealed a significant and large group difference in infants' Sensory Sensitivity ($t(1,29) = -2.330$, $p = 0.027$, Cohen's $d = -0.837$) and Sensation Avoiding (t Welch $(1,25.954) = -2.596$, $p = 0.015$, Cohen's $d = -0.927$; Levene's $F(1,29) = 5.636$, $p = 0.024$) quadrants, but not in the Low Registration (t Welch $(1,20.567) = -1.088$, $p = 0.289$, Cohen's $d = -0.395$; Levene's $F(1,29) = 7.211$, $p = 0.012$) and Sensation Seeking ($t(1,29) = 1.027$, $p = 0.313$, Cohen's $d = 0.369$) quadrants. The Sensory Sensitivity and Sensation Avoiding quadrants make up the only combined quadrant of the questionnaire: the Low Threshold quadrant. Thus, the scores

were combined into a Low Threshold score, which also presented a significant difference between groups: t(1,29) = -2.632, p = 0.013, Cohen's d = -0.946. Moreover, using the time-on-task variable in a continuous manner rather than a binary group split, revealed a significant positive correlation between time-on-task and the Low Threshold combined quadrant score: Pearson's R = 0.373, p = 0.039. This suggested that in their everyday life, infants who *stopped* before the end of the experiment both noticed sensations (Sensory Sensitivity) and felt the need to avoid them (Sensation Avoiding) more than the infants who *did all* the experiment. This was in support of our hypothesis that infants who *stopped* before the end of the experiment were more reactive to sensory stimuli, compared to the infants who *did all* the experiment.

Groups were also compared in terms of General Processing (t(1,29) = 1.027, p = 0.313, Cohen's d = 0.369) and Visual Processing (t(1,29) = -0.051, p = 0.959, Cohen's d = -0.018) scores, for which no significant group difference was found. All in all, these results suggested that infants who *stopped* before the end of the experiment were more reactive to sensory stimuli than those who *did all*, irrespective of the sensory modality.

4.3.1.2. Time-on-task and sensory processing at follow-up (ITSP at 17 months)

Next, we looked at the ITSP scores reported by parents at follow-up in order to assess the stability of these effects. We expected effects visible at both time points, although possibly milder at follow-up due to the increase in noise with time passed between test and follow-up.

Independent samples two-tailed t-tests revealed no significant group difference in any of the quadrants: Low Registration (t(1,20) = 1.134, p = 0.204, Cohen's d = 0.563), Sensation seeking (t(1,20) = -0.592, p = 0.561, Cohen's d = -0.253), Sensory Sensitivity (Mann-Whitney U = 49.000, p = 0.486, Cohen's d = -0.183; Shapiro-Wilk test of normality for the group who *did all*: W = 0.792, p = 0.012), Sensation Avoiding (t(1,20) = -0.547, p = 0.591 Cohen's d = -0.234) or the combined quadrant Low Threshold (t(1,20) = -0.317, p = 0.754, Cohen's d = -0.136). There was also no significant correlation between the combined Low Threshold score and the continuous time-on-task variable (Pearson's R = -0.046, p = 0.838). Tests on processing types also revealed no significant effect of group, neither for the General Processing scores (t(1,20) = 1.096, p = 0.286, Cohen's d = 0.469) nor for the Visual Processing scores (Mann-Whitney U = 68.000, p = 0.617, Cohen's d = 0.133; Shapiro-Wilk test of normality for the group who *stopped*: W = 0.847, p = 0.034). This suggested that the individual differences in infants' Low Threshold scores at test that

related to individual differences in infants' response to the experimental stimuli, were not stable enough to be evidenced at follow-up, 7 months later.

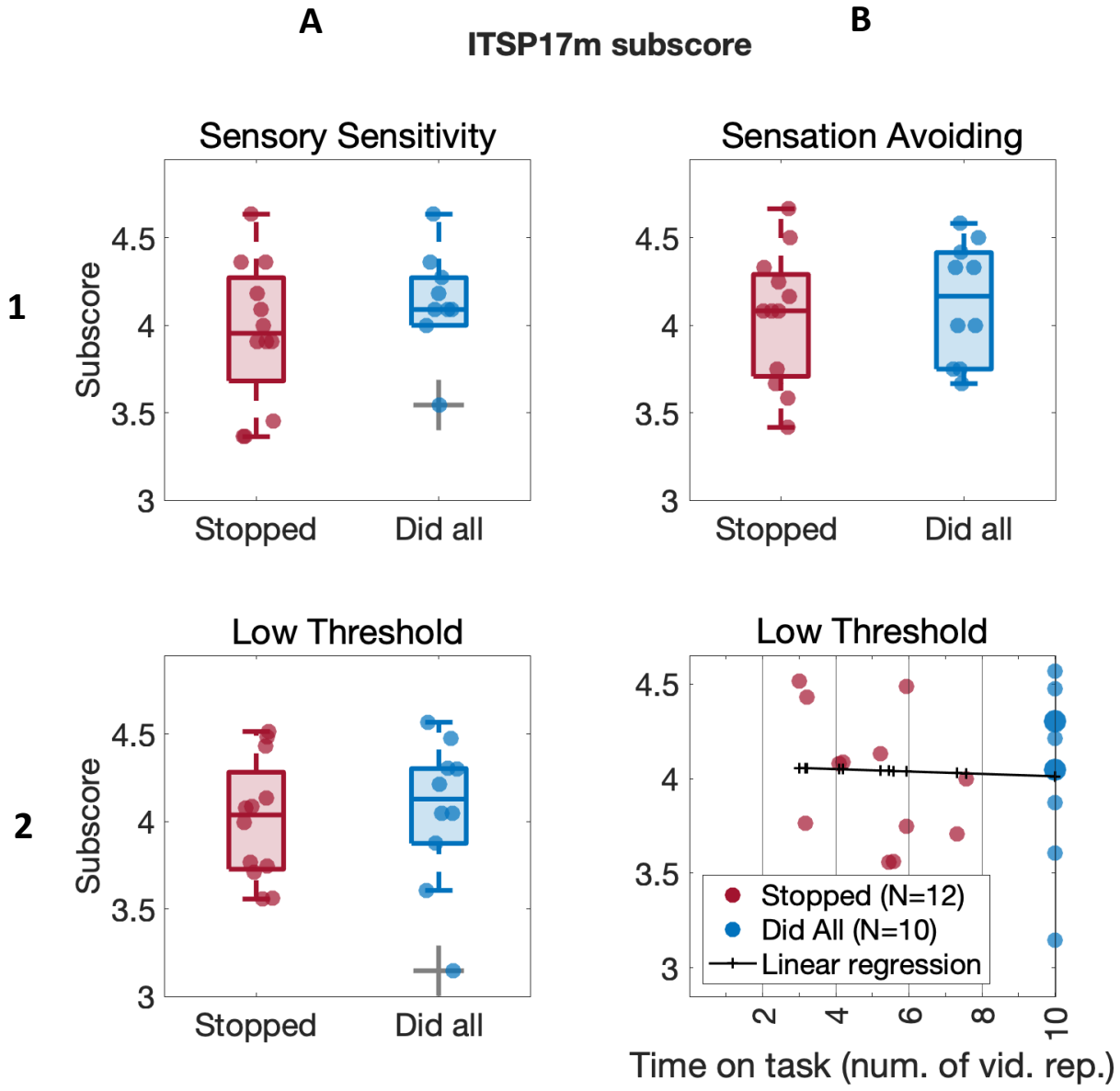


Figure 4.3: ISTP sub-scores at 17 months. The first line represents the sub-scores: Sensory Sensitivity (**1.A**) and Sensation Avoiding (**1.B**), and the second line represents the composite score Low Threshold (**2 A & B**). Each group is shown in a different colour: red: *Stopped*, blue: *Did All*.

1.A, 1.B, 2.A: The boxes represent the range between the 25th and 75th percentiles, the middle line signifies the data's median, and the whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range), which are represented as

grey crosses when present in the data; dots depict all the individuals' sub-score values.

2.B: Individuals Low Threshold scores as a function of the time spent on the task (number of video repetitions); marker diameter is proportional to the number of individual values represented by the same point. The linear regression between both values given individuals data points is depicted as a line on which crosses are drawn to figure the points' projection on the regression line.

Importantly, this change could not be easily explained by the fact that a sub-set of participants did not take part in the follow-up. Indeed, tests reported in Appendix 1 show that all the effects reported in the sample of participants at test date were also evidenced in the sub-sample of follow-up participants. Moreover, we found a significant correlation between the Low Threshold ITSP scores at 10 and 17 months: Pearson's R = 0.522, p = 0.013 (Figure 4.3A). Thus, it seemed that the Low Threshold score at either time point reflected a rather stable trait. However, in the context of this study, only the Low Threshold ITSP score at the former age-point was significantly related to the time-on-task measure, and not at the latter age-point. However, we found a significant correlation between time-on-task and the change in the ITSP Low Threshold score between 10 and 17 months: R = -0.524, p = 0.012 (Figure 4.3B). This was driven by a correlation between time-on-task and the change in Sensation Avoiding scores (R = -0.552, p uncorr. = 0.008, p BF corr. = 0.016) but not with the change in Sensory Sensitivity scores (R = -0.319, p uncorr. = 0.072, p BF corr. = 0.144). This suggested that the particular variations in the Sensory Avoiding scores at 10 months that were able to explain variations in infants' behaviour (time-on-task) during the experiment disappeared between 10 and 17 months, despite the relative stability of the overall Low Threshold scores.

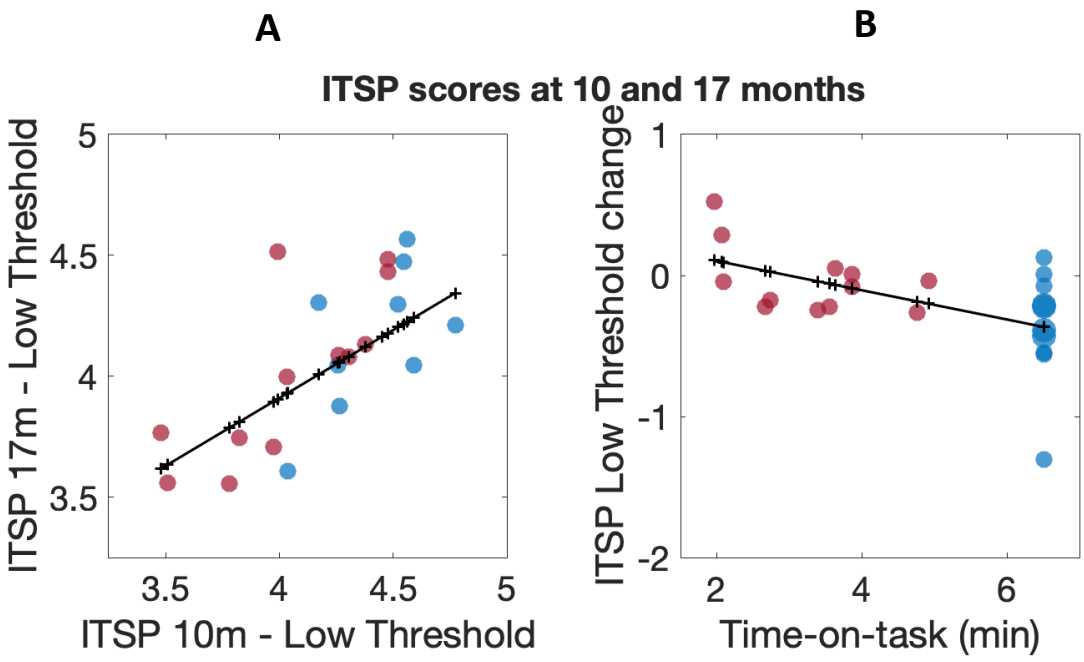


Figure 4.4: Relationship between the ITSP Low Threshold scores at each time point. Each group is shown in a different colour: red: *Stopped*, blue: *Did All*. Each point represents an individual. The linear regression between both values given individuals data points is depicted as a line on which crosses are drawn to figure the points' projection on the regression line.

A: Relationship between ITSP Low Threshold scores at 10 and 17 months.

B. Relationship between the change in ITSP Low Threshold scores between 10 and 17 months, and the time-on-task variable.

4.3.1.3. Time-on-task and temperament at follow-up (ECBQ at 17 months)

Finally, we looked at the ECBQ scores reported by parents at the follow-up point. We expected infants who *did all* the experiment to score higher in the Effortful Control Factor and particularly in the scales comprised in this 5-scale factor that reflect executive function abilities: Attentional Focusing, Attentional Shifting and Inhibitory Control. We did not expect effects in the other 2 scales comprising this factor: Low Intensity Pleasure and Cuddliness.

Independent samples two-tailed t-tests revealed a significant and very large effect of group on the Effortful Control factor ($t(1,20) = -2.919$, $p = 0.008$, Cohen's $d = -1.250$). Post-hoc tests on the 5 subscales revealed a significant, very large effect of group on the Attentional Focusing scale

which survived Bonferroni correction (t(1,20) = -3.052, p uncorrected = 0.006, p BF corrected = 0.03, Cohen's d = -1.307) but no significant effect on the other scales (Attentional Shifting: t(1,20) = -1.625, p uncorrected = 0.120, Cohen's d = -0.696; Inhibitory control Mann-Whitney U = 39.000, p uncorrected = 0.171, Cohen's d = -0.350, Shapiro-Wilk test of normality for the group who *stopped*: W = 0.828, p = 0.020; Low intensity pleasure: t(1,20) = -1.587, p uncorrected = 0.128, Cohen's d = -0.679; Cuddliness: t(1,20) = -1.221, p uncorrected = 0.236, Cohen's d = -0.523). Furthermore, tests performed on the continuous time-on-task variable revealed a significant positive relationship with the Attentional Focusing score: Pearson's R = 0.554, p = 0.008.

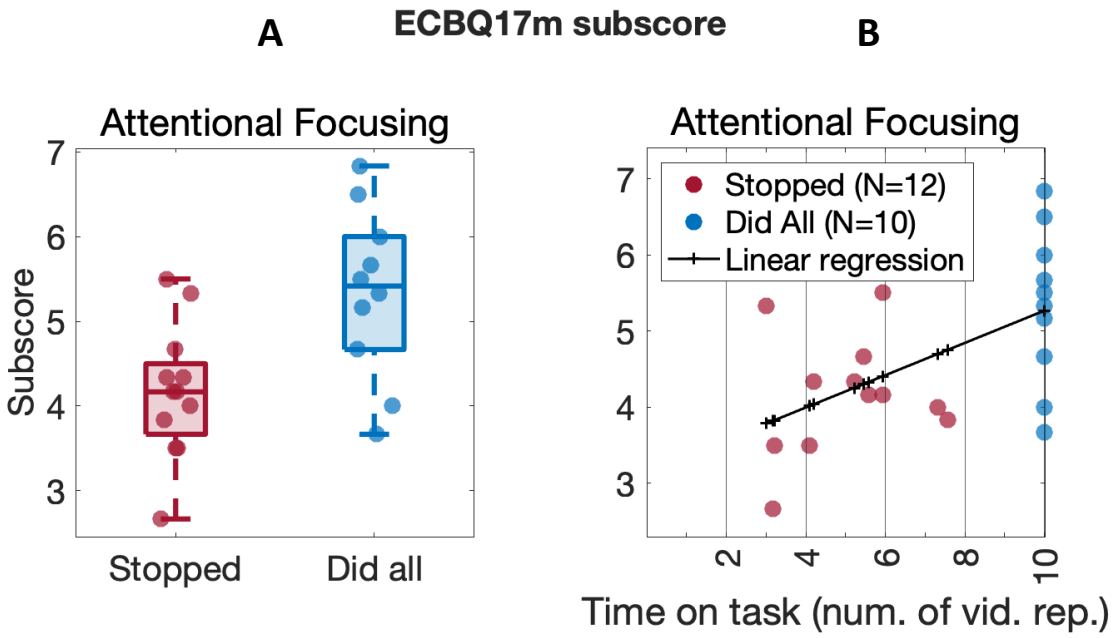


Figure 4.5: ECBQ Attentional Focusing sub-scores at 17 months. Each group is shown in a different colour: red: *Stopped*, blue: *Did All*.

A: The boxes represent the range between the 25th and 75th percentiles, the middle line signifies the data's median, and the whiskers depict the range, discarding outliers (points outside of 1.5 times the interquartile range), which are represented as grey crosses when present in the data; dots depict all the individuals' sub-score values.

B: Individuals Low Threshold scores as a function of the time spent on the task (number of video repetitions); marker diameter is proportional to the number of individual values represented by the same point. The linear regression between both values given

individuals data points is depicted as a line on which crosses are drawn to figure the points' projection on the regression line.

This suggested that the infants who *did all* the experiment tended to sustain their attention to objects and resisted distraction in their daily life significantly more than the infants who *stopped* before the end. This result was in support of our hypothesis that infants who *did all* the experiment would have stronger executive function skills compared to those who *stopped* before the end, although only the Attention Focusing scale and not the Attention Shifting or Inhibitory Control scale showed a significant effect.

4.3.1.4. Relationship between the different questionnaire scores

Next, we investigated the relationship between the different questionnaire scores identified as relevant. We found a significant positive correlation between infants' threshold for responding to sensory stimuli at 10 months (Low Threshold ITSP score, reflecting both their tendency to avoid stimuli and their sensitivity to them) and their ability to focus on stimuli at 17 months (Attentional Focusing ECBQ score): Pearson's R = 0.456, p = 0.033. This suggested that the individual differences evidenced at the two different age points in two different questionnaires were linked.

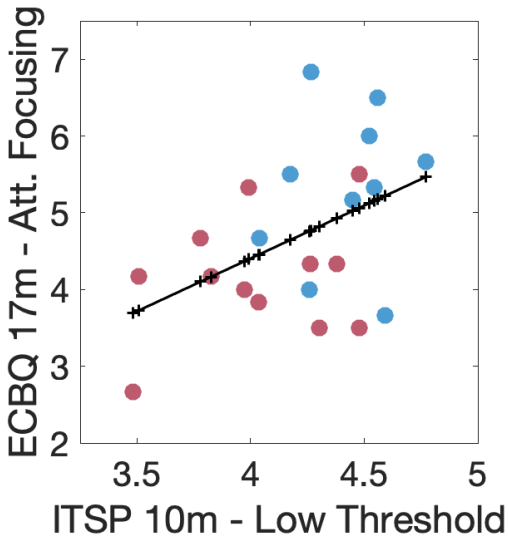


Figure 4.6: Relationship between the ITSP Low Threshold scores at 10 months and ECBQ Attentional Focusing score at 17 months. Each group is shown in a different colour: red: *Stopped*, blue: *Did All*. Each point represents an individual. The linear regression between

both values given individuals data points is depicted as a line on which crosses are drawn to figure the points' projection on the regression line.

4.3.1.5. Effect of gender on the questionnaire scores

Finally, infants were split by gender rather than group to investigate a possible role of gender on the Low Threshold, Sensory Sensitivity and Sensation Avoiding ITSP scores and the Attentional Focusing ECBQ score. Independent sample t-tests revealed no significant effect of gender on either ITSP score at 10 months (Low Threshold: $t(1,29) = -1.692$, $p = 0.101$, Cohen's $d = -0.611$, Sensory Sensitivity: $t(1,29) = -1.602$, $p = 0.120$, Cohen's $d = -0.578$, Sensation Avoiding: $t(1,29) = -1.549$, $p = 0.132$, Cohen's $d = -0.559$), nor on the ECBQ scores at 17 months (Effortful control: $F(1,19) = 0.088$, $p = 0.770$, $\eta^2 = 0.005$; Attentional focusing: $F(1,19) = 1.315$, $p = 0.266$, $\eta^2 = 0.064$). No gender split was performed on ITSP scores at 17 months since there were no group effects to investigate at this data collection point.

4.3.1.6. Relationship between brain response at the lab and behaviour in daily life

To this point, the analysis of the questionnaires data indicated that our experimental measure of behaviour, namely the time-on-task, reflected individual differences in traits that were visible to parents in their infant's daily-life behaviour, and not only variations in attentional states during the experiment. This was true when using the time-on-task variable either as a binary split for group comparisons or as a continuous variable for correlations. Specifically, we found that infants' time-on-task was positively correlated to both their ability not to be distracted by sensations (Sensory Sensitivity) and not to limit the sensations available (Sensation Avoiding), which were aggregated in a composite score reflecting their sensory reactivity (Low Threshold). This relationship was significant at the test date, but not 7 months later. What's more, we also found a significant, very large effect of the time-on-task variable on the infants Attentional Focusing at 17 months. This validated the usefulness of our in-lab measure to capture individual differences that span outside of the lab setup and translate to individual differences in infants' daily life.

In the former chapter, we have also shown that the same experimental measure of attentional behaviour (time-on-task) was related to infants' brain activity in response to distractors and relative to their response to catchy stimuli. Thus, our next step was to combine those results in a final analysis to link infants' in-lab brain activity and their daily-life behavioural questionnaire

scores. Motivated by the analyses described in the former chapter, we used correlations to compare the questionnaires scores to two measures of infants' brain activity: the strength of the gamma response to the distractors, and its alignment (SCA value at lag 0) with the ERP. Infants' brain responses were correlated with the two questionnaire scores identified as relevant in section 4.3.1: their Low Threshold score at 10 months as well as their Attentional Focusing score at 17 months. Similarly to what was done in the analyses of Chapter 3, and because the groups analysed were by definition exposed to different amounts of trials, two steps were taken to both mitigate and account for potential residual effects of the number of trials variable (Ntrials) on the experimental measures. On the one hand, the brain measures were extracted using a downsampled set of trials for the group who *did all*, in order to match the trial number of the group who *stopped*. On the other hand, the analyses were performed using individuals' number of included trials as a conditioning variable for partial correlations.

4.3.1.7. Strength of the gamma response and questionnaire scores

First, we looked at the relationship between infants' gamma response to the distractors and their questionnaire scores. We used the same measures as in Chapter 3: the gamma response (at the maximum frequency and averaged over time) at the start of the experiment (video repetition 1-2) as well as for the whole experiment. We expected the strength of the gamma response (at the beginning of the experiment or over the whole experiment) to be negatively correlated with both the ITSP Low Threshold score at 10 months and the ECBQ Attentional Focusing score at 17 months, meaning that a higher gamma response would be associated with a lower ability to sustain attention but a greater reactivity.

Partial (controlling for variations in Ntrials) Pearson correlations between the strength of the gamma response over the whole experiment and questionnaire scores showed no significant relation between the overall in-lab brain response and the reported daily-life sensory behaviour at 10 months (Low Threshold: $R = -0.247$, $p = 0.189$; Sensory Sensitivity: $R = -0.194$, p uncorr. = 0.305; Sensation Avoiding: $R = -0.277$, p uncorr. = 0.138) or with the Attentional Focusing ECBQ score at 17 months ($R = -0.302$, $p = 0.184$).

However, looking at the relation between in-lab brain response at the start of the experiment only (Vid. Rep. 1-2) and reported daily-life behaviour revealed a significant correlation (computed partially, conditioned on Ntrials) between the strength of the gamma response over the start of

the experiment (video repetitions 1-2) and the Low Threshold ITSP score at 10 months (R = -0.406, p = 0.026). Post-hoc tests on the two subscales revealed a significant effect of Sensation Avoiding but not Sensory Sensitivity (respectively, R = -0.439, p uncorr. = 0.015, p corr. = 0.030; and R = -0.333, p = uncorr. = 0.072, p corr. = 0.144). To note, it is likely that we lacked some power to evidence all the effects at play. Especially in the case of the Sensory Sensitivity score, the large size of the correlational effect and the existence of a trend before correction suggest that this effect might be present but not significantly evidenced in our small sample. Thus, we will not discuss the selectivity of this effect for either the sub-score but for the Low Threshold as a whole. No relationship was found between the strength of the response at the start of the experiment and the ECBQ Attention Focusing score at 17 months though (R = -0.110, p = 0.635).

This analysis showed partial evidence for a link between infants' brain response to distractors in our experiment and their parent-reported daily-life behaviour in response to sensory information, mainly their tendency to avoid sensory information.

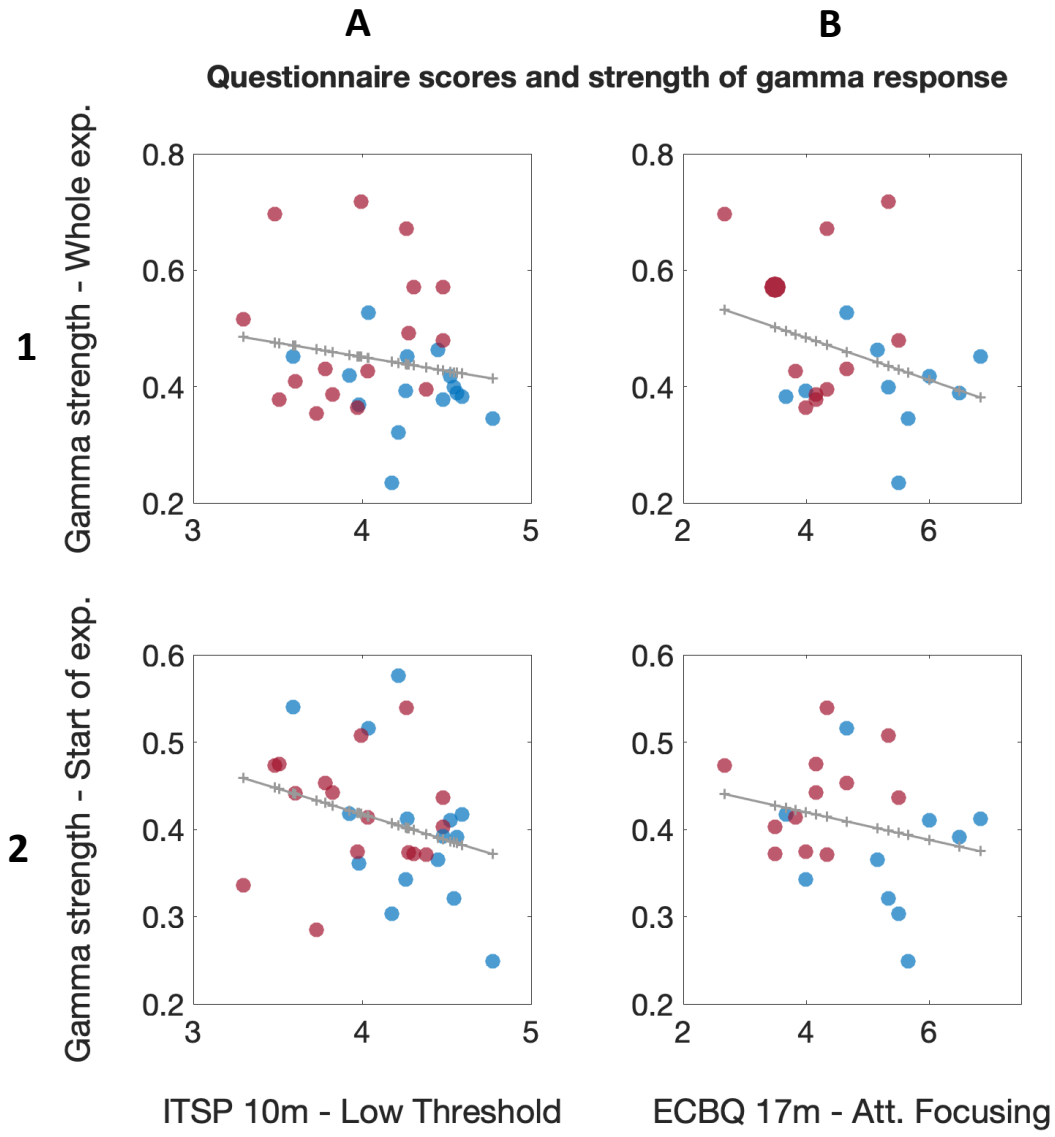


Figure 4.7: Relationship between the scores identified as relevant (horizontal axes) and the strength of the gamma response (vertical axes). Each group is shown in a different colour: red: *Stopped*, blue: *Did All*. Each point represents an individual; marker diameter is proportional to the number of individual values represented by the same point. The linear regression between both values given individuals data points is depicted as a line on which crosses are drawn to figure the points' projection on the regression line. Non-significant correlations are drawn in grey colour.

1: Strength of the gamma response over the whole experiment.

2: Strength of the gamma response over the start of the experiment.

A: ITSP Low Threshold scores at 10 months.

B: ECBQ Attentional Focusing score at 17 months.

4.3.1.8. Alignment of the gamma response and questionnaire scores

Finally, we investigated the relationship between the alignment of infants' gamma response and their questionnaire scores. How infants' fast gamma responses aligned with their slower ERP response was evaluated using Scaled Correlation Analysis (SCA, see Chapter 3 sections 3.2.5.3 for Methods and 3.3.3 for Results) over the start of the experiment (video repetitions 1-2) but not over the whole experiment. Indeed, our previous analyses showed that the alignment of these fast and slow rhythms varied along the experiment and that an overall alignment response would not represent this data well.

Partial (conditioned on Ntrials) Pearson correlations between the gamma-ERP alignment values (SCA) over the start of the experiment and the Attentional Focusing ECBQ score at 17 months showed no significant relation between in-lab brain response and reported daily-life behaviour ($R = 0.316$, $p = 0.163$). However, Pearson correlations between the gamma-ERP alignment values (SCA) over the start of the experiment and the Low Threshold ITSP score at 10 months revealed a trending relationship between in-lab brain response and reported daily-life sensory experience in terms of Low Threshold ($R = 0.340$, $p = 0.066$). As the Low Threshold score is a composite score made of two scores, we went on to investigate the individual contribution of each sub-score to this relationship. We found a significant correlation between the gamma-ERP alignment at the start of the experiment and the Sensory Sensitivity scores ($R = 0.406$, p uncorr. = 0.026, p corr. = 0.052) but not the Sensory Avoiding scores ($R = 0.215$, p uncorr. = 0.255, p corr. = 510). Again, it is possible that we lacked some power to evidence all the effects at play although this time difference in the two sub-scores' effect size was larger as one effect was almost twice as big as the other, and the significance level of the correction not evidenced was further away from the significance threshold. Furthermore, the composite score only revealed a trending relationship, suggesting that combining scores added noise and pointing to a true specificity this time.

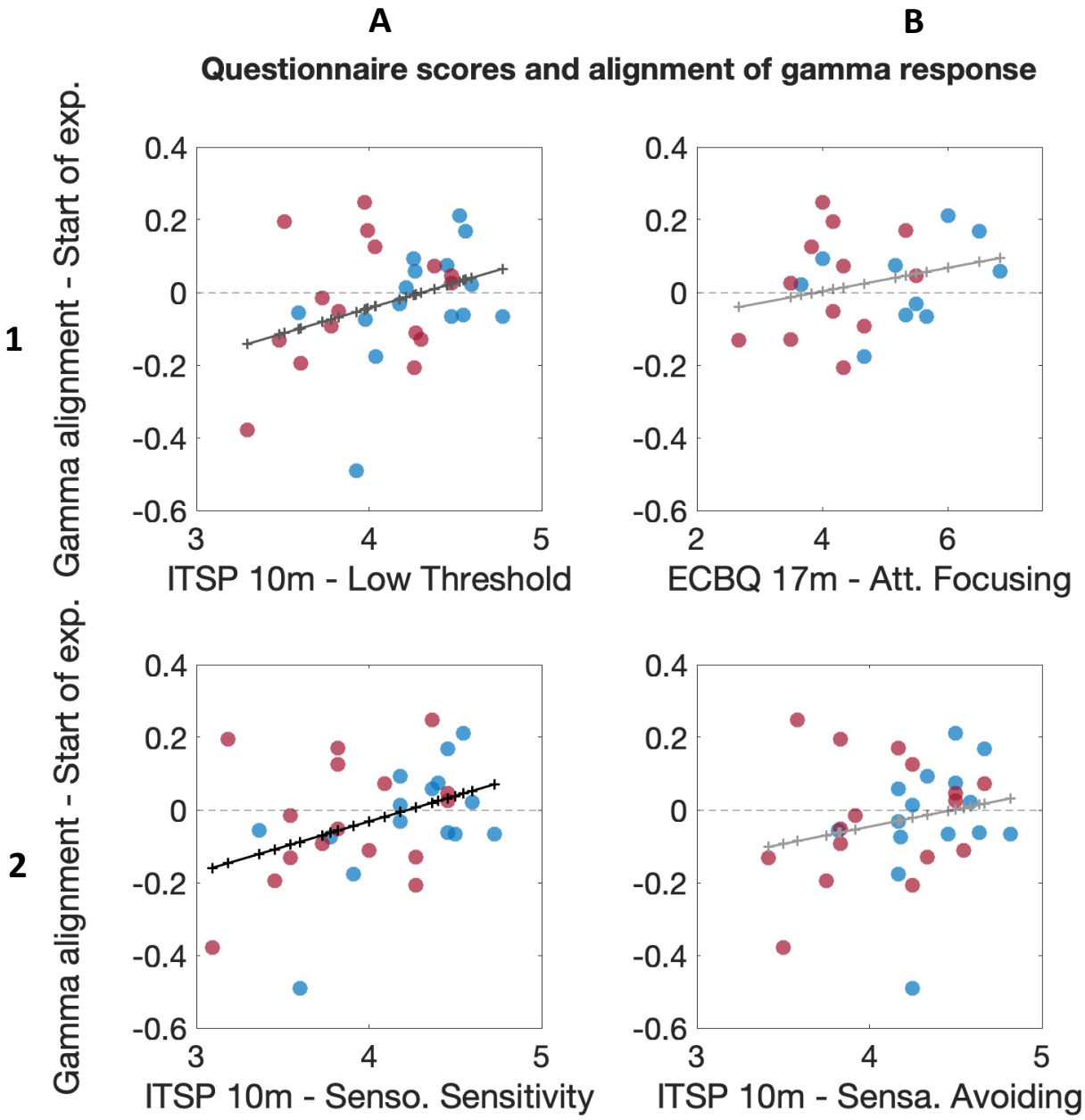


Figure 4.8: Relationship between the scores identified as relevant (horizontal axes) and the alignment of the gamma response (gamma-ERP SCA at lag 0; vertical axes). Each group is shown in a different colour: red: *Stopped*, blue: *Did All*. Each point represents an individual. The linear regression between both values given individuals data points is depicted as a line on which crosses are drawn to figure the points' projection on the regression line. Non-significant correlations are drawn in grey colour and significant ones in black.

1: Main relevant scores: ITSP Low Threshold scores at 10 months (1A) and ECBQ

Attentional Focusing score at 17 months (1B).

2: ITSP Low Threshold sub-scores at 10 months: Sensory Sensitivity (2A) and Sensation Avoiding (2B).

These results suggested that how infants modulated the alignment of their fast, gamma-range activity in response to the distractors and relative to the video, with the slow ERP rhythm, was significantly linked to how they responded to sensory stimuli in their daily-life, and more specifically how they notice and respond to sensations.

All in all, our analysis of the relationship between brain and behaviour in this study showed evidence for a direct correlation between the two. The strength of the gamma response was significantly correlated with their sensory reactivity (Low threshold) at test but not attention (Attentional Focusing) at follow-up. Gamma response also significantly correlated with one sensory experience subscale at 10 months.

4.4. Discussion

4.4.1. Summary of findings

4.4.1.1. Infants' engagement with our paradigm is linked to their attention and sensory processing in everyday life

We have shown that infants' behaviour in the lab when engaging with our paradigm (time-on-task) was linked to their behaviour outside of the lab (ITSP Low Threshold score and ECBQ Attentional Focusing score), as reported by parents, which validates the usefulness of our paradigm. Importantly, significant correlations between two relevant questionnaire scores (Low Threshold and Attentional Focusing) and the time-on-task variable were found both at test (Low Threshold) and at a 7-month follow-up (Attentional Focusing). This suggests that the effects evidenced through our experimental set-up captured robust individual differences, still exhibited by the same infants after a long period of time.

4.4.1.2. Infants' brain response in our paradigm is linked to their sensory Low Threshold in everyday life

We have also evidenced a direct correlation between infants' *brain responses* to distractors during the study at the lab, and their *daily-life behaviour* in response to sensory stimuli, namely their Low Threshold score at test. No link between brain response and Attention Focusing was evidenced. However, this does not rule out a relationship between the two, and these effects might require more statistical power to be evidenced.

4.4.1.2.1. Gamma strength

We found that infants' *strength* of gamma response at the start of experiment but not over the whole experiment was linked to their *Low Threshold* sensory score. This effect might have been specifically driven by their Sensation Avoiding score, although their Sensory Sensitivity score seemed to present a strong yet non-significant link with the brain response too, which we might have lacked statistical power to evidence, and which suggests a general *Low Threshold* effect.

4.4.1.2.2. Gamma alignment

Finally, we found a link between infants' *alignment* of their fast gamma rhythms in response to the distractors with their slow ERP rhythms at the start of the experiment, and their Low Threshold score, which appeared to be specifically driven by their Sensory Sensitivity score.

These results provide strong evidence that, despite the scarcity of research looking into fast rhythms in infants, individual differences in infants' fast brain responses are relevant to their behavioural response to stimuli, both in and out of the lab. Moreover, this finding points to the importance of fast gamma-range activity for information processing and selection in infants, not only on its own but *relative* to a slower, more global rhythm in the infant brain, namely the ERP. While covariations between fast and slow rhythms in the adult brain have been established as central to information processing (Fries, 2015), to our knowledge such a relationship was never evidenced in infants until now. Here, we link our measure of gamma-ERP (fast-slow) alignment not only with infants' behaviour during the study (time-on-task, see Chapter 3) but also with infants' behaviour in their daily life (Sensory Sensitivity score, current chapter), which consolidates this novel finding.

4.4.2. Implications for theories of attention and sensory processing in infants

Our results point out the importance in the infant brain of both gamma activity's *strength* and its *alignment* with slower rhythms for sensory processing and attentional processes, something that was never shown before and has far reaching theoretical and methodological implications for the field. On the one hand, it opens a host of new possibilities for the study of the infant brain, on the other hand, it refines the theories of information processing and selection in infants.

Importantly, gamma *strength*, which was evidenced to relate to infants' behavioural engagement in the former chapter, was selectively linked to *Low Threshold* sensory score at test, linking three levels of infant measures: brain measures, behaviour measures and trait measures. This effect might have been driven by the Sensory Avoiding scores specifically. Gamma *alignment* on the other hand appeared to significantly correlate with a sub-score of the *Low Threshold* score, the Sensory Sensitivity scores at test, rather than the whole scale which remained under the significance threshold albeit a trending effect. The specificity of the responses seems difficult to discuss in the context of this study though, as statistically powered appeared a bit limited to

evidence all the effects. Still, we found that the *Low Threshold* score in general was linked to gamma responses, which generally followed our hypothesis.

However, we had anticipated that *Low Threshold* would more specifically relate to the *strength* of the response, and executive attention scores to the *alignment* of the response, which was not evidenced. Indeed, according to theories of information processing in the brain (Fries, 2015), occipital gamma activity is thought to hold information related the local processing of visual sensory information, while its alignment with slower rhythms represents a means to integrate it given more global information, such as higher cognitive processes like attention.

Firstly, we did not find evidence of a link between brain responses and trait attention. There was however a link between infants' Attentional Focusing score and their engagement with the stimuli (time-on-task), itself linked to brain response differences. This suggests that Attention Focusing was generally relevant and given the relatively limited sample size of this exploratory study, we suggest that future work with more statistical power is needed to rule out or evidence the hypothesised link between brain response and attention in this paradigm. It is also possible that attentional mechanisms are not well described by local couplings in the visual cortex, and that investigating how slow rhythms arise in integrative regions such as the pre-frontal cortex would encompass more of the complexity of these mechanisms than focusing on their influence directly in the visual cortex. An interesting avenue to further investigate is the study of cross-regions cross-frequency couplings in the infant brain. It is also possible that infants' attentional skills were not developed enough to influence their processing of the stimuli at 10 months. Indeed, attentional skills only develop during the second half of the first year of life and are thought to arise around the age that was tested in this study (Lawson & Ruff, 2004; Xie et al., 2018). Future work both with younger and older infants would help to clarify this question.

Secondly, we found that *Low Threshold* did not seem to specifically map onto either measure. It is possible that this lack of specificity is related to the fact that the infant brain is still developing and possibly more generally reactive to stimuli than the adult brain: this reactivity (*Low Threshold*) could then affect not only their local sensitivity to information (*strength*) but also their general selection of information (*alignment*). Indeed, in the visual domain, infants' attention has been shown to undergo developmental changes going from a more stimulus-driven salience-related selection of information, to a more goal-driven selection of information (Colombo, 2001). It might

be the case that these processes diverge and specify later on. Again, future work with older as well as younger children could shed light on how these processed develop and specify.

4.4.3. Power limitation

Aside from these important findings, it's worth noting that an important part of the correlations tested in this chapter came out non-significant. These negative results are difficult to interpret as they do not bring evidence against the hypotheses that motivated them, but rather fail to bring evidence for the presence of the effect tested. Indeed, it is possible that these effects do exist in the population but that our sample size did not provide enough power for evidencing them. Working on a small sample size was useful for explorative work such as ours: we looked at measures motivated by work in the adults' and animals' literature, but rarely or not yet investigated in the developmental field. Using a small sample size enabled a fast exploration of our novel ideas in order to establish their relevance and significance before embarking on a more involved investigation. However, this small sample size is only powerful enough to unveil large-sized effects. Thus, future work to increase the sample size of the study or replicate it in a new, larger sample would be useful, especially for the investigation of the link between questionnaire scores and brain response. This is also motivated by the fact that we worked not only with an infant population but also with fast brain rhythms, both substantial sources of additional noise, which calls for more statistical power in the future.

4.4.4. Distractibility: adaptive behaviour and atypicality

It is also important to note that all the infants included in this study had no reported history of neurological or psychiatric health issues in their nuclear family: the range of individual differences observed in our sample is thought to be representative of variations in the general population. Thus, there is no reason to believe that either group of infants was behaving more optimally than the other. Despite the negative connotation associated with the term *distraction* in our everyday life, paying too much attention to incoming information can be detrimental just as well as paying too little attention and being distracted too easily. Being distractable is useful to avoid the dismissal of new relevant information, while maintaining it a reasonable level helps gaining enough knowledge from one source of information before shifting to another. While infants in our experiment seemed to adopt different strategies for this trade-off, favouring either one or

the other behaviour depending on their group, it is likely that both strategies could be a source of learning opportunities for them. An interesting future direction for this work would be to investigate how this trade-off is made in infants at risk of developmental disorders such as autism spectrum disorder (ASD) or attention deficit and hyperactivity (ADHD), in which sensory processing and attention mechanisms are impaired. Both conditions are usually diagnosed after 2 years of age, so being able to identify neural and behavioural markers of risk at 10 months would be particularly relevant. Notably, ASD has been described as a selective imbalance in short-range vs. long-range circuitry, with overconnected local networks (Peters et al., 2013; Snijders et al., 2013). The analyses developed here offer new tools for tracking infants' local brain activity, with potential for extending ASD diagnostic tools at young ages.

4.5. Conclusion

Combining analyses of infants' behavioural and neural responses to our visual paradigm with parent-reported measures of infants' sensory experiences and attention, we were able to link infants' brain activity during the study (strength and alignment of their gamma responses), their engagement during the study (time-on-task), and their sensory and attentional behaviour in their daily life (ITSP Low Threshold score and ECBQ Attentional Focusing score). We showed that infants who stayed until the end of the experiment exhibited 1) higher Attentional Focusing scores 7 months after and higher Low Threshold scores at test, but also (see Chapter 3) 2) an overall lower gamma response to the stimuli and a more modulated alignment of the gamma activity with the ERP rhythm. There was a direct correlation between Low Threshold scores and the gamma-ERP strength and alignment at the start of the experiment, providing evidence for a direct link between brain activity in the lab and behaviour in the lab on the one hand, and behaviour in the lab and behaviour in daily life on the other hand, but also between brain activity in the lab and everyday behaviour. All in all, this and the former chapter bring in new and important information on basic bricks for active learning: how information selection processes are implemented in infants' brains and how they translate to behavioural changes in the lab and in their everyday life. Yet, the question of whether the mechanisms that dictate which information is selected on a theoretical level remain unexplored. Do infants select information based on the same types of learning theories as presented in the general introduction of this thesis?

Chapter 5: Testing the learning-progress hypothesis in 15-month-old infants

5.1. Introduction

Infants are astounding learners who are constantly gaining new knowledge about the world around them. They do not merely passively absorb the information that they come across but play an active role in selecting which information to learn from. Piaget's seminal rattle experiment (1952) showed that already at 4 months of age, infants are capable of adjusting their own actions to control the amount of stimulation that they are getting. When attaching a rattle to their hand, he found that infants first accidentally move and shake the rattle, fortuitously discovering the relationship between their hand movements and the toy's multisensory stimulation, and then proceed to increase their hand movements in order to get more of this stimulation (Thelen, 1994). Interestingly though, while rattles are universally catchy to young infants, they do not appeal as much to older infants: having learned what outcome to expect from shaking a rattle, older infants lose interest in such simple toys (Oudeyer & Smith, 2016). Indeed, it appears that infants choose their own sources of information depending on their own interest, which can vary in time, with an individual's developing knowledge and abilities, but also between individuals at a given developmental stage. How is it then that infants choose what information to engage with it?

5.1.1. Curiosity theories

Several theories have been advanced to explain what makes information worth sampling for a learner, or in other words, what piques their curiosity. The main ideas and studies that shaped the field of curiosity were presented in more detail in the general introduction of this thesis. They tackle the topic from different angles and focus either on defining what makes information sampling rewarding despite its cost, or characterising the effect of various information properties

on the process. While such theories might seem like largely different takes on the matter, they are not mutually exclusive and could instead be seen as different pieces of the same puzzle. Indeed, on the one hand, reward is a core aspect to making the process of sampling information generally worthwhile, and on the other hand, looking into the properties of the information itself allows to refine the mechanism and predict what specific information will be sampled out of the vast amount of information present in the environment. Theories that focus on characterising what aspects of information drive the learner's interest all gravitate around the idea that worthwhile information offers the agent opportunities for learning i.e., building a coherent picture of its surroundings. Exactly what aspect of the information is evaluated by the agent to come to this conclusion however varies depending on the theories. They can roughly be grouped along three main currents, quickly reframed bellow with references to what has been shown in infants for each theory.

5.1.1.1. Discrepancy theories

Discrepancy theories propose that learners form priors and predictions about their environment, which, when they do not match the input received, create an incongruity, conflict or surprise, which in turn motivates learners to gather more information on the topic and update their internal representations (Berlyne, 1954; Hunt, 1963; Kagan, 1972; McCall & McGhee, 1977). We will refer to this mismatch as prediction error (Den Ouden et al., 2012). Importantly, the information alone is not the driver of the seeking behaviour here, but rather the learning opportunity borne by information that generates a prediction error, to update the agent's conflicting internal model of the world. One piece of information might be crucial for one agent to disambiguate a situation and reframe their internal representations and thus drive their interest, while it might not resolve any conflict given the representations of another agent, thus leaving the latter uninterested in the same piece of information that highly attracted the former. In young populations, these accounts are backed by results showing that 11-month-olds' visual exploration is longer for surprising compared to unsurprising stimuli (Stahl & Feigenson, 2015) and that 2-3-year-olds prefer to play with toys that violated their expectations (Sim & Xu, 2017).

5.1.1.2. Information-gap theories

Information-gap theories posit that a learner's interest is driven by the existence of a gap in their knowledge i.e., a missing piece of information for them to form a full representation

(Loewenstein, 1994). Such theories are close but deviate from discrepancy theories in that the learner does not have to form expectations for the gap to exist. However, in information-gap theories, prior relevant knowledge is central for interest to arise: a learner won't be interested in information that they already know, but it is similarly not possible for them to be curious about something about which they have no knowledge. There appears to be an optimal gap-size that drives agents' interest (Kang et al., 2009), the quantification of which remains to be better defined. In infants, a related "Goldilocks' effect" was evidenced, according to which the relationship between infants' interest and the complexity of the information follows a bell curve, with peak interest at intermediate complexity levels (Kidd et al., 2014).

5.1.1.3. Learning-progress theories

Finally, learning-progress theories posit that an agent's interest in information is driven by how much learning progress they expect to make using the information that they are about to engage with (Kaplan & Oudeyer, 2007b). According to such theories, a learner will look to maximise their learning progress, monitoring it over time to decide when to engage or disengage with a source of information. In this framework, the agent monitors the *changes* in the prediction error and not the error itself. This enables them to be flexible in case the task that they are engaged with is not challenging enough or in case they remain stuck with something that they do not manage to learn from. This idea is closely related to the former theories but offers a quantitative criterion for drawing the function of how interest changes given different types of information and does not require prior knowledge for progress to be expected. This hypothesis has been tested with robot models programmed to optimise their learning progress in environments similar to infants' playgrounds. They showed complex exploratory behaviours akin to the early exploratory behaviours exhibited by infants (Oudeyer et al., 2005). This suggests that infants might be monitoring their learning progress to guide their exploration, but there is currently no direct evidence that they do.

5.1.2. Eye-tracking paradigms to test these hypotheses in infants

Several prominent eye-tracking paradigms exist to test these hypotheses with infants, namely habituation paradigms, novelty response paradigms and violation of expectation paradigms. They typically measure increases or decreases in infants' fixation time as measures of interest. In habituation paradigms, infants are repeatedly presented with the same stimulus until they have

learned as much as they need to and start to disengage with it, as indicated by shorter fixation times. Novelty response experiments build on this finding but go a step further, capitalising on infants' tendency to regain interest in a stimulus that they have been habituated with if a change is introduced. Such experiments are often used to demonstrate whether infants are capable of encoding the dimension of the change within their internal representations. Finally, in violation of expectation paradigms, infants are presented with information from which they can draw predictions about future information to come. In cases when their predictions are violated, they have been shown to fixate the stimuli longer (Baillargeon & Graber, 1987).

These types of paradigms are compatible with any of the three above-mentioned theories: for example, infants will lose interest in information that they have been habituated with because there is no more discrepancy, knowledge-gap or learning-progress for them to engage further. Similarly, new information or information violating their predictions is symptomatic of a discrepancy, knowledge-gap or expected learning-progress. While there exists evidence in infants that is compatible with the first two theories, there isn't any infants' data to back the newest, learning-progress theories.

5.1.3. Current study

This chapter aims to bring in new knowledge about whether infants can indeed monitor their learning progress to guide their interest in engaging with a source of information. To this end, a variation of the former paradigms was used, in which infants learn a rule and are able to progress and improve their performance. By monitoring infants' learning progress i.e., performance improvement throughout the task, in relation to their engagement with the task, we can use this paradigm to test the learning-progress theory in infants: is engagement with the task a function of learning progress? If evidence for the theory is found, this paradigm can help to show that learning progress is sufficient for driving infants' interest and that neither a discrepancy nor a knowledge-gap, absent from this design, is necessary for explaining how infants get interested in information. This however does not disqualify the other theories: interest can be driven by different parameters in different contexts, and more data would be needed to fully tease apart these theories and invalidate any of them.

The paradigm is drawn from a study from Hochmann et al. (2016), itself inspired from Kaldy et al.'s (2015) paradigm based on the game "Memory". In their study, Hochmann et al. sequentially

presented 14.5-month-olds with two shapes drawn onto cards on a TV- screen for 4s, after which a third shape was presented for 2s (test phase, including 1s of animation and 1s of anticipation) which matched either one of the first two options. Finally, the matching option started spinning for 1s in lieu of a short reward that infants would only perceive if they would anticipate it by looking at the matching card before the end of the anticipation period. They showed that at 14.5 months, infants are capable of learning both match and non-match rules (i.e., anticipate the location of the reward), as shown in the matching rule condition by increased fixation times for the matching option as opposed to the non-matching option over the course of the experiment (and vice versa for the non-matching rule).

In the current study, we made critical changes to the paradigm that allow us to follow infants' changes in performance or learning progress throughout the task. To this aim, infants were presented with three options rather than two, lowering the probability that an infant would look at the matching option by chance to 0.33 rather than 0.5 and making it easier to determine whether they perform above chance. To give infants enough time to scan all elements of the scene, a longer 5-s test phase was used instead of the original 2s allowed in Hochmann's paradigm. Moreover, in order to motivate infants to make progress from one trial to the other, we introduced a gaze-dependant outcome offering them a stronger sense of agency. Infants were presented with a shiny, rewarding animation of the target if they sustained their gaze on it (reward), or were presented with a simpler, orienting animation of the target if they did not (feedback). This feedback design enabled them to learn from incorrect trials as well as correct ones. Finally, in order to investigate how infants' autonomously gain or lose interest in information based on their learning progress rather than due to other external factor, the experiment was conducted with no interference from the parent or experimenter meant to drive the child's interest back on the task. Infants notoriously gain and lose interest in computer tasks easily, and various foods, social directives such as pointing to the screen, and soothing stimuli such as bubbles are often used in infants' studies to maintain their interest on the task at hand. Those were all avoided in the current study such that the investigation focused on infants' intrinsic motivation to engage. We recorded data from roughly the same age group as in the original study to avoid introducing more changes to the design. Moreover, in order to monitor their learning progress, infants need to go through enough trials that their performance can be

compared over time. 15-month-old infants appeared to be old enough to be able to go through numerous trials, understand the paradigm's rule and its contingency with their gaze location.

With this paradigm, we formulated two main hypotheses, illustrated in Figure 5.1:

1) Infants' decision to keep sampling information from the task or not i.e., their engagement with the task (number of completed trials), is proportional to their local learning progress (change in performance);
2) Infants' changes in performance over the course of the task follow a bell function, which we describe along three stages that we divide the task in: they first progress steeply and steadily at the beginning stage, after that their performance plateaus in a second stage, before dropping towards the end stage when the infants lose interest and disengage with the task.

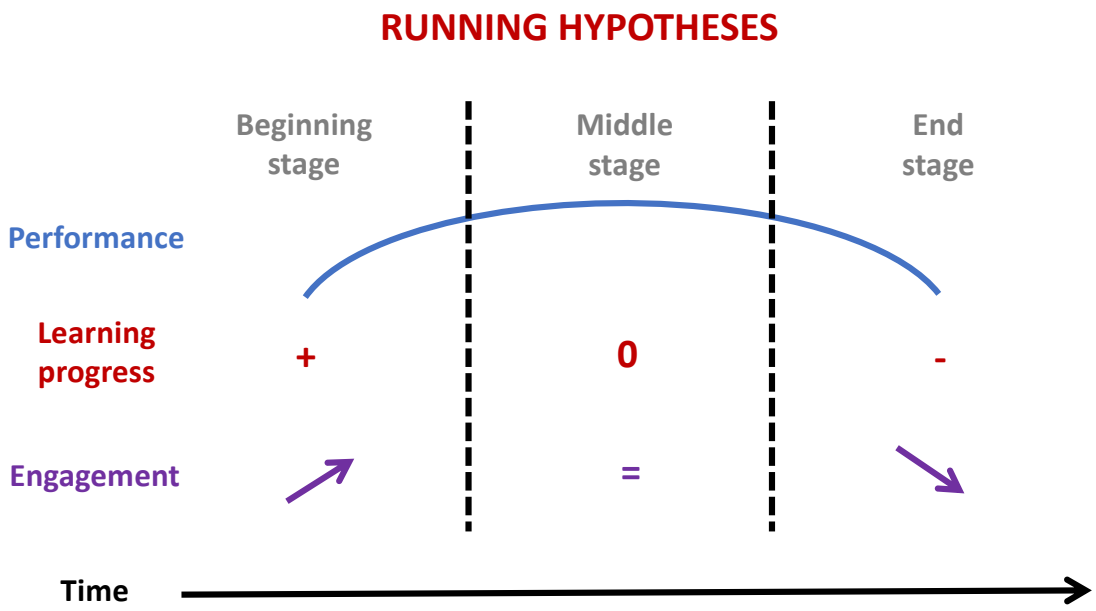


Figure 5.1: illustration of the study's hypotheses: over time, infants' engagement (in blue) is expected to change according to a bell curve, while learning progress (in red) follows the slope in performance and changes in engagement (in purple) follow the sign of the learning progress. Three stages in time are identified, following substantial changes in each variable.

5.2. Methods

This study did not rely on previous work from other people: the experiment's design, data collection (with the help of interns) and analysis were performed by me.

5.2.1. Participants

All the infants who participated in this study were recruited from a database of volunteers at the Centre for Brain and Cognitive Development at Birkbeck, University of London. Parents reported that they were born full-term (37-42 gestation weeks and over 2.5kg at birth), were typically developing and had no history of neurodevelopmental disorders in their nuclear family. Parents' written consent was obtained before any data was collected and the procedures were approved by the Ethics Committee of the Psychological Sciences Department at Birkbeck, University of London (approval no. 181903).

76 infants were tested with the final version of the study. They were 15-month-and-1-day old on average (S.D. 11 days, min. 14 months and 10 days, max. 15 months and 22 days). The study was composed of different tasks and infants sometimes only contributed data to part of the task set. This chapter focuses on the first task of the study, for which 7 participants did not contribute any data (1 lost file, 1 eye-tracking technical failure, 5 didn't cooperate for this task), a further 12 contributed too little clean data to be included in the analysis (less than 3 clean trials) and 16 were excluded for randomly triggering the reward too much or during the first trials – this was considered too confusing for them to grasp the gaze-contingent aspect of the task and its goal (see section 5.2.4.3 for criteria).

Ultimately, 41 infants (22 females) were included in the analysis. Some of the analyses required specific trials to be performed i.e., at least one trial when participants gazed at the target shape (triggered trials; N = 40 infants; 21 females), or three such trials (N = 26 infants; 14 females). Analysis the three stages required at least one clean trial per stage (N = 30 infants; 16 females) or at least one triggered trial per stage (N = 19 infants; 11 females). For all sub-samples of participants, infants were 15-month-and-0-day old on average (S.D. = 11 days). In addition, 25 infants were tested during the preliminary piloting stage of the study which served to refine the design of the tasks and the correct functioning of the equipment.

A sensitivity power analysis was conducted with the software G*Power 3.1(version 3.1, https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower), for each sub-sample and with standard values of significance threshold α = 0.05 and power 1-β = .80. The results of this analysis are summarised in Table 5.1 below and revealed that the study was partly sensitive to small effects for all types of tests and sub-samples (R > 0.3, F > 0.2 and t > 0.2).

Sub-sample	Test	Detectable effect size (critical threshold value)
N = 26	Correlation	R = +/- 0.39 (-)
N = 30	1-group*3-measure ANOVA	f = 0.24 (F = 3.16)
N = 19	1-group*3-measure ANOVA	f = 0.30 (F= 3.26)
N = 40 or 41	t-test comparison to zero	d = 0.45 (t = 2.02)

Table 5.1: Sensitivity power analysis for each sub-sample and each associated test.

5.2.2. Procedure for the study

Upon arrival, infants were given the chance to play with an experimenter and relax while another experimenter explained the study to the parent and answered their questions. After giving their consent to take part in the study, families were taken to the testing booths where infants performed a set of 4 short tasks: 1) First, they performed an eye-tracking, screen-based card-matching task, 2) after that they were taken to another room for a free play session at a table with a set of toys, 3) followed by an inhibitory control task, 4) and finally they went back to the eye-tracking booth to perform the screen-based memory task. The order of the tasks in the study was always the same. There was a planned break before the last task, when infants could stay on the floor of the play booth and play with bubbles and toys. Parents could request to stop the tasks or the study at any point, although the situation did not arise. If the infants became fussy, the experimenter stopped the task and the baby took a break before moving on to the next task. Finally, at the end of the session, the parents were given two questionnaires about their child to fill in at home and return by freepost: the Early Childhood Behavioural Questionnaire (ECBQ) and the Ages and Stages Questionnaire (ASQ). The last 3 tasks of the study will be the object of the

next chapter, in which the detailed procedures for those tasks can be found. The first task of the study (card-matching task) is the focus of the current chapter.

5.2.3. Procedure for the card-matching task

A Tobii TX300 screen-eye-tracker system was used to both follow participants' gaze throughout the task and present stimuli on the 1920x1080 pixels (pix.) display. The system was sampling the gaze at 120Hz and was controlled from a computer using PsychToolBox for Matlab (http://psychtoolbox.org/). Participants sat on their care-giver's lap about 60cm away from the Tobii screen and a 5-point calibration of the eye-tracker was presented, after which the task started. Each trial started with a medium grey screen shown for a randomly drawn duration within the 1000-1500ms interval, before a contingent red fixation spiral appeared on the grey background. The fixation spiral only disappeared after the participant's gaze was detected on it for a continuous 500ms.

The task was composed of 3 blocks: an object familiarisation block (2 trials), a task familiari-sation block (4 trials) and a test block (24 trials). A set of 72 stimuli was used: 12 shapes, each available in 6 colours. In both familiarisation blocks, a different set of 6 colours was used for the same 12 shapes to avoid carrying on effects of colour associations in the test trials.

In each of the two object familiarisation trials (see Figure 5.2), 6 of the 12 shapes were presented for 10s. They were positioned on a circle at a regular interval, with a grey background. Their locations on the circle and colours were pseudo-randomised such that each shape appeared once in the block while each colour was used once per trial.

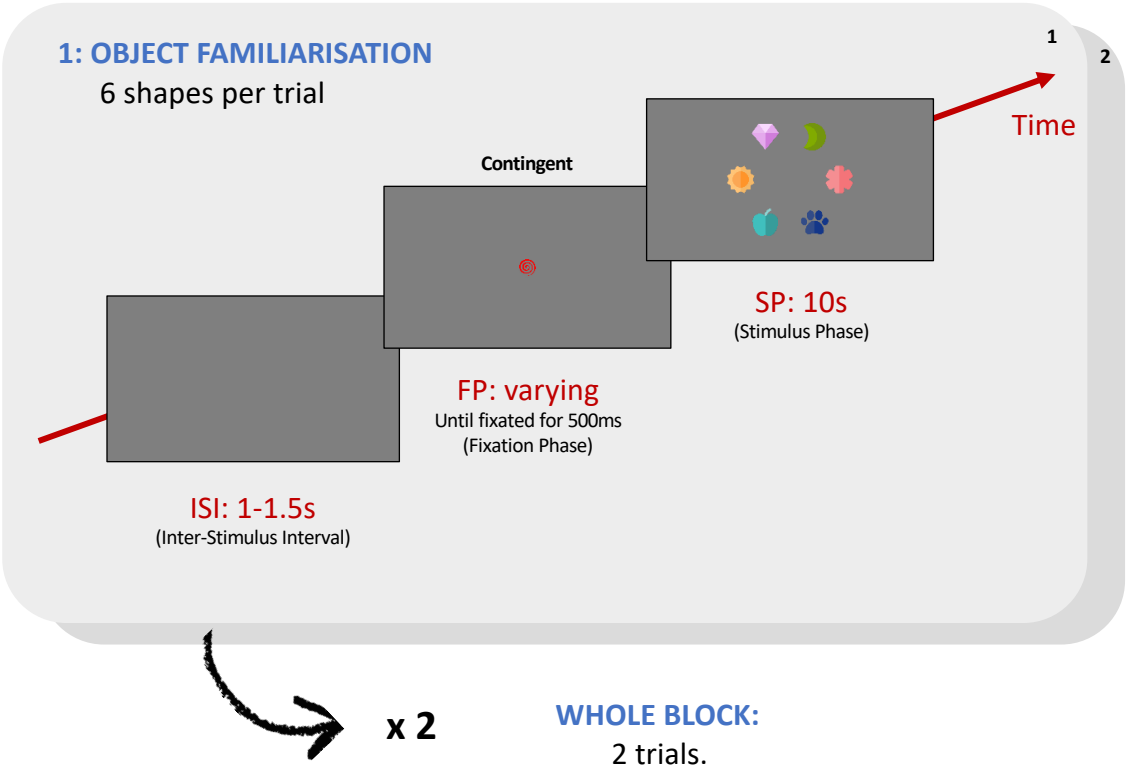


Figure 5.2: Structure of an object familiarisation trial; the trials were composed of an inter-stimulus interval (ISI) of 1-1.5s, a fixation phase (FP) of varying length contingent on the participant's gaze, and a stimulus phase (SP) of 10s. There were 2 such trials, each comprised of 6 shapes.

In the task familiarisation trials (see Figure 5.3), participants were presented with two cards on a beige background, one facing downwards and one facing upward. One of the shapes was drawn on the upward facing card. After 2.5s, the second card was flipped to reveal another shape of the same colour on its upward side. The second shape was either the same (2 matching trials) or different (2 non-matching trials) from the first shape. After 2.5s, an orienting outcome was presented. In the case of matching trials, a chime sound was played while the first card spined and loomed and a shiny purple explosion appeared on the card (rewarding). In the case of non-matching trials, a two-tones sound was played with no visual animation (contrasting unrewarding outcome). The trials' order was randomised and the location of the first card was pseudo-randomised such that each was location was used once in each matching trial and once in each non-matching trial. These trials were used to familiarise the infants with the task and the shape-matching rule.

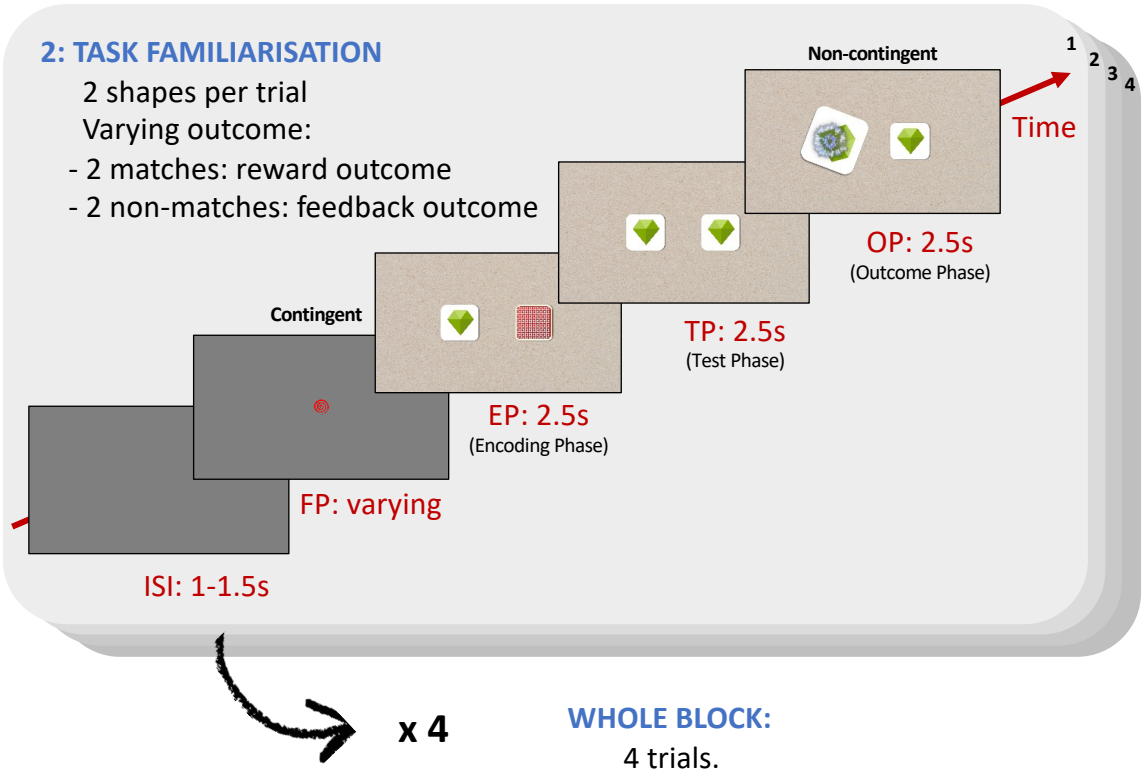


Figure 5.3: Structure of a task familiarisation trial. The trials comprised an inter-stimulus interval (ISI) of 1-1.5s, a fixation phase (FP) of varying length contingent on the participant's gaze, an encoding phase (EP) of 2.5s when one shape was visible, a test phase (TP) of 2.5s when both shapes were visible and an outcome phase (OP) of 2.5s which consisted of a rewarding outcome if the shapes matched or a feedback outcome if they did not, independent of the participant's gaze. There were 4 such trials, 2 matches and 2 non-matches, each of them containing 2 shapes.

In the test trials (Figure 5.4), participants were presented with a central card facing downwards and three peripheral cards, all at an equal distance of the central one. The cards were presented on a beige background and were all facing downwards apart from one of peripheral cards, on which one of the shapes was drawn. Every 2s, one of the hidden cards was flipped and its shape revealed while a flipping sound was played. The central card was shown last and loomed for 1.5s when flipped to drive the participants' attention to it. This screen with all the shapes lasted for an extra 3.5s (5s total) maximum, after which the trial outcome was presented. From the moment that the central shape was shown, if the participant's gaze was caught in an area of 1.5 times the

size of the peripheral matching card, a rewarding outcome was triggered, where the matching card loomed and spun with a shiny purple animation on it and a chime sound was played (as for the task familiarisation block). If the participant's gaze was not caught in the area-of-interest (AOI) for the 5s when the cards were facing upward, a two-tone sound was played (as for the task familiarisation block) and the matching card loomed for 2.5s in lieu of an orientating but non-rewarding outcome.

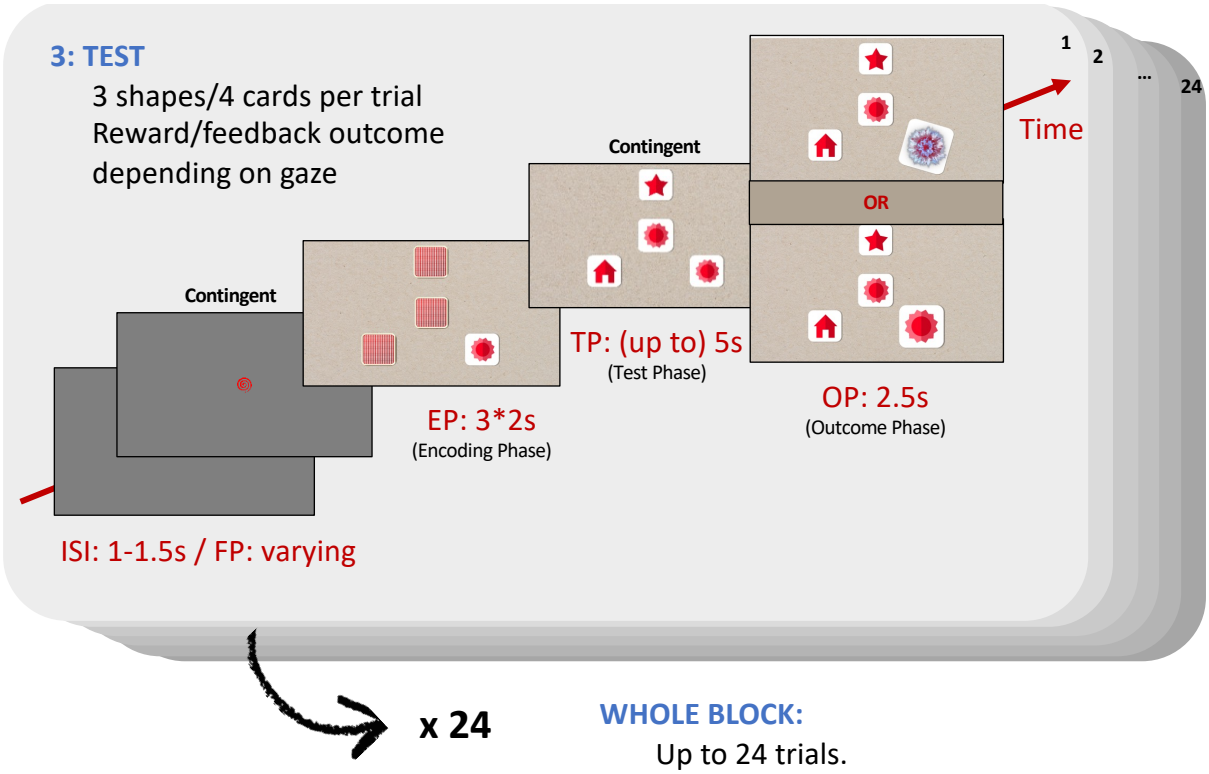


Figure 5.4: Structure of a test trial; the trials were composed of an inter-stimulus interval (ISI) of 1-1.5s, a fixation phase (FP) of varying length contingent on the participant's gaze, an encoding phase (EP) of 6s, 2s for each of the 3 peripheral cards which were flipped one after the other, a test phase (TP) of up to 5s when all four shapes were visible and during which the participant's gaze was tracked in order to terminate the TP and trigger the reward in case their gaze was caught in the target peripheral card for at least 100ms, and an outcome phase (OP) of 2.5s which consisted of a rewarding outcome if the participant's gaze triggered the reward, or a feedback outcome if it did not. There were up to 24 such

trials depending on participant's interest, and each of them comprised 3 shapes on 4 cards.

All the shapes were obtained from a royalty-free online database (flaticon.com). They had the same colour within one test trial, and the test trials colour was randomised such that each colour was used once per set of 6 trials (4 times for the whole 24 test trials). All the shapes were randomised such that each shape was presented on a peripheral card twice every 8 trials (6 times for the whole 24 test trials), and such that they were drawn on the central card once every set of 12 trials. The location of the matching shape was randomised such that each location held the matching shape twice every set of 6 trials (8 times for the whole 24 test trials). The position in time in which each shape was revealed was also pseudo-randomised such that each location was revealed at each position in time twice per set of trials (8 times for the whole 24 test trials) and that the matching shape was revealed at each position in time twice per set of 6 trials (8 times for the whole 24 test trials). At the end of the task, infants were presented with a 9-s rewarding clip of animals clapping, taken from a scene of Classical Baby (HBO).

5.2.4. Pre-processing of the data

The pre-processing pipeline, summarised in Figure 5.5, was performed participant by participant in 4 main stages on 1) the continuous dataset, 2) each trial, 3) each trial's test phase, 4) at the participant level.

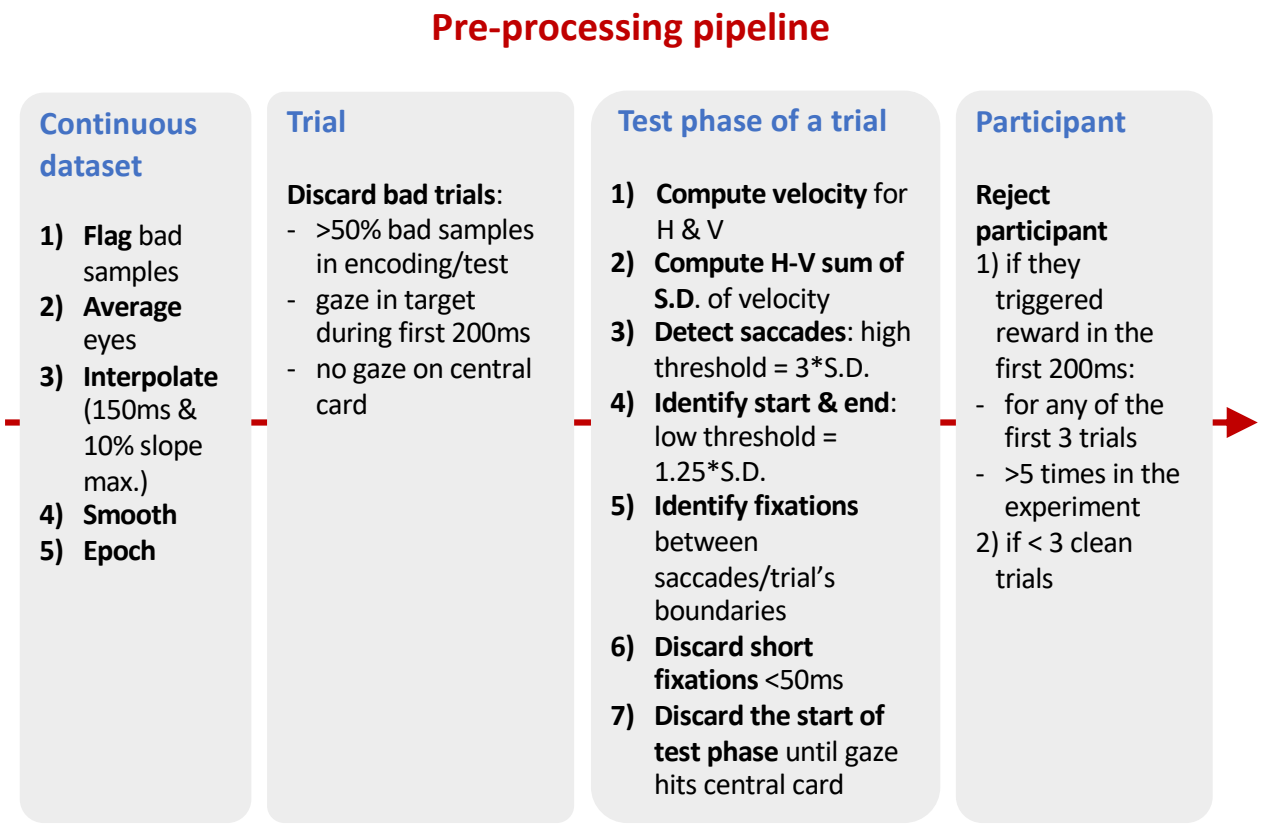


Figure 5.5: Summary of the pre-processing pipeline. It was performed participant by participant, in 4 main stages on 1) the continuous dataset, 2) each trial, 3) each trial's test phase (horizontal H and vertical V coordinates, considered separately until now, grouped in step 2), 4) at the whole participant level. Each stage comprised different steps summarised here and developed in-text.

5.2.4.1. Stages 1 and 2: processing of the continuous dataset and trials

Horizontal (H) and vertical (V) coordinates of the gaze location were obtained from the recordings and processed as two separate data streams. First, they were processed continuously: for each stream of data (H and V, each for the left and right eye), samples for which the gaze was not found by the eye-tracker or caught outside of the screen were flagged as bad. Then, when data was available for both eyes, an average was calculated – otherwise only one eye was used. Next, the missing data was interpolated using a linear method, for segments up to 150ms long with up to 10% slope between the start and the end of the missing portion. Finally, the data was smoothed with a 5-point average and epoched into trials. At this point familiarisation trials were excluded from further analysis and only test trials were included in the following steps. Test trials were

each divided into 5 parts: the inter-stimulus, the fixation, the encoding, the test and the outcome phase (see Figure 5.6A). The analyses were performed on the test part, but the encoding part was also considered for trial rejection. At this point, trials were rejected if: 1) over 50% of the samples were flagged as bad in either the encoding or test phase, 2) the gaze was caught in the target during the first 200ms of the test phase, which was considered too early to be voluntary, 3) there was no gazing of at least 50ms to the central card to be matched. Moreover, within good trials, any data until the gaze was found in the central card for the first time was discarded.

5.2.4.2. Stage 3: identifying saccades and fixations over trials' test phase

Saccades and fixations were identified using a simplified version of Nyström and Holmqvist's algorithm (Nyström & Holmqvist, 2010) as follows. The velocity (first derivative) of the horizontal (H) and vertical (V) locations was computed and summed into one composite H-V value. The standard deviation of the trial's H-V velocity was then calculated to serve as a threshold: a high threshold of 3 times the S.D. was used to detect saccades, and a lower threshold of 1.25 times the S.D. was used to identify the start and the end of the saccades around the previously detected location. Fixations were identified as segments of at least 50ms bound by saccades or trial boundaries. An example plot of the data from one participant is provided for reference (Figure 5.6). Fixations were also identified as within an AOI of interest (four regions considered, around an area of 1.5 times each of the four cards) or not.

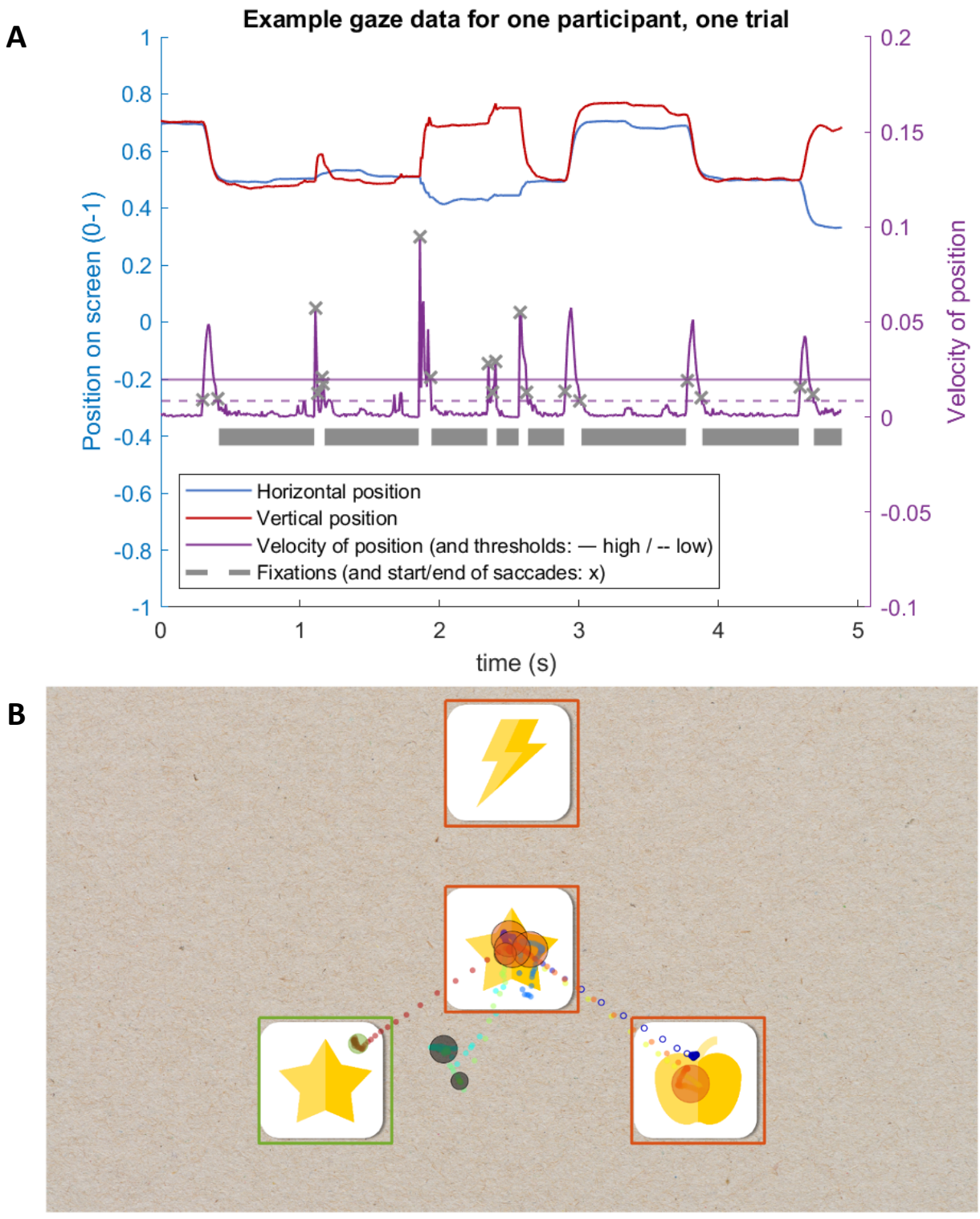


Figure 5.6: example data for one participant's trial.

A. The x-axis represents the time course of the trial, the left y-axis represents the gaze position as a proportion of the screen size (0-1) horizontally (H, in blue) and vertically (V, in purple). The H-V composite S.D. of the velocity is represented along the right y-axis in

orange together with its threshold: low (dotted line) and high (continuous line). The fixations identified from those thresholds are represented as green boxes and the saccades (and fixations) start & end are represented as green crosses.

B. Gaze samples are represented as small circles over the screen seen by the participant during the test phase. Their colour follows the time course of the trial: the first samples are represented in blue, through green, yellow, orange, and red for the last samples. Samples that happened before the central card was hit, and were thus discarded from analyses, are still depicted here as empty circles. Larger circles represent fixations to the target (in green), other cards (in orange) or the background (in black). Fixations to the central card are discounted when analysing participants' fixations to any of the matching options (i.e., the peripheral cards); thus, central card fixations are depicted differently with black outline. The diameter of each fixation's circle is proportional to its duration. Finally, each area of interest around the cards is represented as a square in green (target) or orange (non-target).

5.2.4.3. Stage 4: assessment of the data per participant

Finally, 16 participants were excluded because they triggered the reward during the first 200ms of a trial, either when the trial was one of the first three trials (12 participants excluded) or when this happened for more than 5 trials in the experiment (4 participants excluded). This was considered too confusing for them to grasp the gaze-contingent aspect of the task and its goal. Another 12 participants had less than 3 clean trials after pre-processing and were also excluded.

5.2.5. Analysis of the data

5.2.5.1. Variables extracted

The analyses were performed on the test phase of the trials, discarding any data until the participants looked at the central card. Two types of variables were extracted from the participants' gaze data: binary scores and continuous measures. Two binary scores were computed: *first AOI hit is target*, the percentage of trials for which the first AOI hit was the target AOI, and *gaze hits target*, the percentage of trials for which the target AOI was hit at all. Four continuous average measures were calculated: *time to hit target* in ms, *number of fixations to hit target*, *distance to hit target* in pixels (pix.) and *speed to hit target* in pix./ms The distance was

calculated over the whole path by summing the distance between each included fixation. The speed was calculated as the ratio between distance and time. Binary measures were computed on all included trials while continuous variables were computed on the subset of included trials for which the participants did look at the target (triggered trials). All variables represent a way to measure either how much or how fast the participants were gazing at the target. They were evaluated together to offer a detailed account of infants' eye-movement control during the task, which is argued to help to disambiguate their behaviour (Aslin, 2007).

5.2.5.2. Definition of different windows interest

We took three approaches to investigate these measures: 1) over the whole task, 2) dividing the task into 3 stages, and 3) focussing on only the start (first 3 triggered trials). First, we looked at the key measures over the whole task to assess participants' overall performance. Second, we performed a stage-based analysis to test hypothesis 2 that individual performance varies along the task according to a bell curve – we expected an increase in performance from stage 1 (steep progress stage) to 2 (plateau stage) and a decrease from stage 2 (plateau stage) to 3 (drop and disengagement stage). Similarly to the approach taken in chapters 3 and 4, the stages were defined for each subject: a participant's number of completed trials was divided into 3 stages of equal length – the start, middle and end stages. As for the former chapters, the experiment was designed to allow for participants to complete a different number of trials based on their intrinsic motivation to engage, without interventions from the experimenters or the parent to keep them on the task. Thus, participants completed different numbers of trials. This division into stages on a subject-to-subject basis was helpful to investigate changes over the course of the task in such a group of participants with different trial numbers. Finally, we focused on the begin of the task, when we predict the steepest progress and where the most individual differences in progress can be expected. In this window of interest, we tested hypothesis 1 that progress (change in performance) drives engagement (trials completed). We looked at progress (slope in performance) along the first 3 triggered trials only rather than over the whole of stage 1, such that participants with different trial numbers per stage could still be compared. Indeed, trial number can affect the precision of the slope calculation and needs to be controlled for in this analysis. Slopes were computed by linearly fitting the 3 performance values against their trial

number, using Matlab's polynomial fitting function *polyfit* with a degree of 1. Statistics on any of these variables were performed using the software JASP, version 0.15 (Love et al., 2019).

Participants completed 16.29 trials (S.D. = 5.56) on average, of which an average 9.49 trials were included (S.D. = 5.28) and 8.67 were excluded (S.D. = 5.46).

5.3. Results

Our aim with this study was to investigate the role of infants' learning progress (changes in performance), on their engagement (number of completed trials). First, we looked at the performance of the group over the whole task compared to chance-level performance. Next, we looked at stage-by-stage changes in performance measures to test hypothesis 2 that these changes follow a bell function. Finally, infants' individual learning progress at the start of the task (change in performance over the first 3 triggered trials) was evaluated in relation to their engagement (number of completed trials), in order to test hypothesis 1 that progress and engagement are linked.

5.3.1. Whole-task overall performance

5.3.1.1. Proportion of trials for which the first AOI hit is target

Participants' performance was first measured based on whether they looked at the target AOI, i.e., the matching peripheral card, directly after having looked at the central card. The proportion of trials for which this was the case yielded an individual overall whole-task score, *first AOI hit is target*. Because each trial was composed of three peripheral options, whose locations, shapes and colours were carefully pseudo-randomised across the experiment (see Methods section 5.2.3), the chance level was taken as 1/3 of the trials (33.33%). Overall, participants seemed to perform at chance (see Figure 5.7), as shown by a one-sample t-test performed against the chance value: t(40) = -0.205; p = 0.839; mean = 32.4% (S.D. = 17.5%).

This shown that the group of 15-month-old infants did not seem to be able to perform this task as was anticipated. One possible explanation is that the task was simply too difficult for the infants to perform it correctly. However, it is important to note that if the changes in performance follow the bell curve predicted by hypothesis 1, then they can be expected to average out to the baseline performance level. It could be the case that infants' baseline performance is at chance here, which averages out over the whole task despite above-chance performance at peak time. The stage analysis in section 5.3.2 offers a possibility to further investigate this option while testing hypothesis 2. Finally, it could also be possible that this measure of performance, although it

seemed straightforward to use and interpret, is not the most suited for this task and age group. To look into this, alternative measures of performance were investigated.

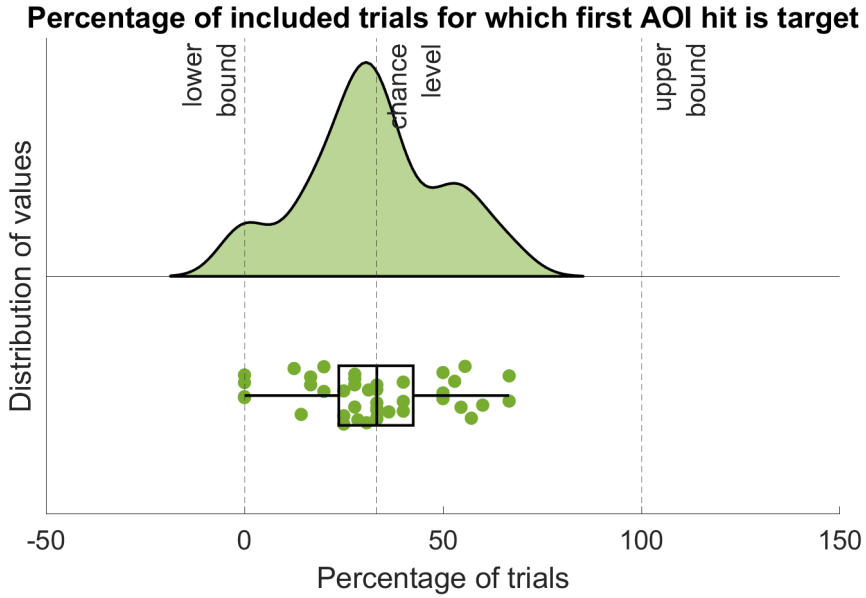


Figure 5.7: Raincloud plot of the proportion of trials for which the target AOI was the first AOI participants looked at after looking at the central card.

Each dot (rain) represents a participant's data point. Their average score over the whole task is plotted along the x-axis, and y-axis positions are randomised to avoid overlapping points. The group's values are depicted as a distribution curve above the points (cloud) as well as a boxplot, in which the box signifies the 25% to 75% percentiles, the line inside the box represents the median, and the whiskers depict the range, discarding any potential outlier (points outside of 1.5 times the interquartile range; non identified for this measure). Vertical dotted lines are drawn at the lower and upper bound values that constrain the measure, and at the 33.33% chance level.

5.3.1.2. Proportion of trials for which gaze hits target

Before looking into other measures of performance, we tested whether participants looked at the target ("triggered" trials) and whether this task yielded a good enough number of triggered trials to further investigate participants' changes in performance. Again, a score was computed over the whole number of included trials: *gaze hits target*, the percentage of trials for which the target AOI was hit (Figure 5.8). Overall, participants seemed to gaze at the target AOI for a majority of the included trials (58.6 % of trials – or 5.56 trials – on average, S.D. = 21.5 %),

indicating that there was a good amount of data to further investigate infants' behaviour. Because participants were provided with a large enough time window to look at several options before the end of the trial, "triggered trials" like these, when participants merely gazed at the matching option, cannot be considered "correct trials", and cannot directly indicate whether infants were able to perform this task better than a pre-determined chance level.

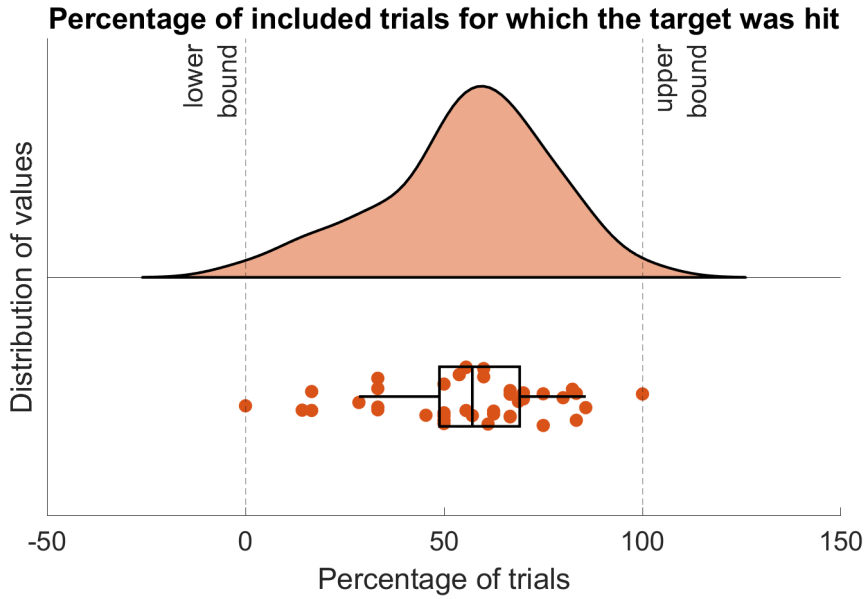


Figure 5.8: Raincloud plot of the proportion of trials for which participants looked at the target AOI. (See Figure 5.7 for a detailed description of the elements of the plot.)

5.3.1.3. Continuous variables: time, number of fixations, distance, and speed to hit target

Four other eye-movement-control variables were used to assess infants' performance, all reflecting how long it took them to gaze at the correct, target option. We looked at *time* (ms), *number of fixations*, *distance* (pix.), and *speed* (ms/pix.) to hit the target (see Appendix 5 for raincloud plots). *Time*, *number of fixations* and *distance* to hit target represent measures of how much of the space is visually explored before reaching the target, while speed reflects how fast this space is explored. Because these variables are continuous and are not scores that can be compared to an absolute chance value, no conclusion on infants' level of performance can be drawn from looking them over the whole task. However, looking at how they change across stages can inform on whether infants learn across time.

5.3.2. Testing hypothesis 2: changes in performance (learning progress) over the three stages

Our next step was to investigate whether infants adapted their performance throughout the task by looking at stage-by-stage changes in the measures introduced above. Stages were defined using the same approach as in Chapters 3 and 4: each infant's total number of completed trials, which varied between individuals, was divided into three stages of equal length for that given infant. This approach has the benefit of enabling the analysis of participants with different trial numbers. Indeed, with the present experimental design, infants are left free to drop out from the task at any time, without interventions from the experimenters or the parent to keep them on the task, resulting in variable numbers of completed trials by design. All continuous variables were standardised by z-scoring participants' stage values to the mean and standard deviation of the whole-task values, which accounted for different ranges and baseline levels between individuals. Stage-by-stage changes in performance were expected to follow hypothesis 1's bell-shaped pattern. We predicted an increase between stages 1 and 2 followed by a decrease between stages 2 and 3 in participants' *first AOI hit is target*, *gaze hits target* or *speed to hit target*, all measures of performance directly proportional to it. For *time, number of fixations* and *distance to hit target*, the opposite pattern was expected because these measures of performance are inversely proportional to it.

These six measures were investigated for an effect of stage (see Figure 5.9) in six repeated measures (3*1) ANOVAs reported in Table 5.2. Descriptive means and standard deviations for each variable and stage can be consulted in Appendix 6. We found an effect of stage on *speed to hit target* ($F(2, 36) = 4.284$, $p = 0.021$, $\eta^2 = 0.192$) in the direction of an overall stage-by-stage decrease, but not on any of the other five measures. This was not unexpected, as we predicted changes along a bell curve which could lead to no overall linear effect. Therefore, planned contrasts were also performed between stages 1 and 2, and 2 and 3, for each variable. These tests are reported in Table 5.3 and revealed a significant decrease between stages 2 and 3 for the *speed* variable only.

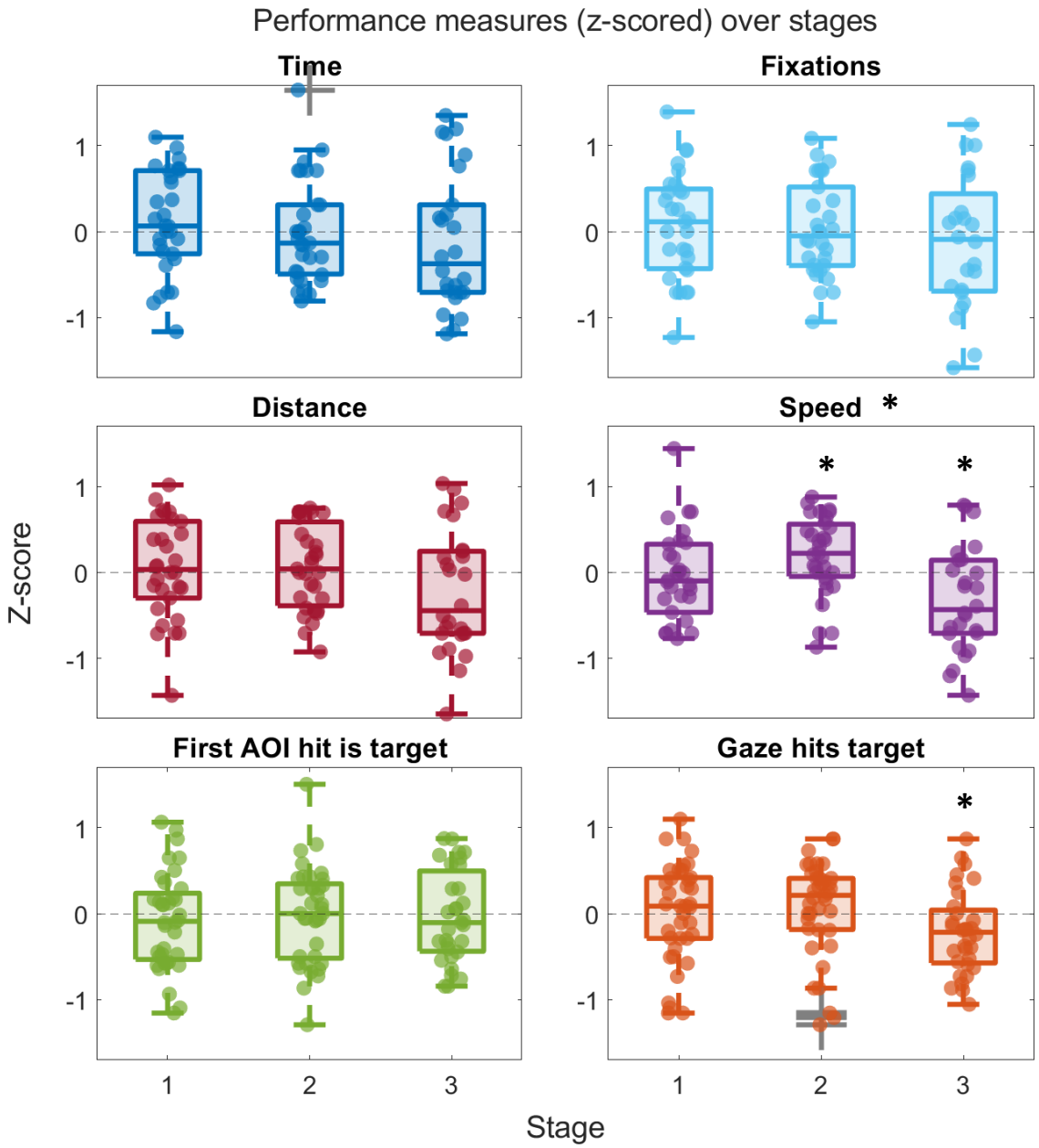


Figure 5.9: Boxplots of the standardised (z-scored) performance based on the six different variables introduced in section 5.3.1, over the three stages of the experiment. Each variable is depicted at a specific row and column, within which each stage is plotted as a different boxplot.

Similarly to the boxplot element in the former raincloud plots, participants' scores are represented as dots, the group's 25 to 75 percentile is depicted as a box within which a line represents the median, and the whiskers depict the range, discarding any potential outlier (points outside of 1.5 times the interquartile range; represented as grey crosses when present). A dotted line is drawn at z-score = 0 to represent the baseline level across

the task. Stars next to variable names represent a significant main effect, while stars over a stage's boxplot indicate a statistical difference to baseline (zero).

Variable	**Statistical test (F)**	**Statistical significance (p)**	**Effect size (η^2)**
Time to hit target	F(2, 36) = 0.402	0.672	0.022
Number of fixations to hit target	F(2, 36) = 0.707	0.500	0.038
Distance to hit target	F(2, 36) = 1.365	0.268	0.070
Speed to hit target	F(2, 36) = 4.372	**0.020**	0.195
First AOI hit is target	F(2, 60) = 1.996	0.145	0.062
Target was hit	F(2, 60) = 0.007	0.993	0.0002

Table 5.2: Results of the 3*1 repeated measures ANOVAs on each performance variable; each variable is reported in a different row, and each statistical value in a different column. Note that the candidate variables reported in the first four rows rely on more stringent conditions (at least 1 triggered trial in each stage) than the scores reported in the last two rows (at least 1 trial in each stage) and are thus computed on different numbers of participants (respectively, 19 and 30 participants). Significant results are highlighted in grey cells and bold font. Means and standard deviations for each variable and stage can be consulted in Appendix 6.

Variable	Stages compared	Estimate (Standard Error, S.E.)	Statistical test (t)	Statistical significance (p)
Time to hit target	1-2	0.109 (0.241) ms	t(1, 18) = 0.453	0.656
	2-3	0.126 (0.247) ms	t(1, 18) = 0.513	0.614
Number of fixations to hit target	1-2	0.039 (0.248)	t(1, 18) = 0.158	0.876
	2-3	0.243 (0.241)	t(1, 18) = 1.009	0.327
Distance to hit target	1-2	-0.012 (0.213) pix.	t(1, 18) = -0.055	0.957
	2-3	0.321 (0.221) pix.	t(1, 18) = 1.451	0.164
Speed to hit target	1-2	-0.118 (0.172) pix./ms	t(1, 18) = -0.682	0.504
	2-3	0.439 (0.201) pix./ms	t(1, 18) = 2.187	**0.042**
First AOI hit is target	1-2	0.019 (0.178)	t(1, 29) = 0.106	0.916
	2-3	-0.013 (0.156)	t(1, 29) = -0.084	0.934
Target was hit	1-2	0.097 (0.176)	t(1, 30) = 0.552	0.585
	2-3	0.212 (0.147)	t(1, 30) = 1.444	0.159

Table 5.3: planned contrasts (paired samples t-tests) for each performance variable between stages 1 & 2, and 2 & 3; each variable is reported in a different row, and each statistical value in a different column. Significant results are highlighted in grey cells and bold font.

Finally, each stage's marginal mean value was compared to zero to further investigate stage-by-stage changes in terms of deviations from baseline performance (z-score = 0). This analysis revealed that infants started at a baseline *speed*, which then significantly rose above baseline at stage 2 to then significantly drop below baseline in the last stage (see Table 5.4). Other measures did not appear to significantly differ from their baseline level at any stage, except for the measure of the *gaze hits target*, which significantly dropped below baseline in the last stage (see Table 5.4 for *gaze hits target* and Appendix 7 for other variables).

Variable	Stage	Mean	Standard Error (S.E.)	Statistical test (t)	Statistical significance (p)
Speed to hit target	1	-0.071	0.126	t(40.926) = -0.562	0.577
	2	0.259	0.126	t(40. 926) = 2.039	**0.048**
	3	-0.369	0.126	t(40. 926) = -2.906	**0.006**
Proportion of trials for which target was hit	1	0.094	0.095	t(68.027) = 0.996	0.323
	2	-0.003	0.095	t(68.027) = -0.030	0.976
	3	-0.215	0.095	t(68.027) = -2.275	**0.026**

Table 5.4: Statistical significance of the marginal means at each stage compared to zero. Only two variables came out significantly different to zero in some stages: *speed to hit target* and *gaze hits target*; other variables were all above the significance p = 0.05 and the trend p = 0.10 thresholds and are not reported in this table to avoid crowding but can be consulted in Appendix 7. Significant results are highlighted in grey cells and bold font.

These results suggest that infants did adapt their performance on this task over time in terms of how fast they were scanning the space i.e., in their *speed to hit target*. Their change in *speed* performance followed a bell curve, as predicted by hypothesis 2: infants got better from the first to the middle stage of the experiment, before dropping in performance at the end of the session. This final drop in performance was also reflected in a drop below baseline in *gaze hits target*, partially supporting hypothesis 2 for this measure too. Interestingly, neither of the other measures of performance showed a significant stage-by-stage change.

The way the stages were defined on a subject-by-subject basis, with varying lengths depending on the individual, is compatible with hypothesis 1: they scale with infants' interest, which is itself expected to be driven by bell-shaped changes in learning progress or performance. The fact that typical bell-shaped changes could be found using this subjective stage division leans toward hypothesis 1, but more direct evidence is needed.

5.3.3. Learning progress: changes in performance over the first 3 trials and relationship with engagement

In a last set of analyses to directly test hypothesis 1, we looked at whether learning progress at the very beginning of the task could predict for how long an infant would engage with the task. The beginning of the task represented the time window when the steepest progress was expected, making it a key time window for the investigation of changes in performance. We computed the slope in performance using linear fitting, over the first three trials in which the target was hit, for the four continuous variables. Binary scores were not included here because an oscillation between two values over a small number of trials cannot yield meaningful slope values. This analysis was performed on the first 3 triggered trials only rather than over the whole stage 1, because trial number can affect the precision of the slope calculation and using a subjective stage rather than an objective amount of trial would bias the results for this analysis.

5.3.3.1. Whole-group learning progress

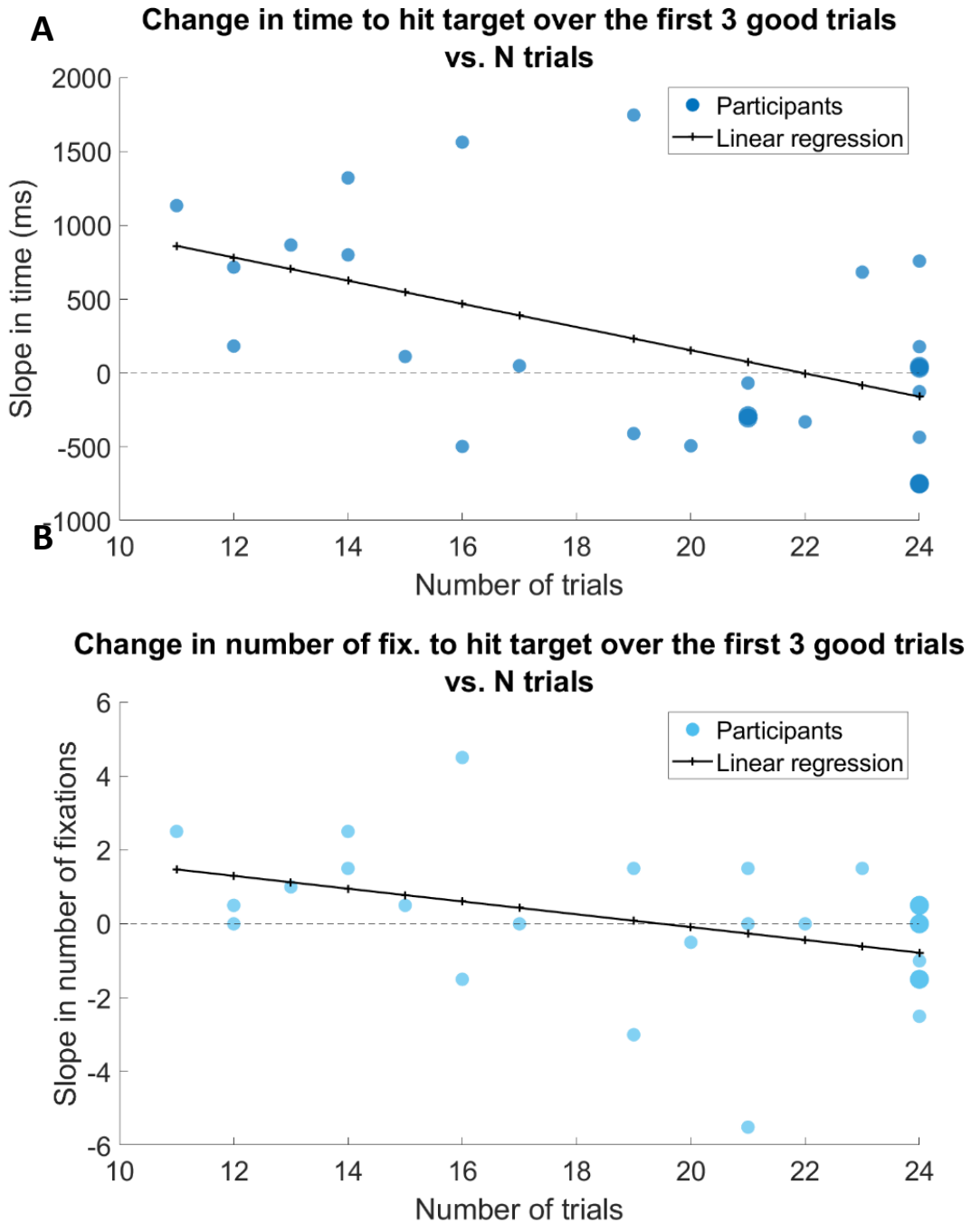
First, we compared the slope values of the participants to zero (no learning) through four one-sample t-tests, which all came out non-significant (Table 5.5). This did not bring substantial evidence that on average, participants were making progress in this task over the first trials.

Variable	**Mean (S.E.)**	**Statistical test (t)**	**Statistical significance (p)**	**Effect size (Cohen's d)**
Time to hit target	221.292 (708.161)	t(25) = 1.593	0.124	0.312
Number of fixations to hit target	0.058 (1.971)	t(25) = 0.149	0.883	0.029
Distance to hit target	126.371 (613.282)	t(25) = 1.051	0.303	0.206
Speed to hit target	0.011 (0.130)	t(25) = 0.426	0.674	0.084

Table 5.5: One-sample t-tests on the slope over the first three trials in the four continuous measures of performance.

5.3.3.2. *Individuals' learning progress related to engagement*

Next, we asked whether how long participants stayed on the task was related to how much progress they made, by looking at correlations between the number of completed trials and the four slopes at start (Figure 5.10). In Chapters 3 and 4, time-on-task was used as a measure of engagement. However, here, the use of a gaze-contingent design through which participants could trigger rewards and shorten the trials' length made the time-on-task variable performance-dependant, thus we used the number of completed trials as a measure of engagement here.



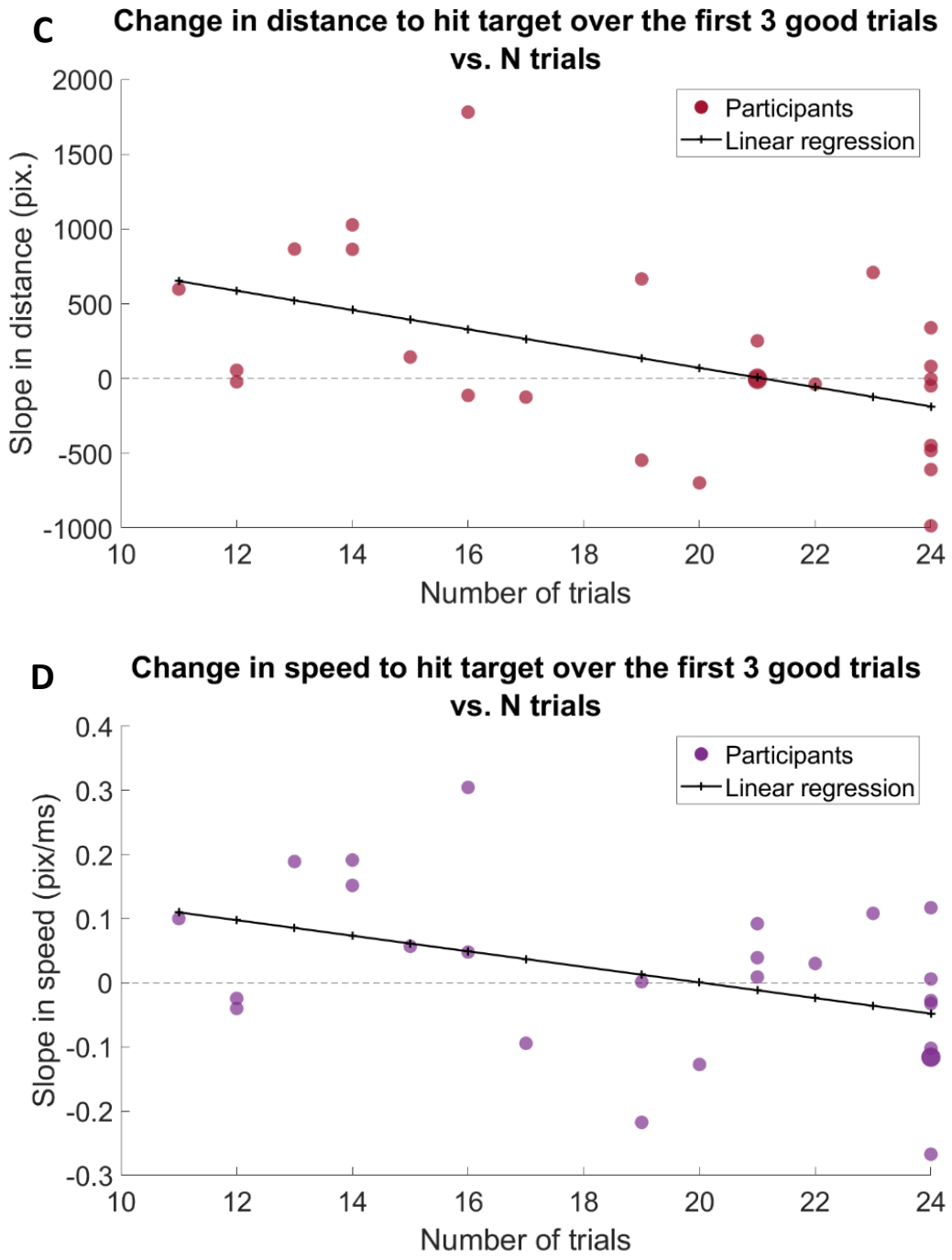


Figure 5.10: Relationship between the slope in performance measured by the four continuous variables introduced in section 5.3.1 and the participants' number of completed trials. Each point represents the slope value for one participant; marker diameter is proportional to the number of individual values represented by the same point in case of overlap. The linear regression between both values is depicted as a black line on which crosses are drawn to figure the points' projections on the regression line. A dotted line is drawn at 0 (no change).

A. Time (ms) to hit the target AOI.

B. Number of fixations made to hit the target AOI.

C. Total distance (pix.) travelled by gaze from fixation to fixation to hit the target AOI.

D. Speed (distance divided by time, pix./ms) to hit the target AOI.

We found strong negative correlations between the slopes and the number of completed trials for all four measures (Table 5.6): negative slopes were associated with more completed trials. The first three measures reflected how much of the space participants explored before looking at the target, so their decrease over the first trials (negative slope) indicated less exploration of alternatives i.e., increased performance and learning progress for reaching the target. However, the *speed* variable was a measure of how fast the space was explored before reaching the target: here negative slopes reflected a slowdown of exploration i.e., decreasing performance, still associated with more completed trials. This last result was unexpected, further comments are made in this Chapter's Discussion section 5.4.

Because the slopes are dependent on the absolute performance value achieved at the last of these 3 trials, a post-hoc analysis was performed, computing these correlations again while controlling for the performance value at 3rd trial. No correction was applied as these tests were not performed to reveal any new effect but to rule-out the implication of a covariate in already evidenced effects. Importantly, the relationship between number of trials and performance slopes did not disappear when controlling for the third trial's value, at the exception of the *speed* and *number of fixations* variables, which however maintained a trending level of significance (respectively, $p = 0.060$ and $p = 0.083$). This indicates that how much infants were willing to stay on this task did not depend on the absolute performance achieved, but rather on their progress itself, which supports the learning-progress hypothesis (hypothesis 1).

Variable	Controlling for	Correlation (Pearson's R)	Statistical significance (p)
Time to hit target	**No variable**	-0.507	**0.008**
	Value at 3rd trial	-0.542	**0.005**
Number of fixations to hit target	**No variable**	-0.402	**0.042**
	Value at 3rd trial	-0.353	**0.083**
Distance to hit target	**No variable**	-0.481	**0.013**
	Value at 3rd trial	-0.451	**0.024**
Speed to hit target	**No variable**	-0.428	**0.029**
	Value at 3rd trial	-0.381	**0.060**

Table 5.6: Correlations between the number of trials completed by the participant and their slope in performance, as measured by the four continuous performance variables. Correlations were computed with or without controlling for the performance value at the 3rd trial. Results under the significance threshold ($p < 0.05$) are highlighted in grey cells and results under the trending threshold ($p < 0.10$) in bold font.

All in all, this last set of analyses revealed that on average the group did not seem to be improving performance over their first trials (section 5.3.3.1), but there was a heterogeneity of behaviours, with some infants showing learning progress and others not. Investigating individual differences in performance change in relation to engagement revealed that the participants who showed more progress also completed more trials, which in in support of hypothesis 1. It is worth noting though that all the infants who completed few trials also had a negative or close-to zero slope in each of the measures tested (Figure 5.10, points at the left side of the plots). However, there was more heterogeneity for the infants who completed more trials (Figure 5.10, points at the right side of the plots): they mainly had negative or null slopes (Figure 5.10, points at the bottom right corner of the plots) but some of them still exhibited positive performance slopes (Figure 5.10, points at the top right corner of the plots). Thus, learning progress at the start of the experiment appeared sufficient to maintain engagement, as proposed in hypothesis 1, but not necessary.

5.4. Discussion

5.4.1. Summary of findings

We used a new paradigm adapted from Hochman et al. (2016) and Kaldy et al. (2015) to track local learning progress in 15-month-old infants on a rule-learning task. Our results bring the first direct piece of evidence that infants are able to monitor their learning progress and adapt their engagement with information accordingly.

5.4.1.1. Evidence for hypothesis 2: performance changes over time following a bell curve

We first tested hypothesis 2, that changes in infants' performance over time follow a bell curve. We found evidence for this in the *speed* eye-movement-control variable, which significantly increased between the beginning and the middle stage, before decreasing from the middle to the end stage. The *gaze hits target* variable partly followed this expected shape, with a significant drop below its baseline level during the end stage, while the start and middle stage value did not significantly differ from baseline. We found no significant stage-by-stage change in the other measures: *first AOI hit is target*, *time*, *number of fixations* or *distance to hit target*. These results support hypothesis 1, but only when looking at performance in terms of how fast participants explored the space (*speed*) or how much they were able to reach the correct answer (*gaze hits target*), and not in terms of how much they explored (*first AOI hit is target*, *time*, *number of fixations*, *distance to hit target*) before reaching the target.

5.4.1.2. Evidence for hypothesis 1: learning progress influences engagement

We also tested hypothesis 1, that infants' willingness to engage with a task depends on their learning progress i.e., change in performance. We found that infants' performance change at the start of the task, as measured by either *time*, *number of fixations*, *distance*, or *speed to hit target*, was negatively correlated with their *number of completed trials*. This indicated that infants who got better over the first three triggered trials and explored less and less of the space before reaching the target, also tended to engage for longer with the task. There was an exception to this relationship with the *speed* variable, for which a negative correlation was also found, but this time it reflected a link between decreased performance (slower exploration) and more

engagement with the task. The *speed* variable thus appeared to represent a possibly more ambiguous measure of performance compared to the other variables, which is further discussed in section 5.4.3. Hypothesis 1 stems from the learning-progress theory which, to our knowledge, had not been directly tested in infants yet. Experiments with robots programmed to learn by maximising learning-progress showed that such algorithms generated infant-like explorative behaviour in robots placed on a toy playground, suggesting that it might similarly be valid for infants. We brought the first direct piece of evidence for this theory in 15-month-old infants performing our rule-learning task.

5.4.1.3. Learning progress could be sufficient but not necessary to drive engagement

Interestingly, while our results support the learning-progress theory in infants, they show that learning progress is sufficient to drive engagement, but does not appear necessary. Indeed, infants who completed few trials all showed little or no progress (positive or close-to-zero slopes), but infants who completed more trials did not all show important progress (heterogeneity: a majority of negative slopes mixed with positive or close-to-zero values). This leaves room for other drivers of infants' engagement to come into play when individuals remain engaged despite an unfavourable context given their learning progress alone. In such cases, other motivators might take on and maintain engagement, such as setting goals different to the designed goal of the task and making progress towards that other direction. Alternatively, there could be noise in infants' ability to monitor their progress, or they could be changing strategies for progressing over trials, improving on different variables depending on the trials, which would not necessarily show on our separate analyses of their performance.

5.4.2. Evidence of infants' learning with this task

One equivocal aspect of our results is the lack of clear evidence for above-chance performance and positive learning progress in the whole group of participants across the whole task, which transpired from our first analysis. In the last analysis as well, which looked at infants' learning progress over the first three triggered trials of the task, we found no significant evidence that the group progressed on average. However, we detected a heterogeneity of behaviours exhibited by the participants. Overall, the whole-group's progress was not above zero for any of the variables tested, but there were infants both above and below the zero line for each variable, and their

level of progress based on these variables influenced how much they were willing to engage as proposed by the learning-progress theory. While we did expect positive progress to lead to more engagement, we did not expect negative performance in this time window but rather a range of more or less steep positive progress following the expected bell shape of hypothesis 2.

5.4.2.1. The case of infants who engage for longer: good evidence of learning progress

Infants who engaged for longer tented to show the highest learning progress i.e., the most negative slopes in performance. While there was some heterogeneity in how much progress those infants who stayed longer made, overall this was in support of our hypotheses and of the learning-progress theory. It brought good evidence both that infants did learn over the course of the task and that their progress did influence their willingness to engage.

5.4.2.2. The case of infants who disengage: an avoidance strategy?

Infants who did not engage for very long however did not show this increase in performance at the start of the task but rather got worse from the start, then quitted early. There seemed to be an aversive effect of, so to speak, learning "regress" which got these infants to disengage. While we did not expect decreased performance already from the start of this task, this relationship between performance decrease and disengagement is part of the learning-progress theory too, and even of the study's hypothesis 2: a decrease in performance at the end of the bell-shaped performance curve is linked to disengagement. It is important to note that this relationship between performance decrease and disengagement does not have to be preceded by the steep learning and the plateau phases. In the case of tasks that exceed the difficulty level that an individual is able to learn from, disengaging early prevents getting stuck on a task that does not serve them, which is one key point of the learning-progress theory. Here indeed, the task did not seem suitable for some infants who did not manage to learn from the start and led to their disengagement, as predicted by the theory. Therefore, the lack of evidence for above-chance performance in the whole-task group average does not prevent to study learning processes in this task, but rather reflects a range of behaviours following different strategies for engaging or not, which still each follow the predictions of the learning-progress theory.

5.4.2.3. Task difficulty: a heterogeneity of strategies masking the effects

It is possible that mixing infants with heterogeneous strategies for, on the one hand, engaging with the task as they progress or, on the other hand, disengaging as they do not manage to learn, could lead to each type of effect cancelling out the other. This would result in an overall stable, baseline performance over the three stages of the task, as was found with this study. For this reason, it would be interesting to investigate this further with a less difficult version of the task that could allow all infants to learn and progress on it. Compared to the original task (Hochmann et al., 2016), the number of options was increased in the current task, which diminished the chance-level and helped the analysis, but increased the difficulty for the infants. Moving back to two options could help to make the task easier. Additionally, in our version of the task, the shapes were always of the same colour within one trial. This helped by decreasing colour salience, hereby avoiding a simple sensory reaction to more salient colours, a possible drawback in the original design discussed by Hochmann et al. (2016). Instead, the current design pushed for a higher-level analysis of the options and the task's rule, but again increased difficulty compared to the original stimuli. Finally, the introduction of a contingent outcome combining triggered rewards and passive feedback to increase infants' sense of agency during the task was appealing but might have introduced too much complexity in the task structure for infants to really grasp it. Our changes to the original task design overall increased difficulty, which then appears as a more evident issue in our case. However, it is equally important to ensure a high enough difficulty level to avoid infants disengaging for the opposite reason: a task that is too easy and thus boring.

5.4.2.4. The role of the reward

In order to distinguish between triggered rewards and passive feedback, a mixture of Kaldy et al.'s (2015) shiny and engaging outcome, and Hochmann et al.'s (2016) simpler outcome was used as respectively the task's reward and feedback. It is possible that the infants who did not engage for long with this task did not associate a high value to the reward and actually tried to avoid it. However, the choice of reward was based on Kaldy et al.'s (2015) design. In their study, Kaldy et al. could present 32 trials including 24 test trials to infants, with an average 11.17 to 12.78 valid trials for their groups of 10- and 8-month-olds, respectively. They appeared to get good engagement with their stimuli with no evidence of an avoidance behaviour, which makes it unlikely that participants would avoid the reward in the current study, and more likely that their

disengagement would be driven by their poor performance as predicted by the learning-progress theory.

5.4.2.5. Mitigated evidence for a learning effect in the original study

Finally, it is also important to note that evidence for infants' above-chance performance and their learning progress was also modest in the original study that inspired the current task. Hochmann at al.'s study used two measures of performance: *cumulative looking time* to the target and *first fixations* (as a binary score: to the target or to the other option). Overall, their results using *cumulative looking time* were more convincing than with *first fixations* – to note, *cumulative looking time* could not be used in our analyses because our paradigm was gaze-contingent. In their study, infants were presented with one of two different rule-learning tasks: match-to-sample (MTS), similar to the current study, and non-match-to-sample (NMTS), when infants had to gaze at the non-matching option instead. They included two very similar versions of the experiment with slightly different transitions (Figure 5.11). The second design also included test trials in which only one of the two options was shown to the infant and shown cards were visible at different times, with a delay, introducing a working memory component. These test trials are not discussed here as their design differs more with ours.

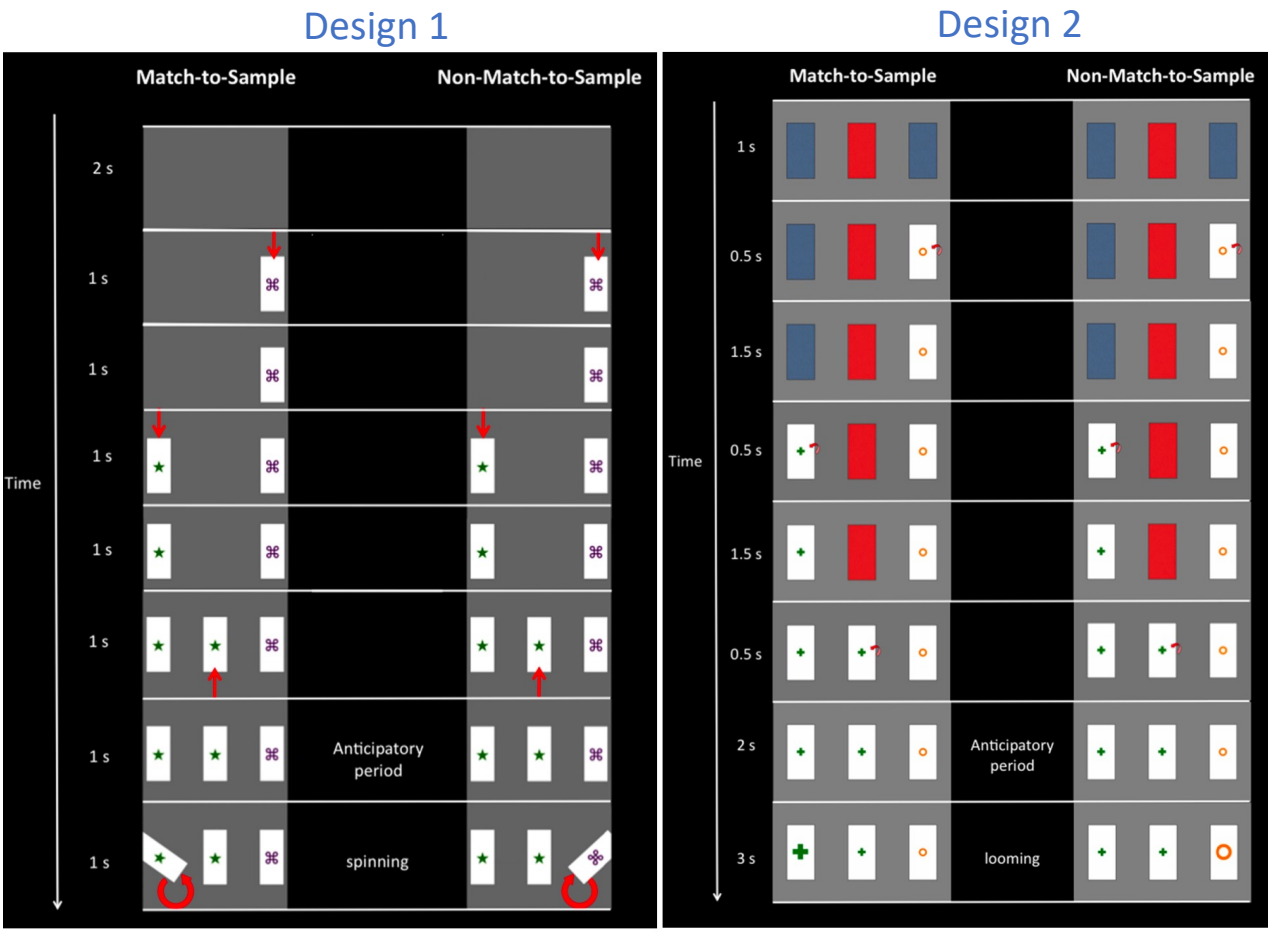


Figure 5.11: Paradigm design from Hochmann et al.'s study – taken with permission from Hochmann et al. (2016). Two slightly different designs were used.

With the first design (Figure 5.12), Hochmann et al. found that infants in the MTS task were at chance in block 1 and 3, and above chance in block 2 using the *cumulative looking time* variable but not the *first fixations*. In the NMTS task, infants' performance was above-chance in block 3 only, using either measure. With the second design (Figure 5.13), infants were found to be *below chance* in the MTS task's first block according to both measures, and at chance in the second block; there was no block 3 for this version of the paradigm. In the NMTS task, infants were at chance in the first block and above chance in the second block, significantly for the *cumulative looking time* variable and marginally for the *first fixations* variable.

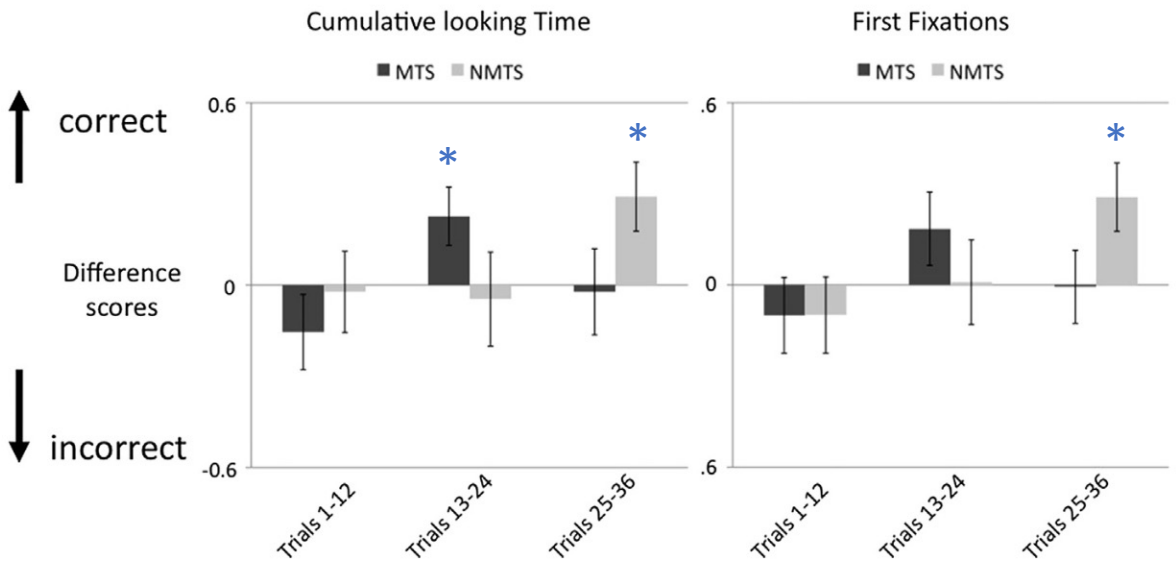


Figure 5.12: Results for the first design – taken with permission from Hochmann et al. (2016); significance stars were added in blue for values above chance.

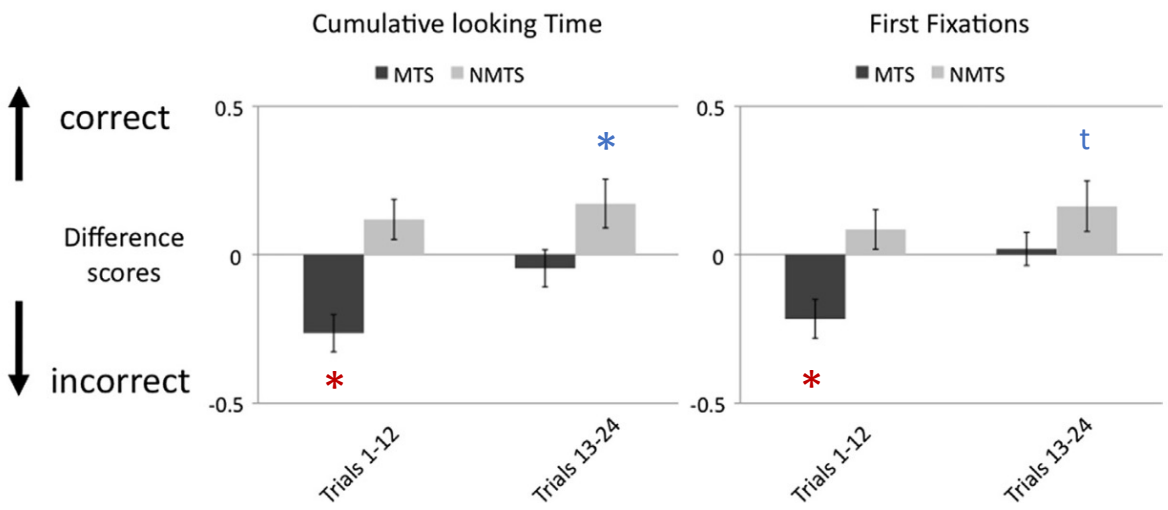


Figure 5.13: Results for the second design – taken with permission from Hochmann et al. (2016); significance stars or "t" for trending effects were added in blue for values above chance and in red for values below chance.

It was not noted whether infants' overall performance lied above or below chance in either design, for either task, and with either measure. However, it is unlikely to be above chance when taking all blocks into account. Infants' bias for the non-matching option in the second design

illustrates how quickly infants' preferences can change with small modifications to a paradigm's design, and how difficult it is to predict the direction of their preference. Thus, it is worth noting that changes in performance over time rather than absolute performance values, appear both informative and independent of infants' initial bias, making them a measure of choice when studying learning in infant populations. The findings from the current study largely rely on such measures and appear to match the level of evidence from similar studies in the literature.

5.4.3. Suitability and significance of the performance measures

In the present study, we used two score measures of performance (*first AOI hit is target* and *gaze hits target*), and four continuous measures (*time*, *number of fixations*, *distance* and *speed* to reach target) to ensure an investigation of gaze behaviour as complete as possible. We will discuss the validity of the score measures first, three of the continuous measures next, and the *speed* measure last.

5.4.3.1. Scores: first AOI hit is target and gaze hits target

Two of our measures were scores: *first AOI hit is target* and *gaze hits target*. Scores represent a fairly straightforward approach to measuring performance, with a simple yes/no answer to whether a participant has succeeded in a given trial. They resemble approaches taken in the school's curriculum grading system, yielding an overall percent correct for the whole experiment when taking all trials together, which is a quantity that one can easily get a good meaning of. However, they can be too reductive to show subtle changes in performance that pertain to the participants' way to get to the correct answer. Because they are binary, they are also ill-suited for looking at trial-to-trial changes and learning slopes, since the scores will oscillate between two extremes rather than slowly get from the low-performing end to the high-performing one, using a full range of values. In our paradigm, scores did not prove to be very informative, although they could be more helpful with an easier version of the task to determine whether infants generally know how to perform this task better than chance, and when they start to be able to if so.

5.4.3.2. Continuous measures: time, number of fixations and distance to reach target

The next most direct approach was to look at how long it took infants to get to the correct option, in terms of *time*, *number of fixations* or *distance*. How much (time, fixation or distance) it takes

for people get to the correct option is also a measure of performance often used in school environments and meaningful. These variables can be seen as measuring how much infants explored the space and searched for the correct option before staying on it. Individuals' exploration of alternatives has been shown to be relevant for learning processes: in a computer-based maths school game, exploration of options measured by mouse-tracking in 5- to 13-year-old children was shown to correlate with error rate (de Mooij et al., 2021). In the current study, none of our performance measures reflecting the exploration of the space before hitting target showed significant changes with the stages of the task. However, for all variables (*time*, *number of fixations* and *distance* to reach target), the change over the first three trials was significantly linked to infants' engagement. This followed the learning-progress theory. Because of the heterogeneity of learning progress exhibited by the group, it is possible that stage-to-stage effects were masked, which might be avoided by using a less difficult version of the task (see earlier Discussion point 5.4.2.3).

5.4.3.3. Continuous measure: speed

The last of our performance variables was a measure of how fast the infants were moving their gaze on their way to the correct option: the *speed* to hit the target. This is a very similar take on performance than the last measures (section 5.4.3.2), however here we measure how *fast* participants explore the space of alternatives rather than how *much*. In our analyses, this measure was the only one that came out significantly changing from stage to stage as well as over the first three trials. It appeared as the best reflection, out of the set of variables investigated, of how infants adapted their performance in our task. However, while stage-to-stage changes in *speed* followed our hypotheses, changes over the start of the task did not, which made it unclear what it reflects. Several possibilities for what the changes in this measure could entail will be discussed below.

5.4.3.3.1. Speed as a measure of processing speed or ability to process information

One way to look at *speed* is as a measure of how fast individuals can process one piece of information before moving on to the next. As stated above, measures discussed in section 5.4.3.2 reflect how much of the space infants explore before reaching the target, while *speed* reflects how fast they are at doing this i.e., how fast they can process the information encountered. This could mean that a decrease in *speed* reflects an overload, a difficulty to process the information.

However, it is likely not the case in this study that infants' decrease in *speed* only reflects a state of overwhelm, because the participants who slowed down the most over the first three trials, as reflected by a negative *speed* slope, were also those who completed the most trials.

5.4.3.3.2. Speed as a measure of engagement or motivation to process information

Another way to look at this is that *speed* could be a reflection of individuals' motivation to process the information. At the start of the experiment, for infants who do not manage to extract relevant information and learn from the task, the information that they are looking at is getting less and less engaging as they continue not to monitor any progress. Thus, as they engage less with the information, they process it less and less, brushing on it faster until they disengage completely and drop out of the task. Inversely, those infants who do show interest to engage with this information might be processing it more slowly on their path to the target in order not to miss out on relevant information. It is possible then that speed is directly inversely proportional to engagement. The size of the correlational effect for the speed variable was comparable to the other performance variables though, and did not indicate a stronger, more direct link than with the other measures.

5.4.3.3.3. Speed as a measure of confidence or task-general knowledge accumulation

Finally, another take on this measure could that *speed* reflects how confident infants are in their exploration strategies. Interestingly, bell-shaped changes in *speed* over the three stages were evidenced despite the heterogeneity of behaviours present in the group, that are argued to be masking the learning effects in other variables (Discussion point 5.4.2.3). If it is the case that all infants, regardless of their strategy to either avoid information because of a lack of progress or engage more and more with it thanks to their progress, all show an increase in *speed* from start to middle and then a decrease from middle to end, then *speed* is likely not a reflection of performance itself. Indeed, if the improvements in *speed* come alone, with no accompanying improvements in *time*, *number of fixations*, or *distance* to reach the target, then the participants are not really reaching the target and triggering it faster or more directly over time: this is not useful on its own for the task's goals and rewarding system.

However, *speed* could still reflect progress relating to how confident infants become in their gaze behaviour, regardless of how optimally they are performing the task. Indeed, over the course of

the experiment, they accumulate knowledge on the general design of the task, even if they do not manage to learn the rule itself. This kind of task-general learning is similar to general priors that were argued to help "random" participants to increase performance over the start of the experiment in Chapter 2. In this study, it can make infants more confident in where they want to explore, only because they know what type of information they are generally presented with, even if they do not know how to process it according to the task's rule. This could lead to faster saccade planning as the experiment goes on, and finally a slower disengaging phase when the resources allocated to performing the task have been cut and the saccade planning takes longer.

5.4.4. Future directions

This study is important and brings the first direct evidence that infants' behaviour does match the predictions of learning-progress theory, yet there remain questions to be answered and uncertainties to clarify that could be tackled in future research. As was argued in section 5.4.2.3, using a less difficult version of the task would be helpful to shed light on changes over the experiment that might have been masked in the current context because of heterogeneous behaviours within the group of participants, with some infants learning and others not. Several options have been proposed: reducing the number of options down to two, using shapes of different colours or renouncing the gaze-contingent outcome to simplify the task structure.

It would also be interesting to replicate these findings with other types of learning task. While emphasis was put on having abstract rule learning in the current task to move away from purely sensory processes, other types of learning possibly less abstract would be interesting to investigate too. Tasks such as sequence learning (ABA tasks) or any type of task on which infants can improve performance trial after trial would be a meaningful context in which to either replicate these findings or enrich them by revealing task differences. In particular, one could ask whether the role of the different variables investigated here would be similar in a different task: could infants adapt their performance along various measures of eye-movement control depending on the context?

Looking at other age groups, especially younger infants, would also take these results further. While the current study validates the learning-progress theory with 15-month-old infants, the question of when this type of learning is available to infants and whether younger infants or even new-borns are already capable of learning in a learning-progress guided manner remains to be

answered. Similarly, how older infants and even adults perform this task, and what gaze variables are relevant for describing how they adapt their performance along the task, is also a question that remains open.

Finally, looking at the relationship between learning progress and engagement in atypical populations, especially infants at risk for learning disabilities such as Attention Deficit and Hyper-Activity (ADHD), Autism Spectrum Disorders (ASD) or Down Syndrome would be a promising avenue for this research. It would not only help to better characterise these disorders but also shed light on the bounds of learning-progress-driven learning, if it did happen to be impeded in atypical populations.

5.5. Conclusion

This study brought the first piece of evidence for the learning progress hypothesis in infants. Indeed, 15-month-old infants' performance change at the start of the task (*time*, *number of fixations*, *distance*, or *speed to hit target*), was negatively correlated with their *number of completed trials*: the more they progressed from the start of the task, the more they remained engaged with it. These findings are important for the field of learning in infancy, but also more broadly in adulthood and in artificial intelligence, because they help to define what the steppingstones for learning are in a young brain that still has to grow and develop into its full potential. A question that remains is whether individual differences in how infants learn can be explained not only by their local progress but also by more general individual characteristics. Are there individual differences in learning stable across time points and contexts? Can they be linked to individual differences in more basic abilities, such as the "basic bricks" presented in the general introduction?

Chapter 6: Individual strategies for exploration and executive functions in 15-month-old infants

6.1. Introduction

6.1.1. Background

In the previous chapter, we pointed to infants' active role in sampling information to learn and showed that the time for which they decided to engage with the stimuli shown depended on their individual learning progress. Another situation in which infants learn through active information sampling in their daily life is when they explore objects. Their exploration can be purely visual through their eye-movements, or physical through mouthing, touching, or manipulating objects.

6.1.1.1. Visual exploration

Visual exploration is considered the earliest active exploration behaviour in infants: even new-borns have been shown to adapt their visual sampling of information depending on the stimuli. Indeed, new-borns adapt their gaze depending on stimuli characteristics and look more at salient stimuli such as high-contrast contours (Kessen et al., 1972), stimuli in motion (Slater et al., 1985), and faces (Valenza et al., 1996). Moreover, they are also able to adapt their looking behaviour to their own knowledge about stimuli: they prefer looking at stimuli that they don't know over ones that they are already familiar with (S. Friedman, 1972; Slater et al., 1982, 1984). This type of familiarity vs. novelty preference has been extensively studied in infants and evidenced in a variety of contexts (see Sirois & Mareschal, 2002 for a review). Provided that the stimuli are of age-adapted complexity, infants' preference has generally been shown to change over time: at first, infants exhibit a preference for familiar objects, until they become habituated and switch to preferring novel objects (Hunter et al., 1983). This type of novelty-familiarity bias has been argued to stem from broader principles of learning and curiosity. For example, optimal-level theories

claim that individuals prefer information of intermediate novelty or complexity (Mather, 2013). This shows that infants, already from birth, are well capable of guiding their own learning experience in order to sample visual information that is most helpful for them to learn.

6.1.1.2. Manual exploration

As they grow, infants build a wider and finer repertoire of motor actions that they can perform to sample information, and engage in more complex multisensory exploration behaviours, notably through manual exploration and play (Gibson, 1988). As for visual exploration, infants' manual exploration of toys has also been shown to be influenced by stimulus properties such as salience, and by infants' own knowledge about them. Pre-schoolers have been shown to manipulate and play more with toys that are presented to them in a way that leaves uncertainty about their functions as opposed to toys which functions are clearly presented, and appear to engage in manipulations directed at learning about the unclear functions (Cook et al., 2011; L. E. Schulz & Bonawitz, 2007). They also appear to play more with objects that have violated their expectations (Stahl & Feigenson, 2015; van Schijndel et al., 2015) and to adapt the type and quantity of their actions on toys to prior information about the toys (Ruggeri et al., 2019; Siegel et al., 2021), again in ways that enable them to disambiguate the properties of the object. Overall, play through manual exploration appears as a primary means for active learning during development, something that had long been argued by psychologists and educators (Groos, 1901; Montessori, 1912; Piaget, 1952; Yogman et al., 2018).

6.1.1.3. Exploration as a predictor for later cognitive development

Several studies have looked at the link between how infants play and manipulate toys, and their later cognitive development. Kopp & Vaughn were amongst the first ones to bring evidence for this link and showed that the time for which preterm infants' engaged with a new toy at 8 months predicted their cognitive development at 2 years as measured by the Bayley Mental Scale (Kopp & Vaughn, 1982). More recently, Bornstein at al. showed that 5-month-old infants' motor maturity and amount of exploration, scored during 50 minutes of mother-child interactive play, was linked to higher academic achievement at 14 years (Bornstein et al., 2013). It was also shown that 5- to 19-month-old infants' exploratory efficiency, the number of set functions that they discovered in a given amount of time when playing with a toy, is stable over 4 visits within 9 months and predicts infants' IQ score at 3 years (Muentener et al., 2018). Finally, Slone et al.

measured 15-month-old infants' manipulation of toys through the variability of the toy's visual images that they generated, which was found to predict their vocabulary growth 6 months later (Slone et al., 2019).

In the visual domain as well, some evidence was brought that visual exploration is linked to later cognition. Infants' maximum look duration towards a puppet toy at 5 months has been shown to be inversely proportional to their composite executive function score at 24, 36, and 48 months (Cuevas & Bell, 2014). Additionally, infants at a high risk for Autism Spectrum Disorder (ASD) have been found to fixate visual stimuli for shorter durations at 6 to 9 months of age, which was also linked to higher probability for an ASD diagnostic at 3 years (Wass et al., 2015), and to return more to previously fixated locations at 8 months (Gliga et al., 2018), compared with infants at low risk for ASD. Hence, there is ample evidence that exploration, both visual and manual, is linked to cognitive development.

6.1.1.4. Linking cognition and exploration at the same point in time

While there have been a few studies looking at the link between exploration in infancy and later cognitive abilities, there is not much work looking at how exploration and cognition relate at the same point in time. Yet, individuals explore in unique sequences, and measures of exploration have proven to be highly variable, such that characterising individual differences potentially driving exploratory behaviours at one point in time would be particularly beneficial. In one of the only studies on the topic (Caruso, 1993), infants' manipulation of toys at 11-12 months was extensively documented through the coding of 14 behaviours such as squeezing, mouthing or combining, from which different measures were derived, notably the exploratory breadth (the number of behaviours infants used) and depth (the average number of times behaviours were used), which were put in relation to infants' problem solving abilities. The authors found that infants' exploratory breadth was the only aspect of exploratory behaviour to relate to problem solving abilities, both in terms of success and sophistication. To our knowledge, there is no work to date that tried to relate more standard measures of infants' cognition, such as executive functions, and individual strategies for exploration at the same point in development.

6.1.1.5. Linking manual and visual exploration

Moreover, Caruso's study focused on a description of infants' exploration that relied on manual play. However, one could ask whether individuals' exploration in different modalities are related, especially for visual and manual exploration which have both been studied extensively. There is work suggesting that how infants' attend to visual objects is influenced by the actions that the infants would be able to perform on those objects, which they termed "graspability" (Kaufman et al., 2003). Conversely, as stated before, when infants manipulate objects, they change the angles from which they are visualising those objects, in a way that predicts their later vocabulary growth (Slone et al., 2019). Thus, infants' exploration in the visual and manual modalities appear to be intertwined, and it is possible that individual variations in one modality could also be found in the other.

6.1.2. Current study

In this context, the current study was designed to investigate two outstanding questions: 1) is infants' exploration in the visual domain linked to how they explore manually? and 2) are individual variations in infants' exploration strategies in either modality linked to their individual executive-function abilities? To this aim, 15-month-olds infants were tested in a four-task study comprising: 1) an eye-tracking learning task which was the focus of the previous chapter, 2) a free-play task, 3) an inhibition task, and 4) a memory task. 15 months appeared as an adequate age point for answering these questions as infants at this age have a sufficiently developed visuo-motor repertoire to explore objects according to their interests without being too limited by their capacities to reach for or grasp them for example. It is also the earliest age point at which their executive functioning can be measured through questionnaires, making this age point particularly suitable to investigate the questions at hand.

Infants' visual exploration was assessed during the familiarisation trials of the two screen-based tasks at the start and the end of the study. Such trials were used to establish infants' familiarity with stimuli before using them in the test trials of the screen-based tasks. However, these trials could also be used to investigate infants' visual exploration of objects, regardless of the upcoming task. Indeed, during these trials, infants were presented with 6 objects at once and left free to visually explore them for 10s at each trial. The fact that there were two different screen-based tasks at the start and end of the study was used to assess the stability of infants' exploration along

the study. We also looked at infants' manual exploration of toys throughout the free-play session, during which they were presented with 8 toys at once for 4.5min, similarly to one of Caruso's set-ups (Caruso, 1993).

We aimed to establish measures of exploration strategies that could be used for each task-type, despite the important design differences. Thus, we used a depth-breadth approach to measure exploration strategies, similar to Caruso's approach but based on simpler, time-dependant rather than action-dependant criteria, such that it could be employed in each task-type regardless of the design differences. *Breadth* was defined as the number of objects that infants explored, either visually or manually, while *depth* was characterised by their maximum time spent on one object. Both measures were standardised: the number of objects was divided by the size of the task's object set, and the time on objects by the task's total exploration time allowed, such that each score was expressed as a percentage of exploration of the space (in terms of object set or time window) and could be compared between tasks despite the different designs. This definition of breadth and depth is substantially different from Caruso's: it is defined not by the various actions that infants performed, which was applicable for play but for which there was no corresponding measure identified in the visual modality, but in terms of overall allocation of attention to each object, which applies to both domains. This appeared as a pertinent definition: indeed, decisions to maintain or switch focus on objects are generally at the core of machine learning algorithms for artificial intelligence (Oudeyer, 2017) and appeared as an important aspect of exploration in the case of infants as well.

Additionally, infants' development and executive functions were also measured through two tasks and two questionnaires. One eye-tracking task, based on Guillory & Kaldy's work (Guillory & Kaldy, 2019), was used to measure infants' visual working memory, while a toy-based "glitter-wand" task based on Friedman et al. (N. P. Friedman et al., 2011) was used to measure inhibitory control. Attentional Focusing and Shifting as well as Inhibitory Control scores were also obtained using the Early Childhood Behavioural Questionnaire (ECBQ) as well as a general development score using the Ages and Stages Questionnaire (ASQ).

Similarly to Muentener et al.'s study, in which the authors chose to remain "agnostic" to the particular relationships that could link their measures of exploration and cognition, we also decided not to form specific hypotheses to link particular measures of infants' exploration to

particular executive-function scores. Indeed, while there is little doubt that executive functions are important for learning and exploration, there is no consensus on the specific role that they might play for exploration strategies. Thus, we formed two general hypotheses:

1) Individual differences in infants' strategies to explore are stable across the lab visit and across modalities, such that, for example, infants who explore more into depth in the first batch of familiarisation trials would also explore more into depth in the last batch of familiarisation trials, as well as during the play session.
2) Individual differences in depth vs. breadth exploration strategies are driven by individual differences in executive functions (memory, inhibition, attention), regardless of developmental stage.

6.2. Methods

The tasks described in this chapter are part of a four-task study, comprising a card-matching task, a play session, an inhibitory task and a memory task, as well as parent-report questionnaires. The card-matching task was described separately in Chapter 5, where general information about the study can also be found. To avoid repetition, only additional information about the other three tasks is added here. However, this chapter does comprise analyses of data from card-matching task (familiarisation trials only), as well as the other tasks. As stated in Chapter 5, this study did not rely on previous work from other people: the experiment's design, data collection (with the help of interns) and analysis were performed by me.

6.2.1. Participants

76 participants were tested for this study, as described in the Methods section 5.2.5 of Chapter 5. Infants were included even when they did not contribute data to all of the tasks, to maximise power since the analyses did not include all tasks at once. Thus, depending on the analyses and the type of data assessed, the groups of participants involved may slightly differ. Differences in exclusion rates and their justifications are described in the next paragraph. Demographics for each group are detailed in Table 6.1.

For the *play task*, 72 out of the 76 infants tested were included in the analyses, while 4 were excluded (3 technical issues, 1 infant did not want to play). For the *glitter-wand* task, 68 infants were included and 8 were excluded (5 technical issues, 2 infants got distressed before or during the instructions, 1 infant was found not to be a primary English speaker). For the *memory task (test trials)*, 60 infants were included and 16 were excluded: (11 infants had less than 3 clean trials, 2 technical issues). For the *memory task (familiarisation trials)*, 73 infants were included and 3 were excluded (2 technical issues, 1 drop-out before the end of the familiarisation phase). For the *card-matching task (familiarisation trials)*, 69 infants were included and 7 were excluded (5 didn't cooperate for this task, 2 technical issues). For the *questionnaires*, 69 infants were included and 7 were excluded because no questionnaires were returned by post after the lab visit.

Group	Total n. (n. fem.)	Mean age (S.D.)	Min. – max. age
Play task	72 (37)	15m, 1.3d (11.0d)	14m, 10.2d – 15m, 21.8d
Inhibitory task	68 (36)	15m, 0.7d (10.9d)	14m, 10.2d – 15m, 21.8d
Memory task (test trials)	60 (31)	15m, 0.38d (11.5d)	14m, 10.2d – 15m, 21.8d
Memory task (familiarisation trials)	73 (37)	15m, 0.7d (11.3d)	14m, 10.2d – 15m, 21.8d
Card-matching task (familiarisation trials)	69 (36)	15m, 0.9d (11.1d)	14m, 10.2d – 15m, 21.8d
Screen-based tasks (familiarisation trials)	75 (38)	15m, 0.8d (11.2d)	14m, 10.2d – 15m, 21.8d
Questionnaires	69 (34)	15m, 3.0d (12.6d)	14m, 10.2d – 16m, 8.3d

Table 6.1: Demographics for each group of participants, depending on the tasks at hand: total number of participants, total number of females, mean, standard deviation (S.D.), minimum and maximum values of age (m: months, d: days).

A sensitivity power analysis was conducted with the software G*Power 3.1(version 3.1, https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower), for each sub-sample and with standard values of significance threshold $\alpha = 0.05$ and power $1-\beta = .80$. The results of this analysis are summarised in Table 6.2 below and revealed that the study was partly to fully sensitive to small effects for all types of tests and sub-samples ($R > 0.3$, $F > 0.2$ and $t > 0.2$).

Sub-sample	Test	Detectable effect size (critical threshold value)
N = 69	t-test for 2-sample matched pairs	dz = 0.34 (t = 2.00)
	Correlation	R = +/- 0.24
N = 72	t-test for 2-sample matched pairs	dz = 0.33 (t = 1.99)
	Correlation	R = +/- 0.23
N = 60	t-test comparison to zero	d = 0.37 (t = 2.00)
N = 72 or 75	Two-tailed multiple regression for 7 factors	R = 0.21

Table 6.2: Sensitivity power analysis for each sub-sample and each associated test.

6.2.2. Procedure

The general procedure for the whole study was described in Chapter 5, section 5.2.2.

6.2.2.1. Card-matching task

The experiment started in an eye-tracking booth with a card-matching task, for which the detailed procedure was described in Chapter 5, section 5.2.3.

6.2.2.2. Play session

Infants were then taken to a playroom and sat on a highchair in front of a same-height wooden table, while a calm instrumental music was played in the background ("Sway" by OCB Relax). The parent and the experimenter both sat behind the participant, respectively to their left and right, at a rough distance of 1.5m. Infants were presented with 11 toys in total (Figure 6.1). The session was divided in two parts. 1) First, infants played with 3 toys (respectively, the cube, the ball and then the chicken toy, see Figure 6.1) separately as a familiarisation phase. The experimenter placed the toys on the table in front the baby and said "Look!". Infants were allowed to play with each toy for 1 min. 2) Then, the experimenter emptied a box containing the remaining 8 toys (Figure 6.1) on the table and let the infant freely play with them for 4min30s. The toys were chosen to be as different as possible in their shapes and usages, while being of equivalent size.

Each time a toy was dropped on the floor, the experimenter picked it up and placed it back on the table behind the other toys. Parents were instructed not to grab toys that their child might hand them and not to talk with their child, but were encouraged to smile when their child turned to them to avoid infants' frustration. The experimenter followed the same guidelines. If the child dropped the same toy more than 5 times, it was put back into the box and, in the case of single toy presentations, the next toy was presented. For the single presentations, the first three toys were always presented in the same order and placed upwards on the table. For the multiple presentation, directly emptying the box by turning it upside down ensured a near-random placement of the toys at the centre of the table. Participants' behaviour was recorded concurrently by three fixed cameras to their front and left & right sides.

Figure 6.1: All toys, pictured together for scale; the set-up during the study is visible in Figure 6.5.

6.2.2.3. Inhibitory task – the glitter wand task

At the end of the play session, the experimenter and the infant placed the toys back in the box and the box was put away. The experimenter came to the right side of the participant to show

them a glitter-wand toy (Figure 6.2), saying “Look!”, moved to the front of the participant across the table and showed them the toy again saying “Look!”, then hid it under the table, ensuring eye-contact with the participant and taking a calm, serious facial expression, said “No. No, don’t touch. No.” while shaking their head and index finger. Finally, the experimenter placed the toy in the middle of the table in the participant’s reach, about 30cm away from them, bowed their head to avoid eye contact and waited for 30s, or until the participants grabbed the toy. If the participant did not grab the toy, the experimenter handed it to them. Participants’ behaviour was recorded by three fixed cameras around them, as in the previous task. This procedure is a typical test for inhibitory control in infants, used from 14 months of age when introduced by Friedman et al. (N. P. Friedman et al., 2011).

Figure 6.2: Glitter-wand toy; the set-up during the study is visible in Figure 6.5.

Participants systematically took a break after this task: they were left to sit or roam on the floor to play with the toys of the previous tasks and with bubbles blown by the experimenters. This lasted for approximately 10min but could be shortened or lengthened at the parent’s request and depending on the child’s needs.

6.2.2.4. Memory task

Finally, participants were taken back to the first booth and again sat on their parent’s lap about 60cm from the Tobii TX300 screen-eye-tracker system. They were systematically given two rice cakes as a soothing and motivational food to help them carry on with the last task of the study. One experimenter sat behind them hiding behind a curtain and handed parents more cakes as infants finished them. The eye-tracker was calibrated using a 5-point presentation before the task started. The task was based on Guillory & Kaldy (2019). The stimulus set was composed of 36

drawings of animals extracted from a database of colourful animal cartoon drawings (Blackleaf Studios, www.mygrafico.com). In test trials, they were presented on a background of light blue room optimally contrasting with any stimulus colour, depicting three shelves on which three different animals were drawn. This background was the same across trials and was realised using the Sweet Home 3D software (https://www.sweethome3d.com).

The task was made of 2 blocks: an object familiarisation block (6 trials), and a test block (9 trials). For both blocks, each trial started with a medium grey screen shown for a randomly drawn duration within the 1000-1500ms interval (inter-stimulus interval, ISI), before a contingent red fixation spiral appeared on the grey background (fixation phase, FP). The fixation spiral disappeared after the participant's gaze was detected on it for a continuous 500ms.

The object familiarisation trials were organised similarly to the card-matching task: after the ISI and the FP, objects were presented for 10s in a circle, 6 at a time, on a grey background.

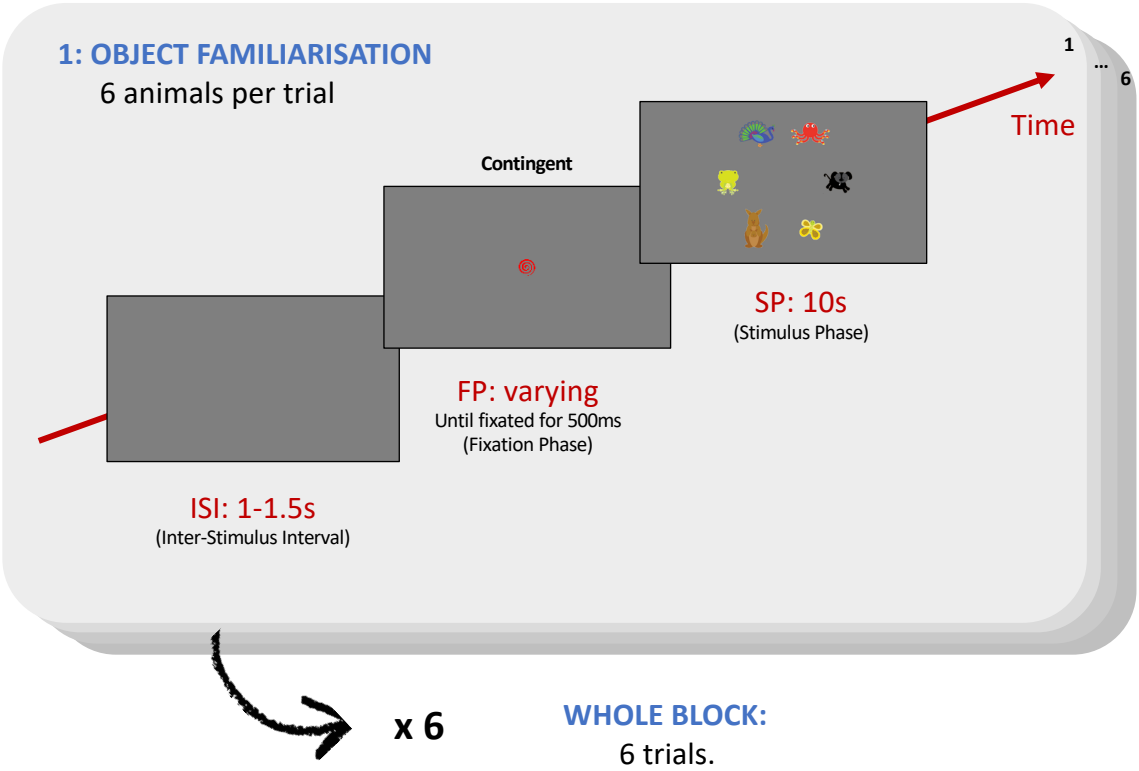


Figure 6.3: Structure of an object familiarisation trial; the trials were composed of an inter-stimulus interval (ISI) of 1-1.5s, a fixation phase (FP) of varying length contingent on the participant's gaze, and a stimulus phase (SP) of 10s. There were 6 such trials, each comprised of 6 shapes.

In the test trials, the ISI and FP were followed by the presentation of a light blue room optimally contrasting with any stimulus colour, depicting three shelves on which three different animals were presented for 6s (encoding phase, EP). A medium grey screen was then shown for 2s (blank). Next, the background room reappeared with two of the formerly shown animals and one new animal replacing one of the former ones (test phase, TP). After 6s, a rewarding outcome was presented for 2.5s: the new animal spun and loomed as a shiny purple animation appeared in the same location (reward phase, RP).

Animals were pseudo-randomised across participants, familiarisation and test trials such that for each participant, they appeared once per type of trial, and two animals of the same colour could not be presented in the same test trial. The location of the changed animal was pseudo-randomised so that each location was the changed location twice within the first 6 trials and once in the last set of 3 trials.

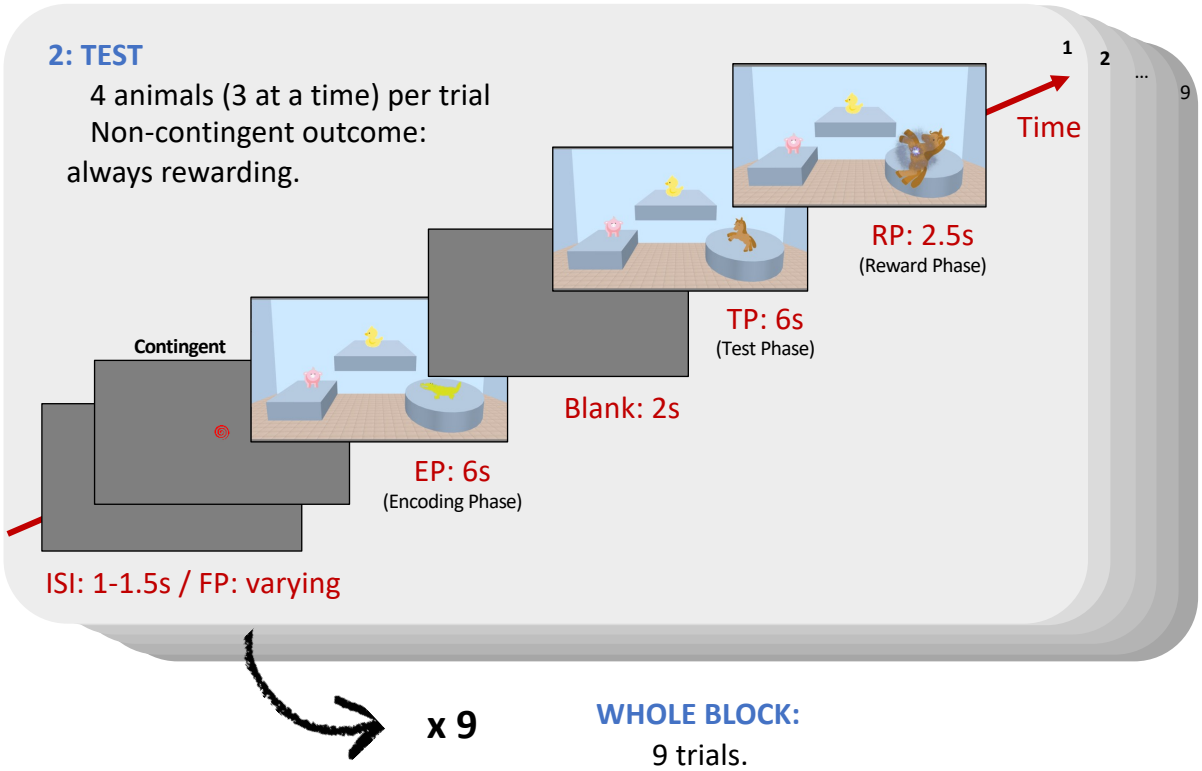


Figure 6.4: Structure of a test trial; the trials were composed of an inter-stimulus interval (ISI) of 1-1.5s, a fixation phase (FP) of varying length contingent on the participant's gaze, an encoding phase (EP) of 6s, a blank (grey screen) of 2s, a test phase (TP) of 6s and an reward phase (RP) of 2.5s. There were 9 such trials, and each of them comprised 4 animals:

3 animals at a time, 1 animal that changed between EP and TP. To note, the outcome of the trials was not contingent in this task and participants always received a reward at a fixed time, independent of their gaze.

At the end of the task, as in the learning task presented in Chapter 5, infants were presented with a 9-s rewarding clip of animals clapping, taken from a scene of Classical Baby (HBO).

6.2.2.5. Questionnaires

At the end of the study, parents were given two questionnaires about their child: the Early Childhood Behavioural Questionnaire (ECBQ) and the Ages and Stages Questionnaire (ASQ). They were asked to fill them in at home, as per the questionnaires' recommendations, and to return them by freepost.

The Early Childhood Behavioural Questionnaire, ECBQ (Putnam et al., 2006) is a parent-report questionnaire that was already used and described in Chapter 4. In short, this questionnaire assesses 18- to 36-month-old infants' temperament according to 18 7-point scales (see questionnaire and scoresheet with all the scales in Appendix 3 and Appendix 4). There is another version of this questionnaire designed for infants until 12 months of age, but there is no specific questionnaire for the age gap of 12 to 18 months, into which our group of 15-month-olds falls. We decided to use the questionnaires for older infants, and informed parents that some questions might be less suited for their infants as they were designed for slightly older infants. We were interested in scales related to infants' executive functions ("Attention Focusing", "Attention Shifting", "Inhibitory Control").

The Ages and Stages Questionnaire, ASQ (Squires et al., 1995; Squires & Bricker, 2009) is also a parent-report questionnaire, which is used to assess developmental progress in children from 1 month to 5.5 years. We used the third revision of this questionnaire (ASQ-3), the latest one at the date of the study, and the 16-month version, designed for infants from 15 months and 0 days to 16 months and 30 days. The age group tested (14.5- to 15.5-month-olds) was between two questionnaires and some of the infants were 15 days younger than the intended date of use for this questionnaire, while others would have been 15 days older if using the 14-month version. The questionnaire comprises questions about daily activities grouped around five themes:

communication, gross motor, fine motor, problem solving, and personal-social skills. It is visible in Appendix 8.

6.2.3. Pre-processing of the data

6.2.3.1. Video data: play and inhibitory task

Infants' behaviour at the playroom's table was recorded at 25Hz from three angles at once, as shown in Figure 6.5. The videos were manually coded offline, frame-by-frame, using the software ELAN (https://archive.mpi.nl/tla/elan) where they could be visualised and coded all together. The coder always used the front angle video as the reference framed and used the other frames as an aid when the view was obstructed or ambiguous in the reference frame.

Coders first marked the start and end of the play session, and then the start and end of infants' manipulation of the toys. Each toy had a different coding stream, such that the coding of each manipulation included information about the start and end time of the event as well as the identity of the related toy. Manipulations of several toys could overlap, did not necessitate contingent gaze to the toy; indirect manipulation was included e.g., when the infants pushed one toy using another toy; involuntary touch was excluded e.g., when the infant briefly touches one toy while reaching for another, and events when the infants pushed all toys away from the table at once, were also excluded.

Finally, coders also marked the start and end of the inhibitory task, from the moment that the experimenter's hand left the glitter wand, to the moment that the infant's hand touched the wand, or at 30s after the start if the infant did not touch the toy before then.

For a randomly picked 10% of the videos (7 infants), infants' interactions with objects were double-coded by another coder to assess reliability. To assess the consistency of the coders' marking for both of the toy manipulations' start and end, and of the experiment's start and end, two separate intraclass correlation coefficients (ICCs: McGraw & Wong, 1996) were computed using Arash Salarian's ICC function for Matlab (Salarian, 2022). The coders were found to be highly reliable for both manipulations ($r = 0.9817$, $p < 0.0001$) and experimental timings ($r = 1.0000$, $p < 0.0001$). There 6 instances out of a total of 168 manipulation events when one coder did not code an event that the other did (and 162 events that they both coded; 3.57% disagreement).

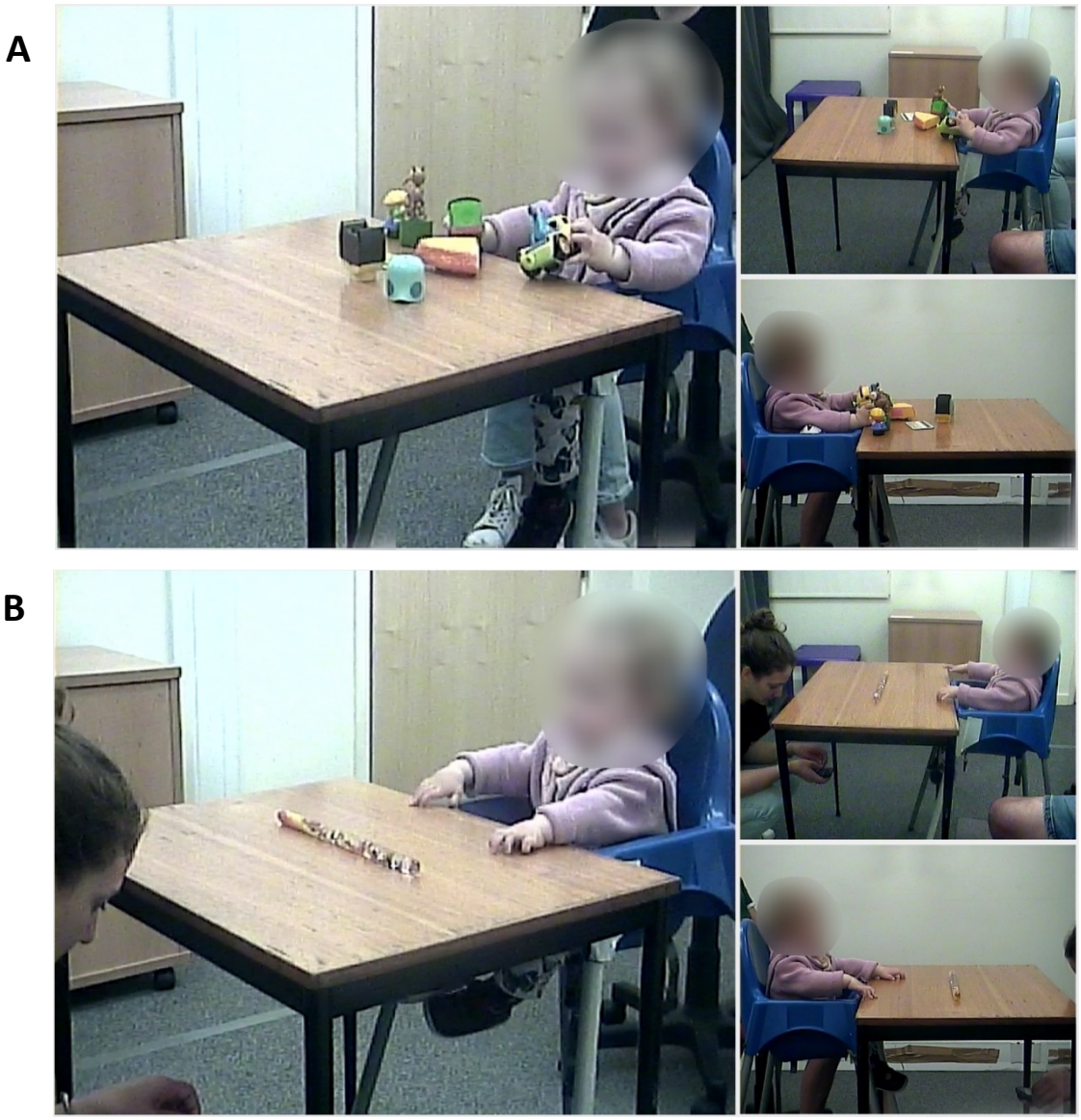


Figure 6.5: Example screenshot of a participant's video recording for the play session (**A**) and the inhibitory task (**B**), recorded simultaneously from three fixed cameras.

6.2.3.2. Eye-tracking data: memory test trials, memory and card-matching familiarisation trials

The eye-tracking data was pre-processed following the same pipeline as for the card-matching task, presented in Chapter 5, except for task-specific steps. This pipeline was used for the test trials of the memory task, as well as for the object familiarisation trials of both the card-matching and the memory task, analysed together to assess free exploration of visual items irrespective of the task to come. The pipeline is summarised in Figure 6.6.

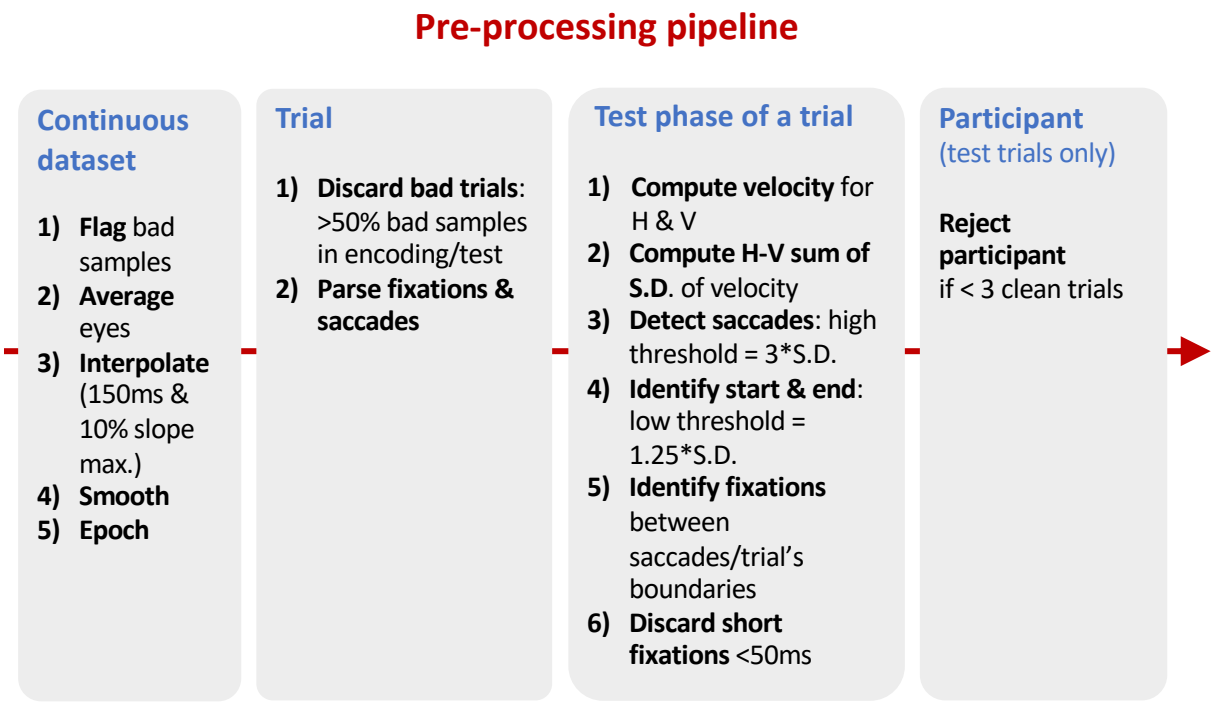
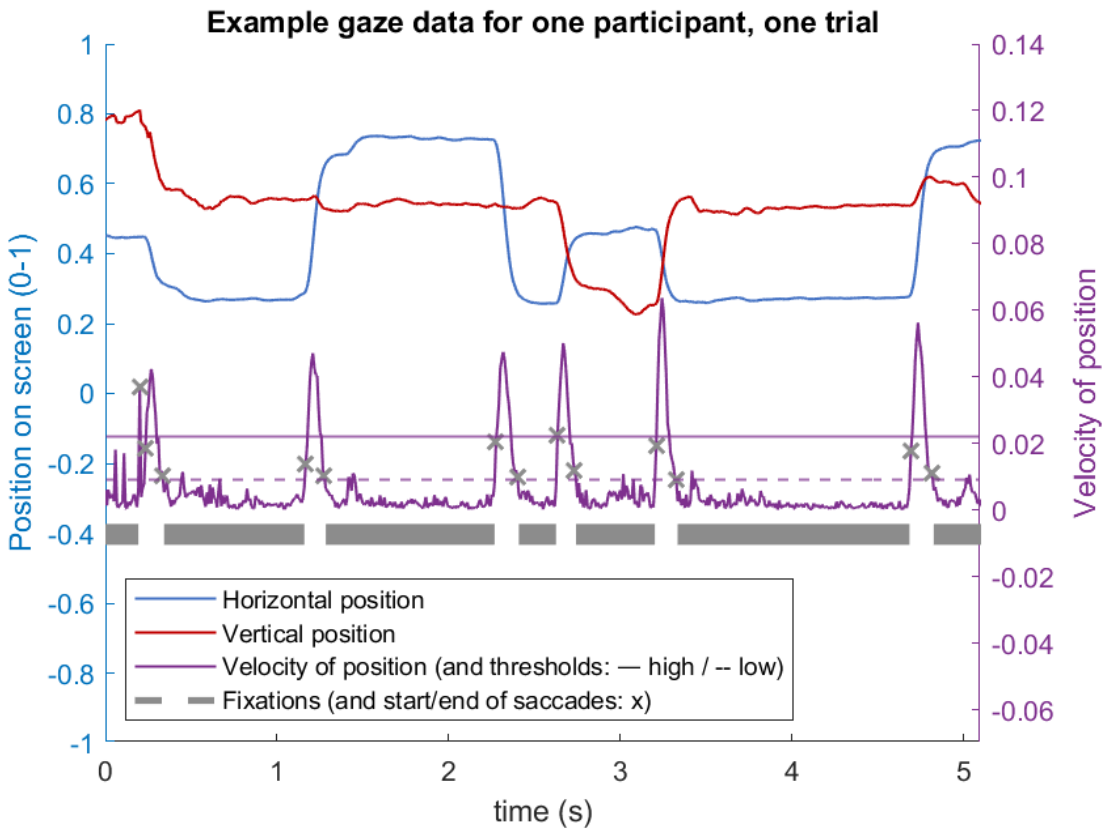


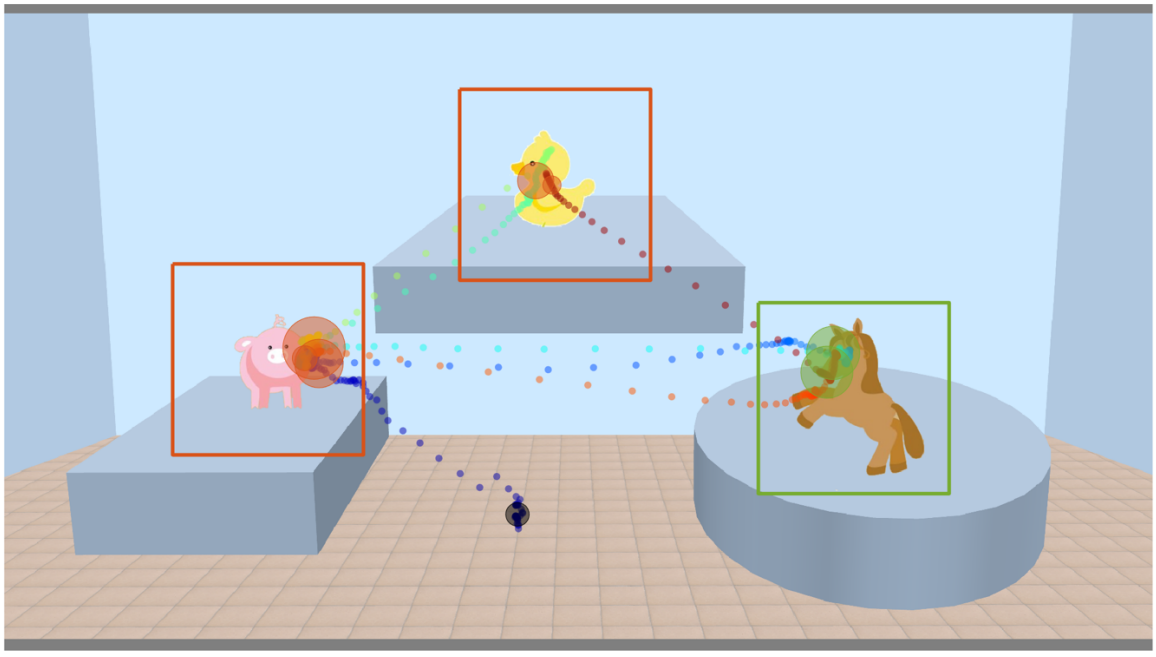
Figure 6.6: Summary of the eye-tracking data's pre-processing pipeline. The pre-processing was performed participant by participant, in 4 main stages on 1) the continuous dataset, 2) each trial, 3) each trial's test phase (horizontal H and vertical V coordinates, considered separately until now, were grouped in step 2), 4) for test trials (not for familiarisation trials): at the whole participant level. Each stage was composed of different steps summarised here and developed in-text below.

The pre-processing was performed participant by participant, in 4 main stages on 1) the continuous dataset, 2) each trial, 3) each trial's test phase, 4) for test trials (not for familiarisation trials): at the whole participant level. The steps taken were the same as described in Chapter 5, with the difference that no constraint was imposed based on looking at a central area of interest at the start of the trials. After pre-processing, 10 participants had less than 3 clean test trials and were excluded. No participant was excluded from the familiarisation trials analysis. The steps were explained more into depth in Chapter 5, section 5.2.4, and are only summarised here in Figure 6.6 to avoid repetition. Example plots of data from one participant's test trial, as well as another participant's familiarisation trial are provided for reference (Figure 9).

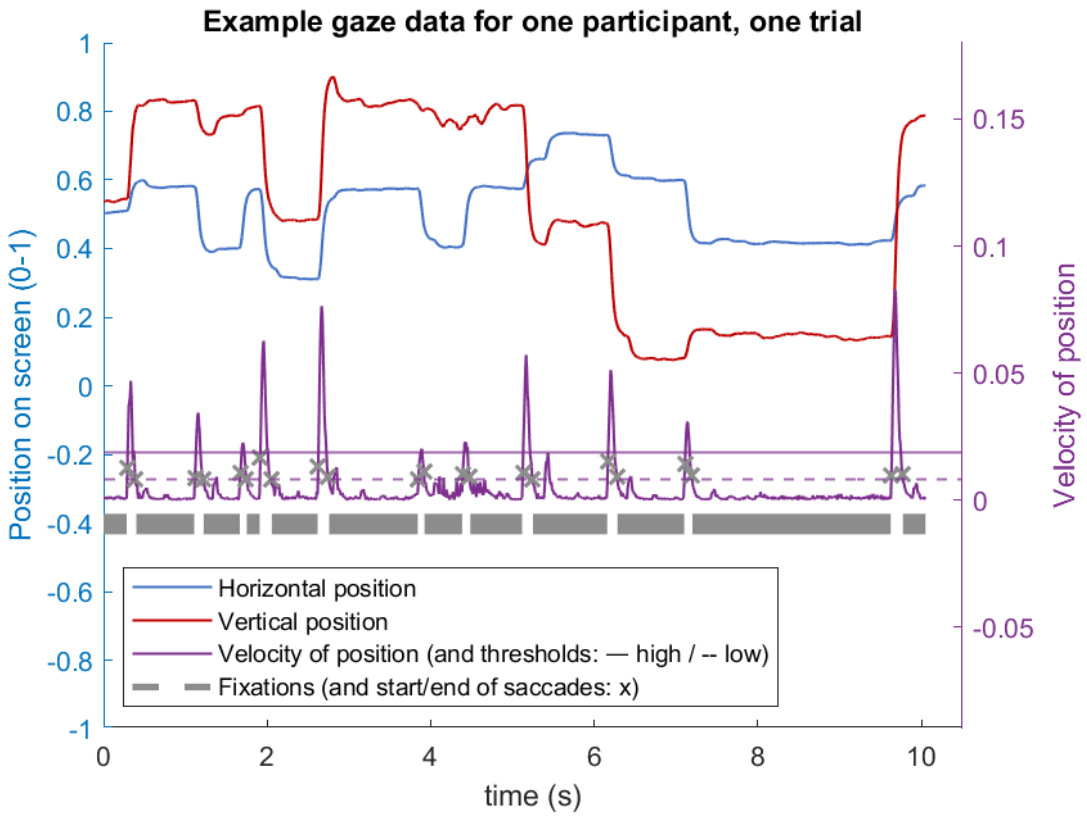
A


Example gaze data for one participant, one trial
Position on screen (0-1)
Velocity of position
time (s)
Horizontal position
Vertical position
Velocity of position (and thresholds: — high / -- low)
Fixations (and start/end of saccades: x)

B



C



D

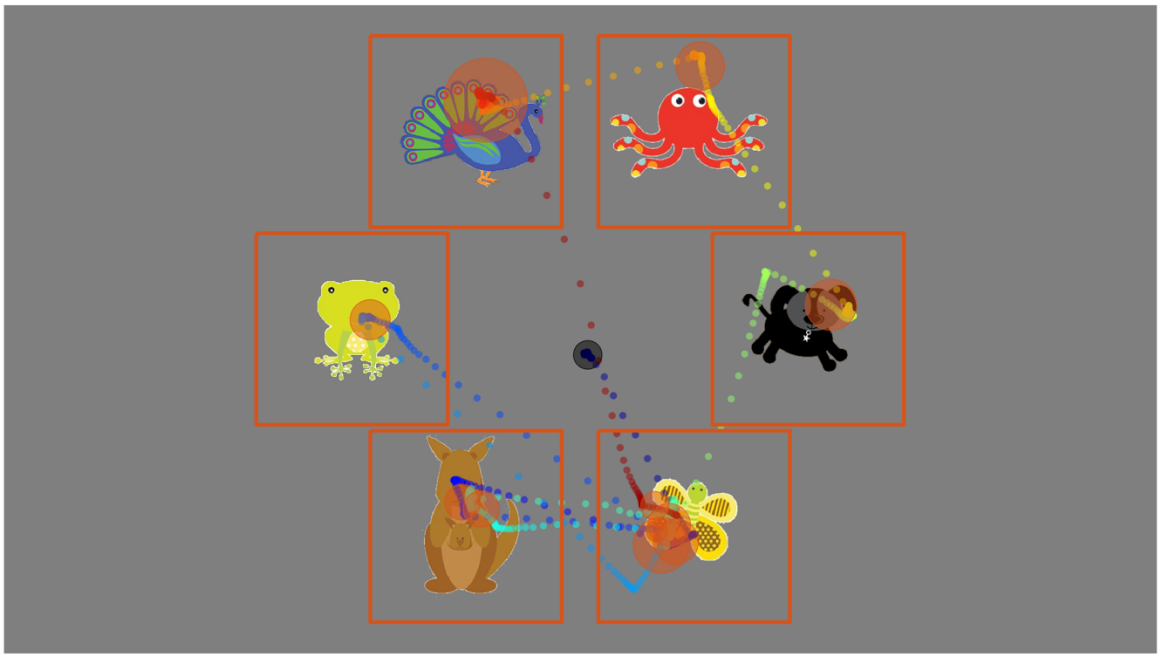


Figure 6.7: example data for one participant's test trial (**A** & **B**) and familiarisation trial (**C** & **D**).

A. & **C.** Gaze location (horizontal and vertical), velocity and fixations over time, illustrated for a test trial (**A**) or a familiarisation trial (**C**). The x-axis represents the time course of the trial, the left y-axis represents the gaze position as a proportion of the screen size (0-1)

horizontally (H, in blue) and vertically (V, in purple). The H-V composite S.D. of the velocity is represented along the right y-axis in orange together with its threshold: low (dotted line) and high (continuous line). The fixations identified from those thresholds are represented as green boxes and the saccades (and fixations) start & end are represented as green crosses.

B. & **D.** Gaze and fixations in space for a test trial (**B**) and a familiarisation trial (**D**). Note that **D** is an illustration of a familiarisation trial for either task, and that the items displayed could be shapes in the case of the learning task or animals in the case of the memory task (case displayed here for reference). Gaze samples are represented as small circles over the screen seen by the participant during the test phase. Their colour follows the time course of the trial: the first samples are represented in blue, through green, yellow, orange, and red for the last samples. Larger circles represent fixations to the target (in green, only applicable for test trials, depicted in **B**), other items (in orange) or the background (in black). The diameter of each fixation's circle is proportional to its duration. Finally, each area of interest around the animals is represented as a square in green (target, only applicable for **B**) or orange (non-target).

6.2.4. Analysis of the data: variables extracted

Several variables were extracted and plotted using the software Matlab version R2020b, and statistical tests were then performed using the software JASP, version 0.15 (Love et al., 2019). Variables extracted can be grouped into two main categories, depending on the tasks that they rely on: exploration-related variables, and executive-function-related variables, each described below.

6.2.4.1. Exploration-related variables

Two types of exploration were investigated in this study: manual exploration during the play session and visual exploration during the familiarisation trials of the screen-based tasks. To favour comparison, we extracted associated measures in each task: there was a one-to-one correspondence between the variables of each task. Outliers (points further than of 1.5 times the threshold for the 25% and 75% quartiles) were removed to avoid spurious results. The variables are reported here in a specific order respecting the correspondence between tasks.

6.2.4.1.1. Play

In the play task, we extracted two different variables of *manual* exploration: 1) the *number of toys* interacted with which is a measure of exploratory *breadth*, and 2) the *maximum overall time* spent on one toy, a measure of exploratory *depth*. The overall time per toy, from which the second variable was derived, was obtained by summing all the individual interactions with a given toy. To standardise variables across designs and sets and facilitate comparisons, these variables were divided by the whole space available: the *number of objects* variable was divided by 8, the object set size, while the *maximum time* was divided by the time window of 4min30s.

6.2.4.1.2. Familiarisation trials (both tasks)

We similarly extracted two variables of *visual* exploration from the familiarisation trials of either screen-based task: 1) the *number of objects* looked at (*breadth*), and 2) the *maximum overall time* spent on one object (*depth*). Again, these scores were standardised: the *number of objects* variable was divided by 6, the object set size, while the *maximum time* was divided by the time window of 10s.

All trials, regardless of how much infants looked away, were included to investigate intrinsic exploration strategies including looking away, a sampling decision in itself. These measures were extracted either at the start of the study during the familiarisation trials of the card-matching task, or at the end of the study during the familiarisation trials of the memory task. Although the displayed objects changed between tasks, the trial structure remained the same. Trials were thus considered similar enough to represent the same measure of infants' visual exploration at two different points during the lab visit. They were compared to assess the stability of infants' visual exploration during their visit.

Additionally, measures from the screen-based tasks were also compared to the play session measures to test whether infants' exploration was stable across modalities. An average of all trials, both screen-based tasks confounded, was computed to obtain one screen-based value per measure.

6.2.4.2. Executive-function-related variables

Scores in executive functioning were extracted from various tasks of this study and related to the exploration measures just described.

6.2.4.2.1. Inhibitory control

The glitter-wand task was used to assess infants' inhibitory control during the study, as measured by their *wait duration* score in ms (N. P. Friedman et al., 2011).

6.2.4.2.2. Working memory

Two scores were extracted from the memory task, using all the clean trials at once, namely the percentage of trials for which 1) the *first AOI hit is the target*, and 2) the *most looked at AOI is the target*. These scores represent two different measures of performance, which were compared to chance level (33.33%) in one-tailed t-tests. We used these two measures following Hochmann et al.'s findings that the *most looked at AOI is the target* reflected performance better than the *first AOI hit is the target* (Hochmann et al., 2016). This finding could not be verified in Chapter 5 given the gaze-contingent design of the card-matching task which did not make it possible to use the *most looked at AOI is the target* variable, and thus was investigated in this memory task instead. We expected both scores to be higher than chance, although possibly less or non-significantly for the *first AOI hit is target* variable,

6.2.4.2.3. Questionnaires

Finally, executive function scores were also obtained from parent-report questionnaires. The Early Childhood Behavioural Questionnaire (ECBQ) was used to measure Effortful Control (EC), one of three factors measured by this questionnaire. Inhibition Control, Attentional focusing and Attentional Shifting subscales within the EC factor were anticipated as the most relevant scales, specifically related to executive functioning. Additionally, the five developmental stage scores were obtained from the Ages and Stages Questionnaire (ASQ) and summed to obtain a total developmental score (Charkaluk et al., 2017; Halbwachs et al., 2013).

6.3. Results

With this study, we aimed to characterise exploration strategies in 15-month-old infants, by relating individual measures in exploration and executive functions. Data from four tasks as well as parent-report questionnaires was put together to this aim. We will first look at the stability of infants' visual exploration over their visit to the laboratory, through the comparison of familiarisation trials from the two screen-based tasks, before comparing screen-based measures as a whole to play-based measures, this time assessing the stability of exploratory tendencies across *modalities*. Finally, we will relate exploratory measures in either modality to executive-function scores extracted from two tasks and questionnaires.

6.3.1. Comparing visual exploration between screen-based tasks

First, we investigated infants' visual exploration of stimuli during the familiarisation trials of screen-based tasks. There were two instances when such trials were presented: at the start of the study during the card-matching task, and at the end of the study during the memory task.

Two variables were extracted from each task, each measuring a different aspect of the exploratory strategy. One measure was considered a measure of the exploration's breadth: 1) *number of objects* explored (%; Figure 6.8A), while another variable was thought to reflect the depth of the exploration: 2) *maximum time* on one object (%; Figure 6.8B). The variables were extracted at the trial level, averaged per participant and task, and divided by the whole space available (i.e., the *number of objects* variable was divided by six, the object set size, while the *maximum time* was divided by the time window of 10s). This way, we obtained percentage scores of exploration, which allowed for comparisons between different designs and sets. This is important for the analysis of the play and screen-based measures in the next section and was thus initiated already in the present analysis.

For both tasks, the trial structure was the same and only the elements presented on the screen differed: shapes were presented in the card-matching task, while animals were shown during the memory task familiarisation trials. This was considered to be similar enough that familiarisation trials from both tasks could be compared as a test and re-test of infants' visual exploration during their lab visit. We verified this by comparing the groups' scores for each task in paired-samples t-

tests, before comparing individual values for each task in correlational analyses to assess the stability of individual variations. Outliers (points further than of 1.5 times the threshold for the 25% and 75% quartiles) were excluded from both analyses to avoid spurious results (see full trial set in Appendix 9 and Appendix 10).

The *maximum time* score appeared significantly non-normally distributed (Shapiro-Wilk test: W = 0.940, p = 0.003), therefore for this score, comparisons were done using a Wilcoxon signed-rank test. Overall, the scores did not appear significantly different (Table 6.3), except for a trending effect in the tasks' *maximum time* variable (Z = 1395.00, p = 0.065). Bayesian statistics were also computed to quantify the amount of evidence for the null hypothesis (the tasks' scores are similar) versus the alternative hypothesis (the tasks' scores are significantly different). Again, the non-normally distributed scores were compared using a non-parametric Wilcoxon signed-ranked test (*maximum time*), while a Student test was used for the normally distributed variable (*number of objects*). We found moderate evidence ($3 < BF_{01} < 10$) for the null hypothesis in the case of the *number of objects* variable, and anecdotal, inconclusive evidence ($1 < BF_{01} < 3$) for the null hypothesis for the *maximum time* variable. We concluded that overall, the tasks seemed to yield similar enough, comparable scores, with some variations between the two tasks which remained limited.

Variable	**Mean (S.D.) memory task**	**Mean (S.D.) card-matching task**	**Test statistic**	**Statistical significance (p)**	**Effect size**	**Bayes Factor**
Number of objects (%)	75.6 (12.8)	73.9 (17.0)	t(65) = 0.319	0.751	Cohen's d = 0.039	BF_{01} = 7.052
Maximum time (%)	33.2 (6.45)	31.7 (12.5)	**W(65) = 1460.00**	**0.065**	MRBC = **0.262**	BF_{01} = 2.590

Table 6.3: Mean (Standard Deviation, S.D.), Pearson's R value, statistical significance, effect size (Cohen's d or matched rank biserial correlation, MRBC) and Bayes Factor of the null hypothesis over the alternative hypothesis for the paired-sampled comparisons between the card-matching and the memory tasks; each row presents the results for one of the two exploratory variables. Results below the trending threshold p = 0.100 are

highlighted in bold font. Outliers were removed from these analyses (distributions including all points can be consulted in Appendix 9-Appendix 10).

Next, we asked whether the variation within each task's scores were reflective of stable tendencies to explore i.e., whether individual scores in one task were informative of individual scores in the other.

Individual values in the two variables extracted from the two tasks were compared through correlations, a positive and significant correlation indicating stability in individuals' values between tasks. We hypothesised that this would be the case for both variables extracted. Outliers were excluded to avoid spurious correlations driven by outlier data (results and plots of distributions and correlations including all points can be consulted in Appendix 9-Appendix 11; there was no substantial change in the findings with or without inclusion of the outliers).

We found a significant correlation between the two tasks' exploratory scores in terms of the *number of objects* explored (*breadth*; Pearson's R = 0.336, p = 0.006), but not for the *maximum time* spent on one object (*depth*; Pearson's R = 0.176, p = 0.157). This correlation was positive, indicating that the *breadth* (i.e., *number of objects*) with which infants explored was stable from one task to the other.

Overall, we found some evidence that inter-individual variations in visual exploration strategies were stable over infants' visit: infants who visually explored more broadly (as reflected by their *number of objects* explored) at the start of the study also tended to explore more broadly at the end. However, there was no evidence that individual exploration tendencies in terms of depth (*maximum time*) were stable from one task to the other.

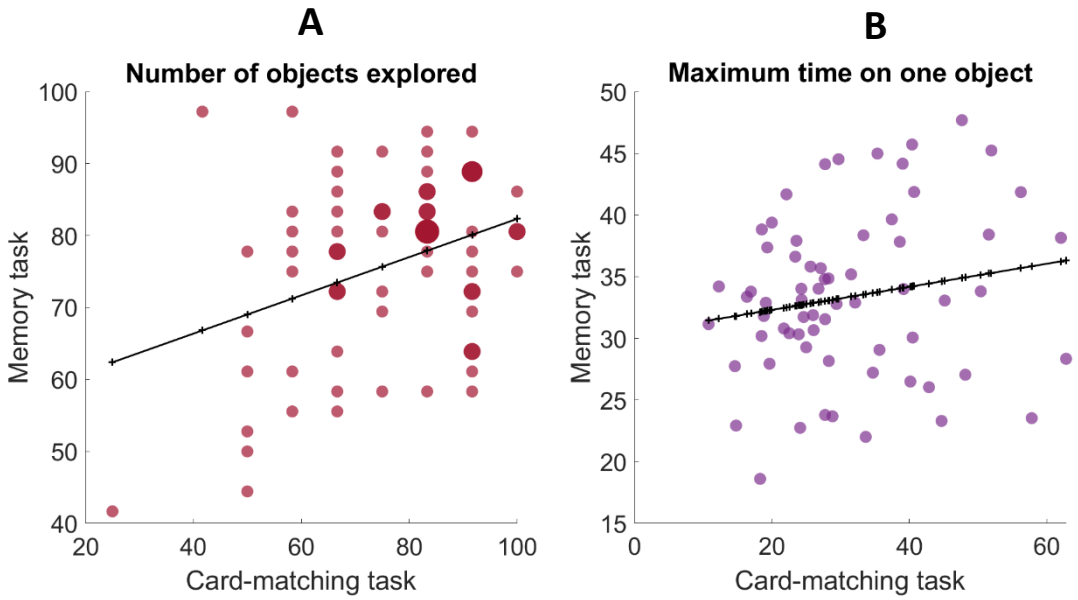


Figure 6.8: Relationship between exploratory measures of the card-matching (y-axis) and memory (x-axis) tasks. Outliers were removed from these plots (distributions and correlations including all points can be consulted in Appendix 9-Appendix 11).

Each point represents one participant's values; marker diameter is proportional to the number of individual values represented by the same point in case of overlap. The linear regression between both values is depicted as a black line on which crosses are drawn to figure the points' projections on the regression line.

A. *Number of objects* explored (%),

B. M*aximum time* on one object (%)**.**

6.3.2. Comparing visual & manual exploration between the play & screen-based tasks

Next, we compared children's exploration strategies between tasks involving different modalities this time: the screen-based tasks described in the former section, and the play task. We compared the same three measures of exploration, using an average of all screen-based trials on the one hand, and measures of exploration in the manual modality with the play task. The use of a screen-based aggregate score was motivated by the fact that although only one of the three exploratory measures showed significantly similar individual variations from one screen-based task to the other, overall Bayes' statistics revealed that the scores between the two tasks were not significantly different: we had no reason to believe that one task would yield substantially

different results compared to the other. Moreover, averaging all trials together had the advantage of increasing the signal-to-noise ratio for individual preferences common to both tasks, compared to averaging separately over smaller numbers of trials.

Again, we first compared the scores for each task in paired-samples t-tests, excluding outliers. Similarly to the previous section, we then compared scores from both modalities using through correlations (Table 6.4). One score (*number of objects*) appeared significantly non-normally distributed (Shapiro-Wilk tests: W = 0.955, p = 0.016), therefore for this score, comparisons were done using a Wilcoxon signed-rank test. Both scores appeared significantly different between tasks (Table 6.4). This was not surprising, as the screen-based and play explorations were considered as substantially different tasks, with different structures in terms of time-window and object set size. Although we computed percentage scores to be able to compare the tasks, we did not effectively expect that the group values would necessarily lead to similar values in each score. Indeed, infants were given much more time during the play task, which allowed them to explore more both in terms of breadth and depth. They explored a higher percentage of objects on average (87.0% vs. 71.8% in the screen-based tasks, see Table 6.5) and also spent more of the allocated time on the toy explored the most into depth, compared to how infants' exploration time was organised in the visual trial (respectively, 46.2% vs. 31.0%).

Variable	**Mean (S.D.) screen-based tasks**	**Mean (S.D.) play**	**Test statistic**	**Statistical significance (p)**	**Effect size**
Number of objects (%)	**71.8 (16.2)**	**87.0 (16.4)**	**W(67) = 1883.00**	**< 0.001**	**MRBC = 0.605**
Maximum time (%)	**31.0 (9.15)**	**46.2 (24.1)**	**t(68) = 4.828**	**< 0.001**	**Cohen's d = 0.559**

Table 6.4: Mean (Standard Deviation, S.D.), Pearson's R value, statistical significance, and effect size (Cohen's d or matched rank biserial correlation, MRBC) of the correlations between the card-matching and the memory tasks; each row presents the results for one of the two exploratory variables. Results below the significance threshold p = 0.05 are

highlighted in grey cells and bold font. Outliers were removed from these analyses (distributions including all points can be consulted in Appendix 12-Appendix 13).

Regardless of task differences, variations between individuals could still carry important strategic choices in exploration which could be compared between tasks, albeit different group averages, which was the next question we investigated. Similarly to the previous section, individual values in the variables extracted from each type of task were compared through correlations, a positive and significant correlation indicating stability in individuals' values between tasks. We found no significant correlation in either exploratory measure between the screen-based and the play tasks (Figure 6.9, Table 6.5). This suggested that there was no evidence for stable exploratory strategies across modalities, based on the two exploratory variables and the two contexts of visual and manual exploration that we assessed.

Outliers were removed from this analysis to avoid spurious results (plots of distributions and correlations including all points can be consulted in Appendix 12-Appendix 14). There was no substantial change based on the inclusion or exclusion of the outliers for the *number of objects*, but to note, a significant correlation ($R = 0.888$, $p < 0.001$) was found in the *maximum time* variable between each task. As reported (Table 6.4), this result did not subsist when removing the outliers, thus it was considered spurious and will not be further commented on (see plots in Appendix 13 and values in Appendix 14).

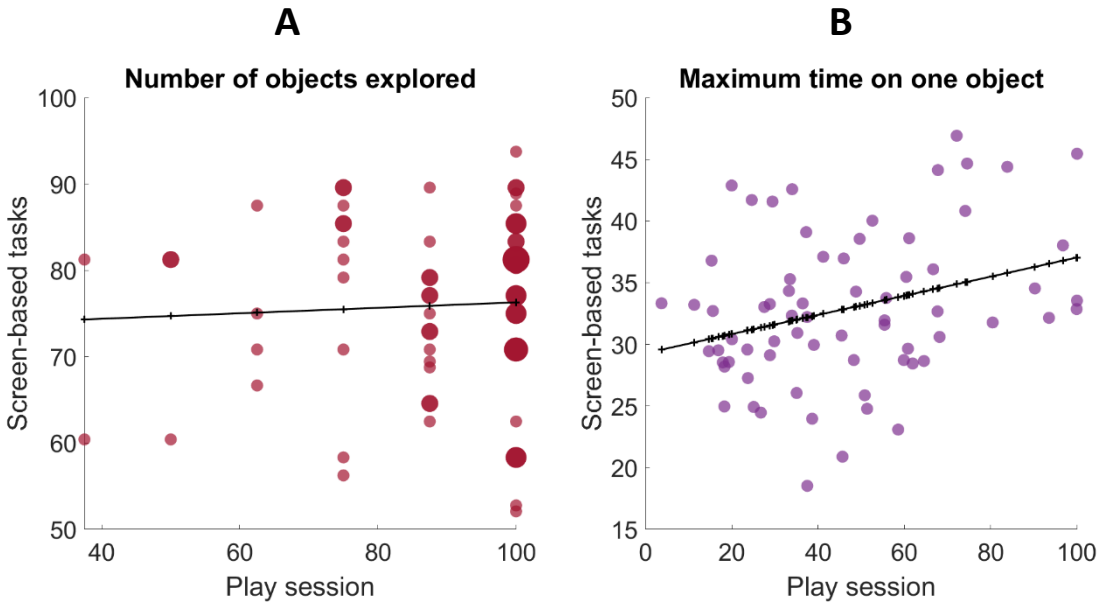


Figure 6.9: Relationship between the exploratory measures of play session (y-axis) and the screen-based tasks (x-axis). Outliers were removed from these plots (distributions and correlations including all points can be consulted in Appendix 12-Appendix 14).

(See Figure 6.8 for a detailed description of the elements of the plot.)

A. *Number of objects* explored (%),

B. *Maximum time* on one object (%).

Variable	**Mean (S.D.) screen-based tasks**	**Mean (S.D.) play**	**Correlation (Pearson's R)**	**Statistical significance (p)**
Number of objects (%)	71.8 (16.2)	87.0 (16.4)	-0.023	0.849
Maximum time (%)	31.0 (9.15)	46.2 (24.1)	0.126	0.292

Table 6.5: Mean (Standard Deviation, S.D.), Pearson's R value, and statistical significance of the correlations between exploration measures of the screen-based and the play tasks; each row presents the results for one of the two exploratory variables. Outliers were removed from these analyses (correlations including all points can be consulted in Appendix 14).

Infants did not appear to be exploring in significantly similar ways between the play and the screen-based tasks, based on our measures of exploration strategies. We decided not to compute an aggregate exploratory score by averaging the play and the screen-based measures together, but rather to look at these visual vs. manual explorations separately in the next analyses. However, because we found a significant relationship between both screen-based tasks in one of the measures of exploration, and because we did not hold strong hypotheses concerning a difference of mechanism between both screen-based tasks, which were mainly compared as the same process at two different points of the study, we kept the card-matching and memory task combined in aggregate screen-based scores.

Overall, results from our analyses comparing the two screen-based tasks, and between screen-based and play-based exploration, suggest that individual preferences in terms of exploration strategies were partly stable between the two visual tasks and not stable between modalities. There was only evidence that individual differences in breadth (*number of objects*) were stable between the two visual tasks, which validated the hypothesis that it might reflect trait differences in exploration. Thus we decided to investigate associations between exploration and executive functions using this measure only, as the other measure did appear stable enough.

6.3.3. Experimental measures of executive functions

Our next step was to investigate whether there was a link between individual differences in executive functions and any of these exploratory measures. Before diving into the analyses of this relationship, we will first describe the results from our two tasks experimentally measuring executive functions: the memory and the inhibitory task. Our aim with these tasks was not only to investigate infants' overall abilities at 15 months of age i.e., the average performance of the group, but rather to look into individual variations in the group: we were hoping for scores that were normally distributed, using a range of values across participants.

6.3.3.1.1. Measuring visual working memory

First, we looked at participants' scores in the memory task (based on Guillory & Kaldy, 2019). Two different scores were assessed, namely the percentage of trials for which 1) the *first AOI hit is target* (Figure 6.10A), and 2) the *most looked at AOI is target* (Figure 6.10B). This is similar to Hochmann et al.'s measures of performance, which were discussed in the previous chapter

(Hochmann et al., 2016). In the study that inspired the current task (Guillory & Kaldy, 2019), the measure used was the average percentage of time spent on the target AOI, relative to the time spent on any AOI. This is similar to our *most looked at AOI is target* score, although our measure is a little more conservative as it requires infants to look more at the target AOI in each trial. Still, we expected similar results as Guillory & Kaldy: we expected infants to be able to remember animals and fixate both more and more quickly on the changed animal after the blank.

Performance was assessed by comparing each score to a chance level, set at 33.3% since there were three otherwise equivalent options (Guillory & Kaldy, 2019). Both scores were found to violate the normality assumption with bi-fid distributions (Shapiro-Wilk tests, respectively W = 0.933, p = 0.003 and W = 0.957, p = 0.035; see also Figure 6.10). Thus, scores were compared to chance using a non-parametric, one-tailed Wilcoxon signed-rank test.

The *most looked at AOI is target* was found to be significantly above chance (mean = 38.7%, S.D. = 19.4%, Z = 1182.00, p = 0.025), which replicated Guillory and Kaldy's findings. They showed that 12-month-olds perform above chance when given 3 or 6s of continuous encoding time, while we confirmed that 15-month-olds also do perform above chance with 6s of continuous encoding time.

The *first AOI hit is target*, a measure that was not tested in Guillory & Kaldy's study, was found not to be significantly higher than chance (mean = 25.2%, S.D. = 19.9%, Z = 547.00, p = 0.997). A post-hoc one-tailed test revealed that the *first AOI hit is target* was actually significantly *below* chance (Z = 547.00, p = 0.003). This was unexpected and suggested that infants tended to *avoid* the new option first, following a possible familiarisation bias at the start of the trials, before shifting their gaze to the new option for an overall longer time over the whole trial. While this result goes contrary to our prediction, it does concur with the former result to show that infants do perform significantly *differently* than chance either at the start or later during the trial: whether they first avoided or later approached the new item, infants did appear to remember which items they already saw in either case.

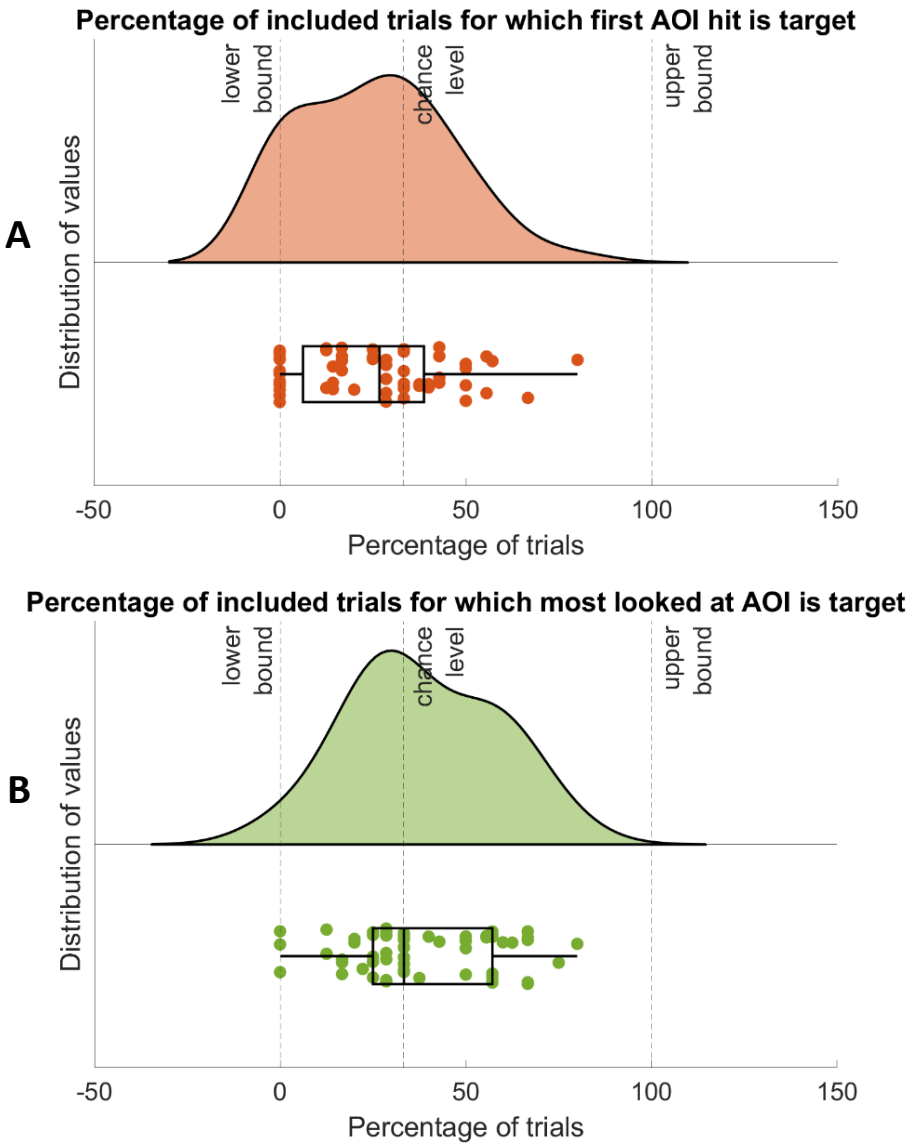


Figure 6.10: Raincloud plot for the two memory scores.

Each dot (rain) represents a participant's data point. Their average score over the whole task is plotted along the x-axis, and y-axis positions are randomised to avoid too much overlapping of the points. The group's values are depicted as a distribution curve above the points (cloud) as well as a boxplot, in which the box signifies the 25% to 75% percentiles, the line inside the box represents the median, and the whiskers depict the range, discarding any potential outlier (points outside of 1.5 times the interquartile range; non identified for this measure). Vertical dotted lines are drawn at the lower and upper bound values that constrain the measure.

A. Percentage of trials for which the *first AOI hit is target*;

B. Percentage of trials for which the *most looked at AOI is target*.

Still, the two scores appeared to show eye-movement control strategies that were going in opposite directions. Thus, we decided not to use a composite measure by averaging both scores, which would cause the opposite effects to cancel out, but instead to use them separately as two experimental measures of visual working memory in our study. As reported above, these scores were significantly diverging from a normal distribution, but the difference remained mild, and the sample was big (over 30: 59 participants included in this analysis), which, following the central limit theorem, granted that parametric tests could still be used where needed in the analyses to come. Moreover, only a trending correlation was found between the two scores (non-parametric Spearman's rho = 0.244, p = 0.060) which suggested that although both measures were linked to infants' memory, they still reflected potentially qualitatively different measures, thus we decided to use both scores separately.

6.3.3.1.2. Measuring inhibitory control

Next, we looked at infants' experimental scores in inhibitory control, as measured by their *wait duration* (0-30s) during the glitter-wand task (N. P. Friedman et al., 2011). Individual values in this score appeared distributed around two peaks (Figure 6.11): most infants did not wait and directly reached for the toy within the first few seconds (median = 2.59s, mean = 10.7s, S.D. = 12.6s), while the ones who did wait mostly waited for the whole 30s allowed, and there were very few infants with intermediate scores between these two extremes. We decided not to use the scores from this task in the further analyses because this highly binary distribution was not a helpful representation of individual differences in inhibitory control (Shapiro-Wilk tests: $W = 0.705$, $p < 0.001$).

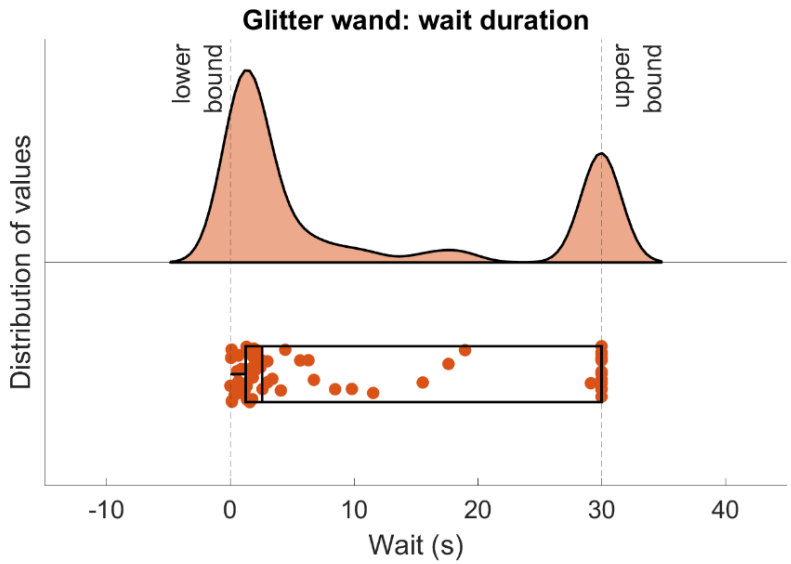


Figure 6.11: Raincloud plot of the *wait duration* (ms) measure of the glitter-wand task. (See Figure 6.10 for a detailed description of the elements of the plot.)

Contrary to working memory, which is not assessed in the ECBQ subscales, we could rely on a questionnaire-based measure for inhibitory control. The ECBQ Inhibitory control score presented values normally distributed and spread across the 7-point scale (median = 3.23, mean = 3.24, S.D. = 0.922; see Figure 6.12). Thus, the task measure was abandoned and we relied solely on the ECBQ score for a measure this function.

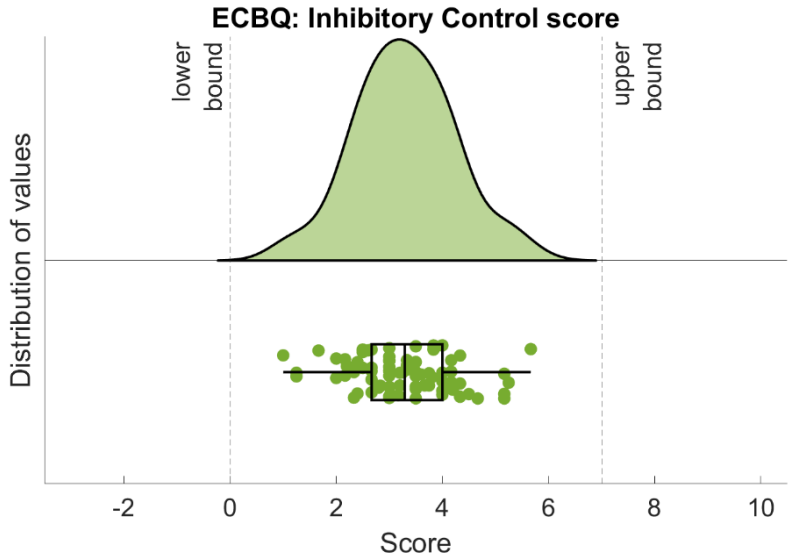


Figure 6.12: Raincloud plot for the Early Childhood Behavioural Questionnaire (ECBQ) Inhibitory Control score. (See Figure 6.10 for a detailed description of the elements of the plot.)

6.3.4. Executive functions and exploration

In our last set of analyses, we looked at the relationship between individual scores in executive functions and in exploration strategies. We remained agnostic as to how each executive function could relate to exploratory measures: as articulated in the general introduction as well as in the introduction of this chapter, executive functions, learning, and exploration are tightly intertwined and even separating the role of one function from another can prove to be difficult. Because only the *breadth* exploratory score (the *number of objects* explored) was found to be stable across the study, the *depth* score was dropped from this analysis, and only relationships between *breadth*, either in the play or the screen-based tasks, and executive functions were investigated. We performed two separate multivariate regressions on the *breadth* measure in each task-type, using seven co-variates: infants' age, their ASQ total score, the three ECBQ executive-function scores (Attention Focusing, Attention Shifting, Inhibitory Control) and the two task-based memory scores. Again, outliers were removed from these analyses.

6.3.4.1. Visual exploration breadth and executive functions

First, we looked at the exploration breadth in the screen-based task and investigated how much of this measure's variance could be explained by the covariates mentioned above. Overall, the model including the seven covariates did not significantly predict the exploration scores: $F(7,44) = 1.54$, $p = 0.180$, $R^2 = 0.197$. However, two covariates appeared to significantly and negatively predict the scores when controlling for variations in the other variables: infants' age ($\beta = -0.329$, $t(44) = -2.11$, $p = 0.040$) and their memory score as measured by the *1st AOI hit is target* variable ($\beta = -0.376$, $t(44) = -2.408$, $p = 0.020$). The rest of the regression coefficients and statistics can be found in Table 6.6. However, a model including only these two variables remained too limited to significantly predict the exploration scores: $F(2,56) = 2.299$, $p = 0.110$, $R^2 = 0.076$. The two variables also appeared uncorrelated (Pearson's $R = -0.013$, $p = 0.923$).

Variable	Standardised regression coefficient (β)	Test statistic (t)	Statistical signifi-cance (p)	Collinearity (tolerance)
Intercept (null hypothesis)	-	63.717	< .001	-
Intercept (alternative hypothesis)	-	3.609	< .001	-
Age	-0.329	-2.113	**0.040**	0.752
ASQ total	0.059	0.401	0.690	0.850
ECBQ: Attentional Focusing	0.121	0.868	0.390	0.947
ECBQ: Attentional Shifting	-0.228	-1.510	0.138	0.803
ECBQ: Inhibitory Control	-0.100	-0.660	0.512	0.800
Memory: *most looked at AOI is target*	-0.002	-0.013	0.990	0.834
Memory: *1st AOI hit is target*	-0.376	-2.408	**0.020**	0.747

Table 6.6: Standardised regression coefficients (β), test statistic (t), statistical significance (p) and collinearity statistic (tolerance) for each covariate. All covariates had tolerance values well above the 0.1 tolerance threshold, which ruled out collinearity. Significant results (Apart from the intercepts) are highlighted in grey cells and bold font.

6.3.4.2. Manual exploration breadth and executive functions

Finally, the same analysis was performed on the measure of exploratory breadth in the play session this time. The overall model did not succeed to significantly predict the exploration scores: $F(7,43) = 0.862$, $p = 0.544$, $R^2 = 0.123$. None of the covariates appeared to significantly predict the scores when controlling for variations in the other variables. The regression coefficients and statistics can be found in Table 6.7.

Variable	Standardised regression coefficient (β)	Test statistic (t)	Statistical signifi-cance (p)	Collinearity (tolerance)
Intercept (null hypothesis)	-	41.493	< .001	-
Intercept (alternative hypothesis)	-	-0.153	0.879	-
Age	0.187	1.122	0.268	0.738
ASQ total	0.002	0.014	0.989	0.858
ECBQ: Attentional Focusing	0.108	0.737	0.465	0.950
ECBQ: Attentional Shifting	-0.020	-0.123	0.903	0.791
ECBQ: Inhibitory Control	-0.105	-0.648	0.520	0.776
Memory: *most looked at AOI is target*	-0.123	-0.791	0.433	0.846
Memory: *1st AOI hit is target*	-0.198	-1.182	0.244	0.728

Table 6.7: Standardised regression coefficients (β), test statistic (t), statistical significance (p) collinearity statistic (tolerance) for each covariate. All covariates had tolerance values well above the 0.1 tolerance threshold, which ruled out collinearity.

Overall, this analysis of how executive functions and exploration breadth might relate was largely inconclusive. In the manual modality, we did not find any evidence for this. In the visual modality, we found weak evidence that age as well as memory, as measured by one of our two scores only (the *1st AOI hit is target* measure), might negatively influence exploration in the visual domain, but the relationship was only significant when controlling for all the other variables.

6.4. Discussion

6.4.1. Summary of findings

With this study, we linked a variety of measures taken from different tasks with the aim to better characterise exploratory strategies and their cognitive correlates in 15-month-old infants.

6.4.1.1. Stability of visual exploration along the study

First, we looked at whether infants' visual exploration strategy at the start of the study, measured in terms of breadth and depth of exploration, predicted their exploration strategy at the end on similar trials. We compared exploratory scores in two same-structure trials (familiarisation trials) from the first screen-based task that they undertook at the start of the study (card-matching task), and the last one (memory task). The group's scores in either task appeared mostly similar, confirming that these two types of familiarisation trials were indeed alike and triggered similar overall patterns of visual exploration. Moreover, individual variations in exploratory breadth (*number of objects*), appeared significantly and positively correlated between the two tasks. This was not the case for the depth of exploration though. Overall, this analysis brought partial evidence that infants' exploration of visual objects is stable over time: individuals appear to maintain a stable level of exploration breadth across time, but not depth. This is in line with the findings from Caruso's original experiment in the context of play, which also found breadth to be the most reliable measure of exploration (Caruso, 1993).

6.4.1.2. Stability of exploration across modalities

We also investigated whether there was a link between how infants explored in the visual modality, and how they did so in the manual modality. We compared exploratory scores combining the above-mentioned screen-based tasks and scores extracted from the play task. These two types of tasks appeared to yield significantly different exploration patterns on average for all three exploration measures. This was not unexpected, as these tasks represent different exploration contexts, involving different object-sets, time-windows, and modalities indeed, to which infants were expected to adapt their exploration strategies, as was the case. Infants overall explored more of the objects and spent more of the allocated time on the toy explored the most into depth in the play session, during which they were allowed more time to explore, compared

to the visual trials. Nonetheless, individual variations within each score could still be compared between tasks to assess if individual exploration values in one modality predict individual values in the other. We found no evidence for this type of cross-modality stability of exploration with our experimental measures and tasks.

6.4.1.3. Measuring individual variations in working memory and inhibitory control

Next, we analysed the results from two executive-function tasks, such that they could be related to exploration measures in the last set of analyses. We used a parent-report questionnaire, the ECBQ, to obtain measures of Attention Focusing, Attention Shifting and Inhibitory Control. Working memory, not measured in this questionnaire, was assessed in a task adapted from Guillory & Kaldy (2019). We assessed memory performance through two scores, the *1st AOI hit is target* and the *most looked at AOI is target* scores. These measures were inspired from Hochmann et al.'s study described in the former chapter (Hochmann et al., 2016). Both scores appeared significantly different from chance. The latter score was above chance, as expected, but the former was below chance, contrary to our predictions. In fact, the *most looked at AOI is target* score appeared to measure infants' *approach* of the new stimuli over the whole trial (novelty bias), while the *1st AOI hit is target* score appeared as a measure of infants' tendency to *avoid* the stimuli at the beginning of trials (familiarity bias). This is similar to the finding that infants first show a familiarity bias when habituation durations are short, and switch to a novelty bias over long enough habituation times (Sirois & Mareschal, 2002). Here however, we show that this tendency is even visible over the course of one trial, which to our knowledge, has not been evidenced before. Because they appeared to measure different behaviours, although both linked to memory, the two scores were used as two separate measures of memory.

We also attempted to measure individual variations in Inhibitory Control through a glitter-wand task (N. P. Friedman et al., 2011), but the task was shown to yield a binary-distributed score unfit for the study of individual variations. Hence, we discarded this measure and only relied on the questionnaire scores for measuring individuals' Inhibitory Control.

6.4.1.4. Relationship between exploration strategies and executive functions

Finally, we looked at the relationship between infants' exploration and their executive function scores. We used a multivariate regression analysis to evaluate the explanatory power of seven

covariates on infants' breadth of exploration: one overall developmental score from the ASQ, three executive function scores from the ECBQ, two task-based measures of memory as well as infants' age. Only the *breadth* measure of exploration, which appeared stable across the study, was used in this analysis. We used both the *breadth* measures, extracted from the visual screen-based tasks as well as from the manual play, in two separate regressions.

We found no evidence for a link between the investigated covariates and exploratory breadth in the manual task. In the visual task, we found weak evidence that age as well as one measure of memory (*1st AOI hit is target*), might negatively influence exploratory *breadth*, but only when controlling for all the other variables. To note, the memory score reflected infants' ability to avoid new stimuli at first, thus a lower score (more avoidance) reflects better memory. This suggests that older infants might explore less broadly, while infants who scored higher in memory might explore more broadly.

6.4.2. Limitations and future avenues

A central issue in this study was to measure individual variations in how infants explore objects, either visually or manually, that can reflect individual strategies for exploration. We developed innovative measures to investigate two important aspects of exploration, its breadth and depth, that allowed for comparisons between tasks involving different designs and exploration modalities. To our knowledge, this comparison was never investigated before. We compared exploration measures between two sets of similar visual trials at two different points in the study, and between the visual and manual modalities. This was helpful to assess the stability of the individual variations and establish the relevance of the measures, before looking into their drivers. However, we only found partial evidence that infants' visual exploration was stable between the two visual tasks, and no evidence when comparing the visual and manual modalities. This appears as a caveat, and several propositions are made below that could help overcome this issue in future work.

6.4.2.1. Power

It is worth noting that the study of individual variations typically necessitates larger, more powerful samples than the investigation of comparisons across groups or conditions. We used a relatively large sample (N = 76) for our study, which was deemed sufficient for our aim to explore

these news ideas. However, further research on larger, more powerful samples is needed to go beyond this exploratory phase: the lack of evidence for stable exploration measures for a large part of the comparisons performed did not refute our hypothesis that infants use stable exploration strategies. It is possible that infants' fatigue over time affected their exploration and introduced individual differences in fatigue that masked underlying stable individual differences in exploration patterns. Studying a bigger sample could provide enough power to investigate this question further.

6.4.2.2. Task design differences between modalities

Although our measures of exploration were similar across modalities, it is possible that the designs remained too different for individual strategic preferences to manifest in similar ways across these contexts. Notably, the tasks involved different time frames (10s in screen-based trials vs. 4.5min of the play task) and sets of objects (6 visual items vs. 8 toys), which might have hindered our ability to compare them despite the use of standardised measures. In particular, in the play session, almost half of the infants played with all the toys (35 out of 72), which might have created a ceiling effect masking individual differences in this measure. Future research involving tasks structures more similar between modalities would be helpful to investigate this further, notably closer exploration durations.

6.4.2.3. Additional measures

It is also possible that exploration depth and breadth are not best reflected by our measures and might necessitate to be described along dimensions specific to the modalities. Indeed, Caruso's original measures in the context of manual play relied on a quantification of the different actions taken by the infants, rather than merely the number of items and the duration of the manipulations. We opted for a simpler measure that was directly applicable to both contexts, but this might have been too reductive in the case of the play task, which allows for a richer exploration that might be better characterised by more complex action-related measures. Further work using an action-based measure of exploration could help to clarify this. In fact, important aspects of infants' behaviour were left undescribed by our measures in both modalities. An interesting avenue to take these results further would be to use additional measures for a better characterisation of infants' exploration.

6.4.2.3.1. For the visual modality

For example, in the visual modality, eye-movement control variables such fixations durations (Wass et al., 2015) or returns to previously visited location (Gliga et al., 2018), could be informative measures of how infants visually explore. Additionally, assessing the stability of infants' visual exploration could be taken further by using a longitudinal design to investigate longer time scales, rather than variations within one visit to the laboratory only.

6.4.2.3.2. For the manual modality

In the manual modality, the identification of action-types taken by the infants when they manipulate the objects might also give more information on how they are exploring. For example, differentiating between sensory actions (stroking, banging, etc.), functional actions (rolling a car, eating pizza, etc.) and combinatory actions (fitting a toy in another toy, stacking them, etc.) would inform on differences in manipulation complexity regardless of duration, something that is currently overlooked in our analysis. The use of motor tracking can also help to precisely quantify and localise the movements taken. Such measures were modality-specific, which was beyond the scope of this study. However, they would be helpful to use in future research for the separate analysis of exploratory behaviours in each modality. Another promising opportunity would be to use head-mounted eye-tracking. This technique does not necessitate infants to remain in one position or look towards a particular direction for their gaze to be tracked. Its use could help to analyse visual-motor interplay during a play task. Additionally, a longitudinal design would also add to the results by offering a chance to assess manual exploration's stability between visits.

6.5. Conclusion

Overall, with this study we found some evidence that infants explore strategically: they maintain similar levels of exploration breadth between two visual tasks in the study. However, this was not evidenced in terms of depth of exploration, neither between modalities. Moreover, their age and memory ability appeared to predict their breadth of exploration, when controlling for the other covariates (developmental stage, ECBQ scores and second memory score). These findings could be strengthened by an investigation of these questions with a larger sample of participants to increase statistical power, as well as the investigation of more complex measures of infants' exploration both in the visual and the manual modalities.

Chapter 7: General discussion and conclusion

This thesis employed different approaches to the investigation of how learners actively sample information according to their own individual circumstances. The topic was investigated with three studies looking at infants as well as adults. The studies investigated intrinsic individual differences as well as experimentally induced ones, visual and manual exploration, using a variety of techniques, namely eye-tracking, EEG brain imaging, and behavioural observations. The results from these studies brought a rich collection of evidence pointing to different processes and drivers for how individuals sample and process information. The key findings and their implications are reviewed in this chapter, first by summarising the main results, secondly by presenting the studies' main methodological contributions, and finally by discussing the work's theoretical contributions to the field, followed by a general conclusion.

7.1. Summary of findings

7.1.1. Randomness impairs specific priors building and priors' influence on visual sampling

In Chapter 2, we looked at the influence of priors on adults' visual information sampling using eye-tracking in the context of a basic object recognition task. Building on previous work using the *Dots* method (Moca et al., 2011), we showed that the behaviour of participants who went through a structured (*Ascending* or *Descending*) paradigm could be explained by a simple model. This model spelled out how priors, incrementally built based on previous access to information, interacted with the level of accessibility to current information (*g*) in order to guide exploration and recognition performance. The model was able to explain even small changes in the structured participants' behavioural responses such as their accuracy, eye-movement control variables such as their fixation number or spread, as well as precise measures of fixation locations' information-content that the *Dots* paradigm enables to reconstruct at each fixation location.

However, this model failed to explain participants' performance in a newly added randomised paradigm. Indeed, when going through this paradigm, participants did not appear to make use of the priors that they were expected to be building given their intermediate information access. Comparing this *Random* group to a theoretical naïve group derived from the model, and to the *Ascending* and *Descending* structured groups at low, middle and high information accessibility (*g*), we showed that the added randomness in the task structure for this group appeared to selectively impair the building and use of object-specific but not task-general priors. Indeed, *Random* participants appeared to improve over the first two blocks only rather than over the whole task, indicating that they did not build priors continuously as would be expected for object-specific priors. Coupled with the fact that they seemed to explore locations generally optimal (comparable to the group with the strongest priors and well above the group with no priors at g = 0.00) yet not containing much of the specific information about the objects' contour, this indicated that they only guided their exploration with general knowledge on the task statistics rather than on each object's identity. This impairment was evidenced despite the randomisation of the stimuli being applied to a dimension orthogonal to the task's goal.

These results, although not anticipated by our original model, ultimately concur with literature showing that randomness generally impairs performance (Barnes & Jones, 2000; Bendixen, 2014; e.g., Correa & Nobre, 2008; Rohenkohl et al., 2012), not only in visual tasks but across modalities. Some authors have suggested that this is due to the higher attentional capture of regular stimuli compared to random stimuli, which facilitates processing of stimuli related to the regular ones even along dimensions uninformative for task goals (Turk-Browne et al., 2009; Zhao et al., 2013). Southwell and colleagues (2017) have questioned this view, arguing that it could not explain why target detection in a predictable information stream was comparable to target detection in a random stream when both streams of information were tracked simultaneously. They proposed instead that randomness increases distractibility because it is more computationally demanding to track and integrate this information compared to predictable events. Our results add to this debate by showing that the task's randomness did not impact all types of information equally. Indeed, it only appeared to affect object-specific but not task-general prior building, at odds with the hypothesis of a general attentional capture with regular but not random stimuli and rather supporting the cognitive load hypothesis. Overall, these results helped to identify two factors that

guide adults' active sampling of visual information: their priors as well as the structure through which the information is presented, which is likely linked to changes in cognitive load.

7.1.2. 10-month-old infants' engagement with visual stimuli is linked to their gamma activity's strength and alignment in response to distracting stimuli

In Chapter 3, we investigated how high-frequency gamma-range activity relates to information processing and selection. We used a visual paradigm combining catchy ongoing information (a cartoon video) with simple disruptive information (black and white checkerboards) known to elicit a range of interest in 10-month-old infants (Piccardi et al., 2020). Infants' time-on-task was used as a measure of their general engagement with the content, reflecting a basic sampling decision: their willingness to keep sampling or not the information from the task. Participants were divided into two post-hoc groups based on their time-on-task variable: those who *did all* the task and remained engaged for the whole duration of the task (6min30s), and those who *stopped* engaging with the content before the end and stayed for a shorter time (2min-5min15s).

We showed that infants who *did all,* overall, responded less strongly to the distractors than infants who *stopped*. Moreover, we found that with increasing stimulus repetition, participants appeared to block the stimuli, as shown by a decreased response. In the group who *did all*, this could be explained by a change of their gamma response alignment relative to a slower rhythm, the ERP. They went from a more aligned state in stage 1 (favouring the amplification of the sensory signal from the distractor) to a more de-aligned state in stage 2 (blocking the transmission of the distractor's signal) and returned to a baseline ground-level alignment in the last stage (possibly due to a fatigue effect). In contrast, participants from the group who *stopped* maintained a stable alignment level all throughout. Gamma strength and alignment seemed to reflect two different aspects of sensory processing in infants. While the strength was found to relate to infants' overall engagement (time-on-task), the alignment appeared to be linked to infants' active information selection to amplify or block the distractors' signal over time, which was only present in infants who engaged more and stayed for the whole task.

These results suggest that gamma rhythms, although rarely studied in infants, hold important information on how infants process information: individual differences in gamma activity in the brain were linked to individual differences in behaviour, namely their engagement as measured

by their time-on-task. Moreover, we showed the existence of couplings between infants' fast and slow rhythms, which had never been evidenced in infants before. In the adult and the animal brain, gamma rhythms have been shown to be central to local information processing, and their coupling with slower rhythms was proposed as a mechanism for information selection and gating (Buzsáki & Wang, 2012; Fries, 2009; Uhlhaas et al., 2010). We show that this complex multi-frequency machinery already exists in 10-month-olds infants, opening new avenues in infant research both in terms of methodology and theory.

7.1.3. 10-month-old infants' Attention Focusing (AF) and sensory Low Threshold (LT) scores are linked to their engagement (AF & LT) and gamma response (LT only) to distracting visual stimuli

In Chapter 4, we took these results further and investigated the link between Chapter 3's evidenced individual differences in infants' engagement and brain response to distractors on the one hand, and individual differences in infants' trait sensory processing and executive function on the other hand. These trait measures were obtained via parent-reported questionnaires which were filled in both at the time of the EEG recording (for sensory processing only) and at a 7-month follow-up (for both). We used the four Infant-Toddler Sensory Profile (ITSP) scores to investigate sensory processing and three of the Early Childhood Behavior Questionnaire (ECBQ) Effortful Control subscales for the executive function measures.

We found that infants' engagement with the stimuli (time-on-task) significantly correlated with the two sensory processing Low Threshold (Sensory Sensitivity and Sensation Avoiding) scores at test but not at follow-up, and with the Attention Focusing subscale at follow-up. These results suggest that infants who were generally less reactive to (less sensitive to and avoidant of) sensory stimuli at test were also engaging with the stimuli more, but this did not appear to remain the case 7-months later. Still, at the 7-months follow-up point, we found a link between infants' attentional skills and their earlier engagement with the task: infants who exhibited stronger attentional focusing skills appeared to be the ones who were engaging more 7 months earlier. This suggests robust individual differences that remained relevant 7-months after test, albeit according to a different measure than at test. The lack of an attentional measure at test does not enable us to speculate on possible developmental changes regarding this scale.

Moreover, we also found a link between infants' brain responses and their trait Low Threshold sensory processing at test, but not their Attentional Focusing skills at follow-up. This correlation was significant both in terms of gamma activity's strength and alignment. Strength appeared potentially more linked to the Sensation Avoiding while alignment seemed more linked to the Sensory Sensitivity, but more statistical power is needed to validate these subscale-specific relationships.

Again, these results point to a powerful role of gamma activity for infants' sensory processing: after showing its relationship with infants' behaviour (engagement) in the lab, we evidenced a link with their trait behaviour in their daily life, which consolidates the findings. Crucially, by showing evidence of strong correlations with trait measures, we validated here again the relevance of studying not only the strength but also the timing relative to slower rhythms of infants' fast gamma responses. This is particularly important given that studies of gamma activity in infants are rare and couplings with slower rhythms have not been evidenced before. This echoes the adult and animal literature's growing interest in these fast rhythms and opens new possibilities for the field of infant research. Moreover, although we did not find evidence for the expected link with attentional processes, the relationship between gamma activity and sensory processing supports current theories of gamma activity's central role for sensory processing (Fries, 2015). Indeed, the influential communication by coherence hypothesis (Fries, 2005, 2015) suggests that synchronised activity between fast and slow rhythms underly information exchange and selection mechanisms. Gamma rhythms are thought to hold local sensory information in sensory cortices, while slow rhythms are thought to regulate the transmission of this local information according to top-down processes. Changes in fast-slow synchronisation have thus been proposed as a mechanism for selective attention (Buzsáki & Wang, 2012), although this relationship was not significantly evidenced here. More work is needed to disentangle the role and development of attentional processes with regards to gamma rhythms, however, our results suggest that gamma rhythms and their alignment are already central to sensory processing in 10-month-old infants.

7.1.4. Learning-progress appears sufficient but not necessary to drive engagement in 15-month-old infants learning a matching rule

In Chapter 5, we tested the learning-progress hypothesis (Kaplan & Oudeyer, 2007a) in 15-month-old infants using a matching-rule learning task (based on Hochmann et al., 2016; and Kaldy et al., 2015) coupled with eye-tracking technology. In each trial, participants were sequentially presented with 3 shapes followed by a fourth central shape to be matched with one of the three peripheral options by sustaining gaze on it to trigger a reward. Participants' performance in this task was measured according to two scores (percentage of trials for which *first AOI hit is target* or *gaze hits target*) and four continuous variables (the total *time*, *number of fixations*, *distance* and *speed* to reach the target).

We found some evidence for a corollary of the learning-progress hypothesis, which proposes a bell-shaped association between performance and engagement in a learning task. This was the case for one of the variables only (*speed*) as well as partially for another one (*gaze hits target*). It is possible that a heterogeneity of behaviours within the group might have prevented the study of such changes. Additionally, we found evidence that performance change over the first three trials was negatively correlated with infants' engagement (number of completed trials) for all four continuous performance variables. Indeed, participants who progressed and improved performance, by decreasing the amount of time needed to reach the target for example, also tended to stay on this task for longer. These results support the learning-progress hypothesis that infants' engagement is driven by their learning-progress.

We found important heterogeneity in the relationship between participants' performance and their engagement. Participants who engaged for a low number of trials all appeared to worsen or remain stable in performance over the first three trials. In contrast, participants who engaged for a high number of trials tended to improve performance over the first three trials, but there were a few cases of infants with worsened performance who still engaged for many trials. This suggests that learning-progress might be sufficient for engagement but not necessary. Indeed, it was absent in the case of participants who disengaged while it was often but not always present in the case of participants who engaged for longer. In the latter case, this might be explained by other drivers influencing engagement. It is also possible that infants' monitoring of their performance was noisy and did not map onto our measures, or that they switched strategies

between trials, tracking improvements in different variables on different trials, while we tracked one variable at a time in our analysis. Alternatively, this could be due to individual variation in how much learning-progress infants' considered sufficient.

To note, there was one variable which did not follow this performance improvement trend: as for the other performance variables, we found a negative correlation between the *speed* variable and infants' engagement, however for this variable, a decreased value indicated reduced performance. This variable, although similar to the other continuous variables, was a measure of how *fast* rather than how *much* infants processed the information on their path to the target. It appeared to be a more ambiguous measure of performance, possibly reflecting other concepts such as processing speed, engagement and motivation to process the information, or confidence and task-general prior accumulation.

Overall, this study brought the first piece of evidence supporting the learning-progress hypothesis in infants. It also offered a more nuanced view of the theory's spectrum. We propose that either 1) the concept of learning-progress, although sufficient to drive engagement, might not always be necessary or 2) learning-progress might be tracked along different measures, with switches from trial to trial. These findings contribute to a better understanding of learning in infancy. Additionally, they provide evidence for what constitutes the basic bricks of active learning in infancy, which can inform theories for how humans in general actively learn, and inspire algorithms in artificial intelligence, which increasingly strive to implement more autonomous information selection processes (Ren et al., 2021).

7.1.5. Visual exploration breadth is linked to memory abilities, and visual and manual exploration were not evidenced to be guided by overarching individual strategies for exploration

Finally, in Chapter 6, we asked whether individual differences in 15-month-old infants' exploratory behaviour were representative of general strategies for exploration which could be found across tasks. We characterised infants' exploration strategies according to two dimensions (similarly to Caruso, 1993): exploratory breadth (*number of objects* explored) and depth (*maximum time* on one object). These measures had the advantage of being applicable to a variety of experimental designs and to allow for comparisons between tasks. We investigated the

stability of these measures in the visual domain by comparing free-viewing visual exploration in two near-identical sets of trials at the start and end of the study. We also tested the stability of these measures between modalities by comparing exploration in the visual trials and in a free-play task. Finally, we looked at the relationship between individuals' executive function abilities and their exploration strategies. We investigated Visual Working Memory as measured through an eye-tracking memory task (Guillory & Kaldy, 2019), as well as Attention Focusing, Attention Shifting and Inhibitory Control measured by a parent-reported questionnaire (ECBQ). We also attempted to measure Inhibitory Control using a glitter-wand task (N. P. Friedman et al., 2011), but we dropped this measure as it appeared highly binary and unsuited for the study of individual differences as a continuum. Indeed, infants either reached for the wand right away or inhibited this behaviour for the total task duration (30s), with few in-between values.

The results from the memory task only partly followed our prediction. When presented with three items, one of which was changed to a new item after a blank, infants tended to look at the new item for significantly longer than the others over 6s (*most looked at AOI is target*), replicating findings from the original study (Guillory & Kaldy, 2019). However, infants' first fixations were significantly less likely to land on the new item than on the other options (*first AOI hit is target*), which was unexpected and had not been investigated in the original study. These results suggested a switch from a familiarisation bias at the start of trial to a novelty bias over the rest of the trial, something that has not been evidenced within trials to our knowledge but which resembles infants' switch from a familiarity to a novelty bias in short habituation times compared to longer habituation times (Sirois & Mareschal, 2002).

Our analyses of infants' exploration showed a significant correlation between the two sets of visual exploration trials in terms of exploratory breadth but not depth, suggesting that the variations in our measure of breadth were stable within the visual modality. We did not find evidence for cross-modality stability between the manual and the visual modality based on either the depth or the breadth measures. Only visual exploratory breadth was investigated for relationships with executive functioning, as it was the only measure that appeared stable. We found a significant negative correlation between exploratory breadth and age, as well as memory (*first AOI hit is target*). This suggests that younger infants as well as infants who better remembered items visually explored more objects.

Overall, in this study, we developed innovative measures of exploration adapted from previous work on infants' play (Caruso, 1993) that enabled us to use the same measure in the visual and in the manual modality to compare exploration across modalities, which has not been documented before. However, our measures appeared only partially stable within the visual modality and there was no evidence for cross-modality stability. We believe that these results do not rule out the possibility for overarching exploration strategies across modalities, given that several issues might have limited our ability to investigate this question. First, although we used a relatively large sample for an exploratory study, our statistical power remained limited to evidence individual variations in a noise-sensitive population such as infants. Second, while we maintained similar measures across modalities, the designs were different in terms of object-set sizes and even more so for time frames. Infants appeared to adapt their exploration to the design and e.g., explored more of the objects in the play task which offered more time. This might have masked variations related to individual preferences. Finally, the standardised measures might have been too reductive to encompass the complexity of exploratory strategies in both modalities, and we suggest that eye-movement control variables on the one hand and action-related measures on the other hand might be needed to better describe exploration strategies in the visual and manual modality, respectively.

7.2. Methodological contributions

The work in this thesis contributed to enrich the range of methods available for the study of information sampling both in infants and adults, where no suitable method was available to answer open questions of the field. We will discuss this work's contributions in terms of novel measures first, and in terms of novel or further validated paradigms next.

7.2.1. Novel measures for the study of information sampling

7.2.1.1. Measuring infants' brain response to visual stimuli: gamma amplitude and gamma-ERP alignment

High-frequency gamma rhythms are rarely studied in infants, although studies have shown that they are important for infants' visual processing and learning (Csibra et al., 2000; Kaufman et al., 2003; Pomiechowska & Gliga, 2021; K. A. Snyder & Keil, 2008). In research involving adults or animal models, these fast oscillations are gaining increasing interest, in particular the study of their relative timings compared to slower rhythms. The brain's ability to generate rhythms at different frequencies appears central for the processing of information at different levels (Buzsáki & Wang, 2012; Fries, 2005, 2015; Singer, 1999). Fast oscillations are thought to convey mainly low-level information involving local networks e.g., specialised areas of the occipital cortex for vision. These oscillations are more noise-sensitive and travel less accurately to remote locations, however, their fast pace enables a rich and time-resolved encoding of information as they complete multiple cycles in the same time window that a slow oscillation completes one. Slow oscillations, on the contrary, are less sensitive to noise and can travel across brain regions to convey high-level top-down information from integrative regions such as the pre-frontal cortex to specialised regions such as the sensory cortices. Such cross-frequency couplings, although they appear crucial in the adult and animal brain, have not yet been investigated in infants. This is likely due to the fact that high-frequency rhythms are particularly sensitive to noise (Moca et al., 2021), while infant populations already present increased noise-handling issues (Frank et al., 2017).

For the first time, we evidenced fast-slow cross-frequency couplings in infants which were significantly related to infants' behaviour, both in terms of experimental measure (engagement) and trait measure (sensory Low Threshold). Thus, we linked three levels of observations and

validated the relevance of studying fast and slow couplings in the infant brain. This was made possible thanks to the use of signal-enhancing steps to increase the signal-to-noise ratio in our measure of gamma activity. First, the experimental design made use of highly contrasted and regular stimuli known to elicit important brain responses in the visual cortex. Second, we used a non-standard rhythm for the study of the fast-slow couplings, namely the Event Related Potential (ERP). This signal is a potential, which is not an oscillation per se but the collapse of phase-locked brain oscillations into one signal with regular peaks roughly every hundreds of milliseconds, similarly to a slow ~10Hz oscillation. It involves the activity of large networks of neurons (Makeig et al., 2002), again similarly to slow oscillations (von Stein & Sarnthein, 2000). Because it is time-averaged, the ERP is less sensitive to noise than oscillations, however it loses information about oscillations happening at different phases, which spectral measures can account for. Thus, using the ERP was helpful to increase the robustness of our signal and evidence fast-slow couplings in the infant brain for the first time. However, because it is not a true oscillation and because the mechanisms that generate this signal are less well understood than for oscillations (Cohen, 2014, p. 54), future work investigating slow coupling with more standardly defined oscillations, such as alpha- or theta-range oscillations, would strengthen these findings. Another avenue for future research would be to extend these findings by investigating infants' fast-slow couplings in other modalities, such as touch or audition, expected to rely on the same machinery than evidenced here for vision. Additionally, looking at these processes in both younger and older infants would help to establish how and when they develop.

7.2.1.2. Measuring exploration across tasks: inconclusive standardised measures of breadth and depth

Another context in which novel measures were introduced in this thesis was the study of infants' exploration in Chapter 6. To our knowledge, no work to date looked at the relationship between measures of individuals' exploration in different modalities, even though theories of curiosity and exploration rarely discuss it per modality and seem to consider exploration as an overarching process independent of the modality. We developed standardised measures of exploration in order to compare exploration strategies across tasks. We measured exploration in terms of depth and breadth, following an approach adopted by Caruso (1993) in the context of infant play. The measures were adapted such that they would be applicable in the context of visual exploration

as well as manual exploration. Thus, while Caruso measured breadth and depth relative to the different actions performed by the infants, here we defined these measures based on the proportion of time that infants spent on each object. This approach appeared limited, and the measures showed little stability. Exploratory breadth appeared stable within the visual modality but not across the visual and manual modalities. Caruso similarly found that exploratory breadth was the most informative measure in his study. However, it is possible that time-related measures such as ours were too simplistic to describe the complexity of exploration strategies in each modality. Thus, these new measures were found inconclusive, and more work is needed to establish whether overarching strategies for exploration exist across modalities. We believe that investigations using more complete measures based on eye-movement-control and action-related variables in the visual and manual modality, respectively, would be particularly informative to answer this question.

7.2.2. Novel and further validated paradigms for the study of information sampling

7.2.2.1. New usage of the Piccardi paradigm to study gamma activity and engagement

One of the paradigms used in this thesis (Chapters 3 and 4) was developed by Piccardi et al. (2020). They studied individual differences in infants' responses to both ongoing videos and disruptive distractors. They measured infants' frontal theta activity in response to the videos and occipital ERP P1 peaks in response to the distractors, and compared the measures along four stages of the experiment. Piccardi et al. (2020) found that both theta and P1 amplitude changed with stimulus repetitions. However, neither of these changes alone could predict infants' visual seeking behaviour as measured by the ITSP questionnaire and only the combination of theta and P1 changes, proposed to reflect infants' trade-off between the processing of the video and the distractors, was predictive of infants' infants' visual seeking behaviour. Here we broaden the scope of this paradigm and show that it can also be used to measure infants' fast gamma-range responses to the distractors relative to the videos in relationship trait sensory processing.

Moreover, Piccardi et al. (2020) observed that this paradigm elicited a range of interest in the 10-month-old infants that were tested. This was considered a difficulty in their case, and they actively tried to bring infants' attention back to the task with external motivators. In the present study, we made use of this property of the paradigm: no external motivators were used and infants'

intrinsic willingness to engage with the task was used as an experimental variable. We showed the paradigm's validity to study intrinsic engagement, a measure rarely used yet highly informative of one of infants' most fundamental decisions for information sampling: when to engage or to disengage.

7.2.2.2. Model-based validation of the Dots paradigm's ability to measure precise visual information sampling in adults

We also used a paradigm developed by Moca et al. (2011) to investigate visual exploration for object recognition in a precise, quantitative manner. We used a simple model of the interaction of participants' priors (previous access to informative stimuli) and their current access to information (visibility *g*), as well as its impact on performance. This model was able to predict even small changes in Moca et al.'s (2011) participants' exploration and recognition of the stimuli. While we did not modify this paradigm, we validated the precision with which it enables to quantify information content at each fixation, in a way that followed the predictions of our simple model. In particular, being able to extract both the *physical information* (local dots displacement or LDD) and the *hidden information* (local contour density or LCD) at each fixation appeared as a powerful tool, since it helped us disentangle the selective impairment of specific and not general priors building in a new group of *Random* participants. The *Dots* paradigm offers a unique possibility to quantify information at each fixation, which is particularly informative when investigating why and how observers sample visual information. We hope that this further validation of the paradigm's precision will inspire its use in future studies, such as Suzuki et al.'s (2018) investigation of insight and implicit object recognition using *Dots* videos.

7.2.2.3. New paradigm to study learning-progress and engagement in infants

Finally, we developed a new paradigm based on Hochmann et al. (2016) and Kaldy et al. (2015) to track infants' learning progress in a matching-rule learning task. This is an important methodological advancement because learning progress is rarely studied in infants compared to overall scores. This paradigm enabled us to demonstrate the relationship between learning-progress and engagement infants, bringing the first piece of evidence for the learning progress hypothesis in infants. However, the design of this task came with some limitations, the main one being that 15-month-old infants' performance did not appear to be above chance on average, and there appeared to be some heterogeneity in infants' ability to learn in this task. This task is a

first step towards a more refined paradigm. Current limitations could be overcome in future versions e.g., by decreasing the difficulty (for example, by reducing the number of matching options to two, using shapes of different colours or removing the gaze-contingency in the outcome to simplify the task structure). Future research refining the task to reduce its difficulty would be helpful. It would help to reduce the heterogeneity in the group and increase the proportion of participants learning and progressing on this task, hereby potentially helping to uncover changes over the course of the experiment that might have been masked in the current version. Several options have been proposed: reducing the number of matching options to two, using shapes of different colours or removing the gaze-contingency in the outcome to simplify the task structure.

Taken together, this thesis validated pre-existing paradigms as well as developed new methods where the current possibilities were limited, hereby helping to better understand active information sampling both in infants and adults.

7.3. Theoretical contributions

Finally, the work of this thesis contributed to the refinement of current theories of how adults and infants learn and sample information. We will review the two main topics introduced in the General Introduction Chapter of this thesis in light of our results: theories of active learning and the role of basic bricks for active learning. Finally, we will discuss the question of whether active learning in infants is comparable to active learning in adults.

7.3.1. Active learning theories

In this thesis' chapter 5, we tested one of the most recent theories about active learning: the learning-progress hypothesis (Kaplan & Oudeyer, 2007a). For the first time, we provided evidence for this theory in infants. All of our four continuous measures of performance significantly correlated with infants' engagement, but one measure (*speed* to reach target) did not follow the expected direction: decreased performance in terms of *speed* over the first trials was linked to more engagement. The ambiguity of this measure was discussed, with possibilities that it might rather reflect other aspects of infants' behaviour during the task, such as motivation or confidence. However, for the other three measures, we showed strong correlations with engagement, which were in the direction that was predicted by the learning-progress hypothesis: infants who progressed more in the start of the study also engaged with the task for longer. This suggests that learning-progress is a strong drive, already present in infancy. It opens new avenues for research in infancy, but also confirms the idea that the learning-progress drive is a core aspect of cognition that is available to infants in order for them to guide their behaviour and learn best from their environment. Future work looking at other ages throughout development would be particularly useful, especially work with younger infants to uncover when this ability arises.

One corollary of the learning-progress hypothesis is that when learners have reached their peak performance, their progress plateaus as they cannot learn more, and thus their drive to engage declines. Hence, following the learning-progress theory, performance when engaging with and learning from a source of information is expected to follow a bell curve: learners first increasingly progress as they learn more about the information at hand, before reaching a plateauing phase, and dropping towards the end as they lose interest and finally disengage. We found limited evidence for this in our task. Only one measure (*speed*) followed this trend, while another one

only appeared to decrease at the end. It is possible that infants were tracking their performance along various measures at once during the task, and for example switched measure from trial to trial, which would not show when analysing the measures separately. It is also possible that infants within the sample tested might have adopted different behaviours and that this heterogeneity masked the effects overall. For example, some infants were shown to worsen in performance over the first trials, which was indeed linked with less engagement, as predicted by the learning-progress hypothesis, but which we did not expect in this study. These infants adopting a strategy to avoid the stimuli as they are not making progress, their performance changes might have cancelled out effects shown in infants who on the contrary approached the stimuli as they were progressing. Finally, it is also possible that the three separate phases of learning-progress described above did not last for equal amounts of time. Thus, our division of the experiment into three stages of equal length per infant, although it had the advantage of accommodating individual differences in overall engagement (creating comparable stages for infants with different trial numbers which were still thought to be compatible with the expected changes following the learning hypothesis), did not account for potential individual differences in how long these phases would last for in each infant. All in all, we did not clearly validate nor invalidate this bell-shape corollary.

Outstanding questions regarding both the learning-progress hypothesis and its bell-shape corollary, notably are: 1) how infants evaluate their progress and whether they can switch from tracking one measure to another from trial to trial; 2) individual differences in the plateau phase: is this phase of fixed duration, or is its duration related to the time that infants took to reach this level of performance?; 3) the same question subsists regarding the drop-out phase: are there individual differences in the time it takes infants to decide on fully disengaging once their interest and performance start dropping?; 4) are these processes impaired (and if so how) in atypical populations with learning disabilities such as individuals with Autism Spectrum Disorder (ASD) and Attention Deficit and Hyperactivity (ADHD)? Another open question that persists on the topic is the role of grit and its influence on engagement and these bell-shaped curves. Grit is defined as "*perseverance and passion for long-term goals*" (Duckworth et al., 2007) and has been found to explain inter-individual variations in success in terms of professional or educational attainment better than Intellectual Quotient (IQ) measures (Duckworth et al., 2007). It appears that it could

play an important role in this theory of how learning-progress drives engagement, but it is unclear what parameters it would influence.

Finally, we did not test other theories for active learning in this thesis, hence we cannot compare their validity. As we noted in the general introduction, there is often an overlap between these theories, which largely do not contradict each other. Indeed, implementations of the learning-progress theory with robots relied on and adapted the older theories' reinforcement-learning framework (Oudeyer et al., 2005; Sutton & Barto, 1998) and the intermediate difficulty level favoured by learning-progress theories reminds medium-level theories such as intermediate discrepancy. Still, one result in this study can aliment the debate, although not having the breadth to clearly settle it. We found that none of the infants who disengaged showed an increase in performance, suggesting that learning-progress in itself might be sufficient to drive engagement. However, infants who engaged for longer did not always progress over the start: although they were more likely to progress, some of them actually regressed over the first trials and yet remained engaged and completed many trials. We discussed above the possibility that this could be due to individual variation in how much learning progress is considered sufficient, or infants tracking different measures of performance at once, but it could also suggest that infants switched drive between trials or used different drives between individuals, although nothing in the experimental design suggests that they could have been driven by e.g., novelty or complexity. Still, the debate remains open as to whether learners always use the same drive to guide their exploration and whether one theory prevails on the others. The learning-progress theory appears as the most flexible and seems to be a core ability present early in development. However, it is possible that different learning contexts might call for different behaviours, sometimes favouring learning-progress as the most optimal drive for information sampling, and other times favouring other heuristics such as novelty or complexity might suffice. The question remains open as to whether some of these heuristics are simpler to compute in some situations, and whether one is always more optimal or not.

7.3.2. Basic bricks for active learning

7.3.2.1. Memory

Memory appeared as a strong driver for exploration in two of our studies, both with adults (Chapter 2) and with infants (Chapter 6). In Chapter 2, we showed that a model of how

participants build priors based on their previous access to information was able to explain even small changes in how they freely explored the stimuli, but that randomness in the task structure destroyed this ability to guide exploration with priors. This confirms the importance of memory processes for active information sampling, and suggests that individual variations involving memory, such as the differences in priors experimentally induced in this paradigm, can cause dramatic differences in how individuals' experience of the world since they already sample it differently and thus come to different conclusions when integrating this information. Together, these findings point to the importance of catering for different stages of learning and memory, as well as presenting information in a structured manner.

In chapter 6, memory appeared as our only measure of executive function that influenced 15-month-old infants' visual exploration in a free viewing paradigm. While these results do not rule out the possible influence of other executive function variables on visual exploration, they confirm the crucial role of memory processes. We found that infants who remembered items better in a visual working memory task explored more items in the free-viewing task. We did not formulate specific hypotheses on how exploration and executive functions would relate given the exploratory aspect of the study, and because although the relationship between the two is quite established in the literature (e.g., Radulescu et al., 2021), there is a lack of work looking at how exactly they relate. However, this relationship is unsurprising a posteriori, as working memory is central to holding information in mind, even more so for information about different items (Awh et al., 2007).

All in all, we have confirmed the importance of memory for information sampling and showed that individual differences in the former could drive individual differences in the latter. To note, we focused on short-term memory only both in Chapter 2 and 6 and we did not assess the role of long-term memory for exploration. In Chapter 2, we evoked the role of long-term memory as well as short-term memory for the acquisition of specific priors (Henderson, 2003). Moreover, long-term priors are known to play a role in information sampling as well (Leckart, 1966). Future work including measures of long-term memory might complement and refine these findings.

7.3.2.2. Attention

Throughout this thesis, we looked at active learning in different ways and often hypothesised a role of attention for information sampling. However, often we did not find evidence for this attentional involvement.

In Chapter 4, building on findings from Chapter 3, we showed that 10-month-old infants' Attentional Focusing ECBQ score was positively correlated to their time-on-task variable measuring engagement the Piccardi paradigm (2020). This relationship followed our predictions: infants who presented stronger attention skills engaged with the stimuli for longer. However, looking at infants' brain response to these stimuli, we did not find a significant link between the ECBQ's Attentional Focusing score and neither the strength nor the alignment measure of the gamma response, which however appeared significantly correlated to another measure (Low Threshold sensory processing scores form the ITSP questionnaire). This was the case despite our strong predictions for a link between the two, based on work in adults and animals showing that gamma activity and its alignment with slower rhythms appears as a central mechanism for attentional processes and information selection (Fries, 2005, 2015; Lakatos et al., 2008; Singer, 1999). This result was surprising; however, we suggest that more work with higher statistical power is needed to resolve this question. Indeed, this study was performed on a relatively small sample which suited the exploratory purpose of the study to demonstrate the general role of fast-slow oscillations in the infant brain, but might have been too limited for the study of trait individual differences. Moreover, it is possible that the age group tested was too young to be exploiting the attentional processes anticipated. Indeed, attentional processes are thought to develop during the second half of the first year of life around the age at which the infants were tested (Lawson & Ruff, 2004; Xie et al., 2018). Future work on an older age group would help to clarify this question. Thus, this study remains inconclusive regarding the link between fast oscillations and attentional processes in infants. Still, the relation with engagement which was itself linked to difference in gamma activity, is encouraging and pushes for further work on the matter.

In chapter 6, we looked at 15-month-old infants' exploration strategies in terms of breadth and depth both in the visual and the manual modality. However, only the visual breadth measure appeared stable and was thus further investigated to identify potential links with executive

functioning. Here again, no link with attention was identified although a link with memory was significantly evidenced. Similarly to the previous study, we propose that these results do not advocate that attention is not involved in visual exploration processes, but rather suggest that further work is needed with more statistical power to clearly investigate such relationships. Indeed, here again this study was exploratory and although a larger sample size was used, it did not appear sufficient for the analyses performed. Moreover, the measures of exploration appeared limited (see section 7.2.1.2 above), which likely also restricted this analysis.

Finally, although no measure of attention was recorded in Chapter 2, the interpretation of the findings in light of the literature did bring questions about a potential involvement of attentional processes. Indeed, we found that introducing randomness in the structure of our visual object recognition paradigm impaired participants' both performance and exploration of the stimuli. This has been proposed to be due to a relatively higher attentional capture for regular stimuli compared to random ones (Zhao et al., 2013). However, other work suggested that this was not the case and that performance disparities between regular and random stimuli were better explained by differences in computational load (Southwell et al., 2017). Our results show that randomness selectively impaired selective but not general priors, which is in support of the cognitive load hypothesis over a general attentional bias.

Overall, although attentional processes were often not evidenced in the studies from this thesis, statistical power was often identified as a limitation that restricted our analyses. Thus, the absence of significant results involving attentional processes in relation to active information sampling is not to be taken as a conclusion that attention is not involved in the processes at hand, but rather that our studies were inconclusive in investigating these links and that more work is needed, at the exception of Chapter 2.

7.3.2.3. Executive functions

Finally, two of our studies (Chapters 4 and 6) looked at measures of executive functions in relation to active learning. In Chapter 6, we showed that the glitter-wand task (N. P. Friedman et al., 2011) was not suited to measure individual differences at 15 months. Further executive-function-related results from these two studies were already described above in terms of working memory and attentional processes, and no other executive function was identified as having a role in these processes. Notably, the inhibitory control score did not appear to be significantly linked to either

infants' engagement with the Piccardi paradigm (2020), or their visual exploration. However, here again, caution is needed before excluding a potential role of this executive function in exploration and engagement due to statistical power limitations.

7.3.3. Adult-like active learning in infants?

Finally, this work also allows one to consider whether infants and adults learn in the same way when they actively sample information. Two of our studies (Chapter 3 and 5) were designed to look at the existence of core mechanisms that have been evidenced as important for information sampling and processing in adults. In both cases, we evidenced the existence of this machinery in infants: 10-month-old infants did appear to use complex fast-slow oscillations alignment to adapt their processing of visual stimuli (in the group who engaged for more trials) and 15-month-old infants also did adapt their engagement with a task depending on their learning-progress on this task. This suggests that infants are capable of using the same core mechanisms as adults for active learning.

Does this mean that infants and adults explore and learn in the same way? This does not seem to be the case. For example, infants have been shown to learn unusual rules more easily than children, who themselves learned better than adolescents while adults struggled the most (Gopnik et al., 2017; Lucas et al., 2014). There has also been work showing that children were more eager to explore than adults, despite this affecting them by getting less rewards (E. Schulz et al., 2019). These differences have been suggested to be driven by a change in the balance between exploration and exploitation over development, going from more exploration first which opens children to more opportunities to learn and avoid missing opportunities because of immature priors, while adults' strategy enables them to optimise their learning by relying more on their extensive knowledge (Gopnik, 2020). Thus, infants and adults do appear to learn and explore in somewhat different ways which are important to characterise. However, we show that an entirely different machinery doesn't seem necessary to explain these differences, and that state-of-the-art theories of how adult humans learn seem to apply to infants too. Future work is needed both to delineate what are the basic blocks that sustain learning processes already from infancy, but also what defines a more experienced adult learner as opposed to a more naïve infant one.

7.4. General conclusion

Throughout this thesis, different approaches were adopted for the investigation of active learning in adults and in infants. Combining different methods (EEG brain imaging, eye-tracking and behavioural observations) and paradigms, this work brought converging evidence that learners generally tailor the way that they explore and sample information to their own individual circumstances, both in terms of experimentally induced differences and intrinsic trait variations. Individual variations in terms of priors, attentional focusing, sensory processing, visual working memory or learning progress were identified as key drivers for how individuals actively learn. Notably, these studies brought novel evidence for the existence of two core mechanisms already evidenced in adults that can guide infants' learning and exploration. On the one hand, we brought the first piece of evidence that 15-month-old infants are capable of tracking their learning progress to adapt their overall engagement with stimuli, supporting the learning-progress hypothesis in infants. On the other hand, we showed for the first time that fast-slow oscillations couplings were used by the infant brain to adapt their processing of visual stimuli, in a way that was related to infants' trait sensory processing as reported by parents. We also demonstrated that adults sampling of visual information was guided by their priors, but that this ability was impaired by the introduction of randomness in the task structure. Finally, we showed that 15-month-old infants visually explored more broadly when they showed stronger visual working memory abilities. The results from these studies help to build a coherent picture of how individuals actively sample information and learn. We hope that the theoretical and methodological contribution of this work will help to advance the field's knowledge and inform future research on this topic.

Bibliography

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure & Function*, *218*(3), 611–643. https://doi.org/10.1007/s00429-012-0475-5

Ahadi, S. A., Rothbart, M. K., & Ye, R. (1993). Children's temperament in the US and China: Similarities and differences. *European Journal of Personality*, *7*(5), 359–378. https://doi.org/10.1002/per.2410070506

American Psychological Association. (2022). Cognition. In *APA Dictionary of Psychology*. https://dictionary.apa.org/cognition

American Psychological Association. (2022). Learning. In *APA Dictionary of Psychology*. https://dictionary.apa.org/learning

Amso, D., & Scerif, G. (2015). The attentive brain: Insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience*, *16*(10), 606–619. https://doi.org/10.1038/nrn4025

Andersen, P., Morris, R., Amaral, D., Bliss, T., & O'Keefe, J. (Eds.). (2006). *The Hippocampus Book*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195100273.001.0001

Anderson, J. R. (1995). *Learning and Memory: An Integrated Approach*. Wiley.

Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*(1), 48–53. https://doi.org/10.1111/j.1467-7687.2007.00563.x

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual Working Memory Represents a Fixed Number of Items Regardless of Complexity. *Psychological Science*, *18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Baddeley, A. (1986). *Working memory* (pp. xi, 289). Clarendon Press/Oxford University Press.

Bailey, C. E. (2007). Cognitive Accuracy and Intelligent Executive Function in the Brain and in Business. *Annals of the New York Academy of Sciences*, *1118*(1), 122–141. https://doi.org/10.1196/annals.1412.011

Baillargeon, R., & Graber, M. (1987). Where's the rabbit? 5.5-month-old infants' representation of the height of a hidden object. *Cognitive Development*, *2*(4), 375–392. https://doi.org/10.1016/S0885-2014(87)80014-X

Baranes, A., Oudeyer, P.-Y., & Gottlieb, J. (2015). Eye movements reveal epistemic curiosity in human observers. *Vision Research*, *117*, 81–90. https://doi.org/10.1016/j.visres.2015.10.009

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention, and Time. *Cognitive Psychology*, *41*(3), 254–311. https://doi.org/10.1006/cogp.2000.0738

Bartos, M., Vida, I., & Jonas, P. (2007). Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nature Reviews Neuroscience*, *8*(1), 45–56. https://doi.org/10.1038/nrn2044

Bauer, M., Oostenveld, R., Peeters, M., & Fries, P. (2006). Tactile Spatial Attention Enhances Gamma-Band Activity in Somatosensory Cortex and Reduces Low-Frequency Activity in Parieto-Occipital Areas. *Journal of Neuroscience*, *26*(2), 490–501. https://doi.org/10.1523/JNEUROSCI.5228-04.2006

Bendixen, A. (2014). Predictability effects in auditory scene analysis: A review. *Frontiers in Neuroscience*, *8*. https://doi.org/10.3389/fnins.2014.00060

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *British Journal of Psychology*, *41*(1), 68.

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology*, *45*(3), 180.

Berlyne, D. E. (1965). *Structure and direction in thinking*. Wiley.

Bernardo, J. M. (1979). Reference Posterior Distributions for Bayesian Inference. *Journal of the Royal Statistical Society: Series B (Methodological)*, *41*(2), 113–128. https://doi.org/10.1111/j.2517-6161.1979.tb01066.x

Blair, C., & Razza, R. P. (2007). Relating Effortful Control, Executive Function, and False Belief Understanding to Emerging Math and Literacy Ability in Kindergarten. *Child Development*, *78*(2), 647–663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Blanchard, T. C., Hayden, B. Y., & Bromberg-Martin, E. S. (2015). Orbitofrontal Cortex Uses Distinct Codes for Different Choice Attributes in Decisions Motivated by Curiosity. *Neuron*, *85*(3), 602–614. https://doi.org/10.1016/j.neuron.2014.12.050

Börgers, C., & Kopell, N. (2003). Synchronization in Networks of Excitatory and Inhibitory Neurons with Sparse, Random Connectivity. *Neural Computation*, *15*(3), 509–538. https://doi.org/10.1162/089976603321192059

Bornstein, M. H., Jager, J., & Putnick, D. L. (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. *Developmental Review*, *33*(4), 357–370. https://doi.org/10.1016/j.dr.2013.08.003

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Bragin, A., Jando, G., Nadasdy, Z., Hetke, J., Wise, K., & Buzsaki, G. (1995). Gamma (40-100 Hz) oscillation in the hippocampus of the behaving rat. *Journal of Neuroscience*, *15*(1), 47–60. https://doi.org/10.1523/JNEUROSCI.15-01-00047.1995

Brascamp, J. W., Knapen, T. H. J., Kanai, R., van Ee, R., & van den Berg, A. V. (2007). Flash suppression and flash facilitation in binocular rivalry. *Journal of Vision*, *7*(12), 12. https://doi.org/10.1167/7.12.12

Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain Dopamine Neurons Signal Preference for Advance Information about Upcoming Rewards. *Neuron*, *63*(1), 119–126. https://doi.org/10.1016/j.neuron.2009.06.009

Brown, C., Tollefson, N., Dunn, W., Cromwell, R., & Filion, D. (2001). The Adult Sensory Profile: Measuring Patterns of Sensory Processing. *American Journal of Occupational Therapy*, *55*(1), 75–82. https://doi.org/10.5014/ajot.55.1.75

Bruno, R. M., & Sakmann, B. (2006). Cortex Is Driven by Weak but Synchronously Active Thalamocortical Synapses. *Science*, *312*(5780), 1622–1627. https://doi.org/10.1126/science.1124593

Bull, R., & Scerif, G. (2001). Executive Functioning as a Predictor of Children's Mathematics Ability: Inhibition, Switching, and Working Memory. *Developmental Neuropsychology*, *19*(3), 273–293. https://doi.org/10.1207/S15326942DN1903_3

Burgess, P. W. (1997). Theory and Methodology in Executive Function Research. In P. Rabbitt (Ed.), *Methodology Of Frontal And Executive Function* (pp. 81–116). Psychology Press. https://doi.org/10.4324/9780203344187

Buswell, G. T. (1935). *How people look at pictures: A study of the psychology and perception in art* (p. 198). Univ. Chicago Press.

Butler, R. A. (1953). Discrimination learning by rhesus monkeys to visual-exploration motivation. *Journal of Comparative and Physiological Psychology*, *46*(2), 95.

Buzsáki, G. (2010). Neural Syntax: Cell Assemblies, Synapsembles, and Readers. *Neuron*, *68*(3), 362–385. https://doi.org/10.1016/j.neuron.2010.09.023

Buzsáki, G., Lai-Wo S., L., & Vanderwolf, C. H. (1983). Cellular bases of hippocampal EEG in the behaving rat. *Brain Research Reviews*, *6*(2), 139–171. https://doi.org/10.1016/0165-0173(83)90037-1

Buzsáki, G., & Wang, X.-J. (2012). Mechanisms of Gamma Oscillations. *Annual Review of Neuroscience*, *35*(1), 203–225. https://doi.org/10.1146/annurev-neuro-062111-150444

Caruso, D. A. (1993). Dimensions of quality in infants' exploratory behavior: Relationships to problem-solving ability. *Infant Behavior and Development*, *16*(4), 441–454. https://doi.org/10.1016/0163-6383(93)80003-Q

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 6–6. https://doi.org/10.1167/9.3.6

Charkaluk, M.-L., Rousseau, J., Calderon, J., Bernard, J. Y., Forhan, A., Heude, B., Kaminski, M., & Group, on behalf of the E. M.-C. C. S. (2017). Ages and Stages Questionnaire at 3 Years for Predicting IQ at 5–6 Years. *Pediatrics*, *139*(4). https://doi.org/10.1542/peds.2016-2798

Charpentier, C. J., Bromberg-Martin, E. S., & Sharot, T. (2018). Valuation of knowledge and ignorance in mesolimbic reward circuitry. *Proceedings of the National Academy of Sciences*, *115*(31), E7255–E7264. https://doi.org/10.1073/pnas.1800547115

Chentanez, N., Barto, A., & Singh, S. (2005). Intrinsically Motivated Reinforcement Learning. *Advances in Neural Information Processing Systems*, *17*. https://proceedings.neurips.cc/paper/2004/hash/4be5a36cbaca8ab9d2066debfe4e65c1-Abstract.html

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology*, *17*(2), 177–184. https://doi.org/10.1016/j.conb.2007.03.005

Ciuparu, A., & Mureşan, R. C. (2016). Sources of bias in single-trial normalization procedures. *European Journal of Neuroscience*, *43*(7), 861–869. https://doi.org/10.1111/ejn.13179

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. MIT Press.

Cohen, M. X., Elger, C. E., & Fell, J. (2008). Oscillatory Activity and Phase–Amplitude Coupling in the Human Medial Frontal Cortex during Decision Making. *Journal of Cognitive Neuroscience*, *21*(2), 390–402. https://doi.org/10.1162/jocn.2008.21020

Colgin, L. L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., Moser, M.-B., & Moser, E. I. (2009). Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature*, *462*(7271), 353–357. https://doi.org/10.1038/nature08573

Colombo, J. (2001). The Development of Visual Attention in Infancy. *Annual Review of Psychology*, *52*(1), 337–367. https://doi.org/10.1146/annurev.psych.52.1.337

Colombo, J., & Mitchell, D. W. (1990). Individual differences in early visual attention: Fixation time and information processing. In *Individual differences in infancy: Reliability, stability, prediction* (pp. 193–227). Lawrence Erlbaum Associates, Inc.

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*(3), 341–349. https://doi.org/10.1016/j.cognition.2011.03.003

Correa, Á., & Nobre, A. C. (2008). Neural Modulation by Regularity and Passage of Time. *Journal of Neurophysiology*, *100*(3), 1649–1655. https://doi.org/10.1152/jn.90656.2008

Crottaz-Herbette, S., & Menon, V. (2006). Where and When the Anterior Cingulate Cortex Modulates Attentional Response: Combined fMRI and ERP Evidence. *Journal of Cognitive Neuroscience*, *18*(5), 766–780. https://doi.org/10.1162/jocn.2006.18.5.766

Csibra, G., Davis, G., Spratling, M. W., & Johnson, M. H. (2000). Gamma Oscillations and Object Processing in the Infant Brain. *Science*, *290*(5496), 1582–1585. https://doi.org/10.1126/science.290.5496.1582

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience* (Vol. 1990). Harper & Row New York.

Cuevas, K., & Bell, M. A. (2014). Infant Attention and Early Childhood Executive Function. *Child Development*, *85*(2), 397–404. https://doi.org/10.1111/cdev.12126

Dăbâcan, A., & Mureşan, R. C. (2017). Robust Analysis of Non-stationary Cortical Responses: Tracing Variable Frequency Gamma Oscillations and Separating Multiple Component Input Modulations. In S. Vlad & N. M. Roman (Eds.), *International Conference on Advancements of Medicine and Health Care through Technology; 12th—15th October 2016, Cluj-Napoca, Romania* (pp. 189–194). Springer International Publishing. https://doi.org/10.1007/978-3-319-52875-5_42

Darchen, R. (1957). Sur le comportement d'exploration de Batella germacia. Exploration d'un plan [On the exploratory behavior of cockroaches. Exploration of a plane]. *Journal de Psychologie Normale et Pathologique*, *54*, 190–205.

Dashiell, J. F. (1925). A quantitative demonstration of animal drive. *Journal of Comparative Psychology*, *5*(3), 205–208. https://doi.org/10.1037/h0071833

De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review*, *20*(4), 631–642. https://doi.org/10.3758/s13423-013-0386-3

de Mooij, S. M. M., Raijmakers, M. E. J., Dumontheil, I., Kirkham, N. Z., & van der Maas, H. L. J. (2021). Error detection through mouse movement in an online adaptive learning environment. *Journal of Computer Assisted Learning*, *37*(1), 242–252. https://doi.org/10.1111/jcal.12483

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer Science & Business Media.

Den Ouden, H. E. M., Kok, P., & De Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Frontiers in Psychology*, *3*, 548. https://doi.org/10.3389/fpsyg.2012.00548

D'Esposito, M., Postle, B. R., Jonides, J., & Smith, E. E. (1999). The neural substrate and temporal dynamics of interference effects in working memory as revealed by event-related functional MRI. *Proceedings of the National Academy of Sciences*, *96*(13), 7514–7519. https://doi.org/10.1073/pnas.96.13.7514

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Doan, T., Castro, A., Bonawitz, E., & Denison, S. (2020). "Wow, I did it!": Unexpected success increases preschoolers' exploratory play on a later task. *Cognitive Development*, *55*, 100925. https://doi.org/10.1016/j.cogdev.2020.100925

Doebel, S. (2020). Rethinking Executive Function and Its Development. *Perspectives on Psychological Science*, *15*(4), 942–956. https://doi.org/10.1177/1745691620904771

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Dudai, Y. (1989). *The neurobiology of memory: Concepts, findings, trends* (pp. xi, 340). Oxford University Press.

Dudai, Y. (2007). Memory: It's all about representations. In *Science of Memory: Concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195310443.003.0002

Dunn, W. (1997). The Impact of Sensory Processing Abilities on the Daily Lives of Young Children and Their Families: A Conceptual Model. *Infants and Young Children*, *9*(4), 23–35.

Dunn, W., & Daniels, D. B. (2002). Initial Development of the Infant/Toddler Sensory Profile. *Journal of Early Intervention*, *25*(1), 27–41. https://doi.org/10.1177/105381510202500104

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H. J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biological Cybernetics*, *60*(2), 121–130. https://doi.org/10.1007/BF00202899

Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, *2*(10), 704–716. https://doi.org/10.1038/35094565

Engel, A. K., König, P., Kreiter, A. K., & Singer, W. (1991). Interhemispheric Synchronization of Oscillatory Neuronal Responses in Cat Visual Cortex. *Science*, *252*(5009), 1177–1179.

Eysenck, M. W. (1981). Learning, Memory and Personality. In H. J. Eysenck (Ed.), *A Model for Personality* (pp. 169–209). Springer. https://doi.org/10.1007/978-3-642-67783-0_6

Fair, D. A., Posner, J., Nagel, B. J., Bathula, D., Dias, T. G. C., Mills, K. L., Blythe, M. S., Giwa, A., Schmitt, C. F., & Nigg, J. T. (2010). Atypical Default Network Connectivity in Youth with Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *68*(12), 1084–1091. https://doi.org/10.1016/j.biopsych.2010.07.003

Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., & Posner, M. I. (2003). Cognitive and Brain Consequences of Conflict. *NeuroImage*, *18*(1), 42–57. https://doi.org/10.1006/nimg.2002.1319

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage*, *26*(2), 471–479. https://doi.org/10.1016/j.neuroimage.2005.02.004

Fastrich, G. M., Kerr, T., Castel, A. D., & Murayama, K. (2018). The role of interest in memory for trivia questions: An investigation with a large-scale database. *Motivation Science*, *4*(3), 227–250.

Feldman, J. (1993). Perceptual models of small dot clusters. In I. Cox, P. Hansen, & B. Julesz (Eds.), *Partitioning data sets* (Vol. 19, pp. 331–357). American Mathematical Society.

Feldman, J. (1997). Curvilinearity, covariance, and regularity in perceptual groups. *Vision Research*, *37*(20), 2835–2848. https://doi.org/10.1016/S0042-6989(97)00096-5

Feldman, J. (2001). Bayesian contour integration. *Perception & Psychophysics*, *63*(7), 1171–1182. https://doi.org/10.3758/BF03194532

Findlay, J. M., Findlay, C. for V. and V. C. D. of P. J. M., Findlay, J. M., Gilchrist, I. D., & Gilchrist, R. in N. D. of E. P. I. D. (2003). *Active Vision: The Psychology of Looking and Seeing*. Oxford University Press.

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738–743. https://doi.org/10.1038/nn.3689

Foulsham, T. (2015). Eye movements and their functions in everyday tasks. *Eye*, *29*(2), 196–199. https://doi.org/10.1038/eye.2014.275

Fowler, H. (1965). *Curisosity and exploratory behavior*. The Macmillan Company.

Foxe, J., & Snyder, A. (2011). The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism during Selective Attention. *Frontiers in Psychology*, *2*. https://www.frontiersin.org/article/10.3389/fpsyg.2011.00154

Frank, M. C., Bergelson, E., Bergmann, C., Cristia, A., Floccia, C., Gervain, J., Hamlin, J. K., Hannon, E. E., Kline, M., Levelt, C., Lew-Williams, C., Nazzi, T., Panneton, R., Rabagliati, H., Soderstrom, M., Sullivan, J., Waxman, S., & Yurovsky, D. (2017). A Collaborative Approach to Infant Research: Promoting Reproducibility, Best Practices, and Theory-Building. *Infancy*, *22*(4), 421–435. https://doi.org/10.1111/infa.12182

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental Trajectories in Toddlers' Self-restraint Predict Individual Differences in Executive Functions 14 Years Later: A Behavioral Genetic Analysis. *Developmental Psychology*, *47*(5), 1410–1430. https://doi.org/10.1037/a0023750

Friedman, S. (1972). Habituation and recovery of visual response in the alert human newborn. *Journal of Experimental Child Psychology*, *13*(2), 339–349. https://doi.org/10.1016/0022-0965(72)90095-1

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

Fries, P. (2009). Neuronal Gamma-Band Synchronization as a Fundamental Process in Cortical Computation. *Annual Review of Neuroscience*, *32*(1), 209–224. https://doi.org/10.1146/annurev.neuro.051508.135603

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, *88*(1), 220–235. https://doi.org/10.1016/j.neuron.2015.09.034

Fries, P., Nikolić, D., & Singer, W. (2007). The gamma cycle. *Trends in Neurosciences*, *30*(7), 309–316. https://doi.org/10.1016/j.tins.2007.05.005

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of Oscillatory Neuronal Synchronization by Selective Visual Attention. *Science*, *291*(5508), 1560–1563. https://doi.org/10.1126/science.1055465

Fuentes, L. J. (2004). Inhibitory Processing in the Attentional Networks. In *Cognitive neuroscience of attention* (pp. 45–55). The Guilford Press.

Gibson, E. J. (1988). Exploratory Behavior in the Development of Perceiving, Acting, and the Acquiring of Knowledge. *Annual Review of Psychology*, *39*(1), 1–42. https://doi.org/10.1146/annurev.ps.39.020188.000245

Gilbert, S. J., & Burgess, P. W. (2008). Executive function. *Current Biology*, *18*(3), R110–R114.

Gliga, T., & Dehaene-Lambertz, G. (2005). Structural Encoding of Body and Face in Human Infants and Adults. *Journal of Cognitive Neuroscience*, *17*(8), 1328–1340. https://doi.org/10.1162/0898929055002481

Gliga, T., Smith, T. J., Likely, N., Charman, T., & Johnson, M. H. (2018). Early Visual Foraging in Relationship to Familial Risk for Autism and Hyperactivity/Inattention. *Journal of Attention Disorders*, *22*(9), 839–847. https://doi.org/10.1177/1087054715616490

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1803), 20190502. https://doi.org/10.1098/rstb.2019.0502

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/10.1073/pnas.1700811114

Gottlieb, J., Lopes, M., & Oudeyer, P.-Y. (2016). Motivated Cognition: Neural and Computational Mechanisms of Curiosity, Attention, and Intrinsic Motivation. In *Recent Developments in Neuroscience Research on Human Motivation* (Vol. 19, pp. 149–172). Emerald Group Publishing Limited. https://doi.org/10.1108/S0749-742320160000019017

Gray, C. M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proceedings of the National Academy of Sciences*, *86*(5), 1698–1702. https://doi.org/10.1073/pnas.86.5.1698

Groos, K. (1901). Hat Kant Hume's Treatise gelesen? *Kant-Studien*, *5*(1–3), 177–181. https://doi.org/10.1515/kant-1901-0119

Gross, J., Schnitzler, A., Timmermann, L., & Ploner, M. (2007). Gamma Oscillations in Human Primary Somatosensory Cortex Reflect Pain Perception. *PLOS Biology*, *5*(5), e133. https://doi.org/10.1371/journal.pbio.0050133

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of Curiosity Modulate Hippocampus-Dependent Learning via the Dopaminergic Circuit. *Neuron*, *84*(2), 486–496. https://doi.org/10.1016/j.neuron.2014.08.060

Grün, S. (2009). Data-Driven Significance Estimation for Precise Spike Correlation. *Journal of Neurophysiology*, *101*(3), 1126–1140. https://doi.org/10.1152/jn.00093.2008

Grün, S., Diesmann, M., & Aertsen, A. (2002). Unitary Events in Multiple Single-Neuron Spiking Activity: I. Detection and Significance. *Neural Computation*, *14*(1), 43–80. https://doi.org/10.1162/089976602753284455

Guillory, S. B., & Kaldy, Z. (2019). Persistence and Accumulation of Visual Memories for Objects in Scenes in 12-Month-Old Infants. *Frontiers in Psychology*, *10*, 2454. https://doi.org/10.3389/fpsyg.2019.02454

Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (2014). Sins of omission: Children selectively explore when teachers are under-informative. *Cognition*, *132*(3), 335–341. https://doi.org/10.1016/j.cognition.2014.04.013

Hahamy, A., Behrmann, M., & Malach, R. (2015). The idiosyncratic brain: Distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nature Neuroscience*, *18*(2), 302–309. https://doi.org/10.1038/nn.3919

Halbwachs, M., Muller, J.-B., Tich, S. N. T., Rochebrochard, E. de L., Gascoin, G., Branger, B., Rouger, V., Rozé, J.-C., & Flamant, C. (2013). Usefulness of Parent-Completed ASQ for Neurodevelopmental Screening of Preterm Children at Five Years of Age. *PLOS ONE*, *8*(8), e71925. https://doi.org/10.1371/journal.pone.0071925

Harel, G., & Koichu, B. (2010). An operational definition of learning. *The Journal of Mathematical Behavior*, *29*(3), 115–124. https://doi.org/10.1016/j.jmathb.2010.06.002

Harlow, H. F. (1950). Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *Journal of Comparative and Physiological Psychology*, *43*(4), 289.

Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., & McCormick, D. A. (2005). Inhibitory Postsynaptic Potentials Carry Synchronized Frequency Information in Active Cortical Networks. *Neuron*, *47*(3), 423–435. https://doi.org/10.1016/j.neuron.2005.06.016

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*(4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009

Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory* (pp. xix, 335). Wiley.

Hebb, D. O. (1955). Drives and the C. N. S. (conceptual nervous system). *Psychological Review*, *62*(4), 243–254. https://doi.org/10.1037/h0041823

Hebb, D. O. (1958). The motivating effects of exteroceptive stimulation. *American Psychologist*, *13*(3), 109.

Henderson, J. M. (1993). Eye movement control during visual object processing: Effects of initial fixation position and semantic constraint. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *47*(1), 79–98. https://doi.org/10.1037/h0078776

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504. https://doi.org/10.1016/j.tics.2003.09.006

Hendry, A., Johnson, M. H., & Holmboe, K. (2019). Early Development of Visual Attention: Change, Stability, and Longitudinal Associations. *Annual Review of Developmental Psychology*, *1*(1), 251–275. https://doi.org/10.1146/annurev-devpsych-121318-085114

Hermes, D., Miller, K. J., Wandell, B. A., & Winawer, J. (2015). Stimulus Dependence of Gamma Oscillations in Human Visual Cortex. *Cerebral Cortex*, *25*(9), 2951–2959. https://doi.org/10.1093/cercor/bhu091

Hochmann, J.-R., Mody, S., & Carey, S. (2016). Infants' representations of same and different in match- and non-match-to-sample. *Cognitive Psychology*, *86*, 87–111. https://doi.org/10.1016/j.cogpsych.2016.01.005

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory*.

Hunt, J. McV. (1963). *Motivation inherent in information processing and action*.

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

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, *19*(3), 338–352. https://doi.org/10.1037/0012-1649.19.3.338

Hwang, K., Velanova, K., & Luna, B. (2010). Strengthening of Top-Down Frontal Cognitive Control Networks Underlying the Development of Inhibitory Control: A Functional Magnetic Resonance Imaging Effective Connectivity Study. *Journal of Neuroscience*, *30*(46), 15535–15545. https://doi.org/10.1523/JNEUROSCI.2825-10.2010

Isomura, Y., Sirota, A., Özen, S., Montgomery, S., Mizuseki, K., Henze, D. A., & Buzsáki, G. (2006). Integration and Segregation of Activity in Entorhinal-Hippocampal Subregions by Neocortical Slow Oscillations. *Neuron*, *52*(5), 871–882. https://doi.org/10.1016/j.neuron.2006.10.023

ITSP Technical Report (Technical Report No. 1-800-872–1726). (2005). Harcourt Assessment, Inc. https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/sensory-profile/itsp-technical-report.pdf

Jia, X., Tanabe, S., & Kohn, A. (2013). Gamma and the Coordination of Spiking Activity in Early Visual Cortex. *Neuron*, *77*(4), 762–774. https://doi.org/10.1016/j.neuron.2012.12.036

Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye–Hand Coordination in Object Manipulation. *Journal of Neuroscience*, *21*(17), 6917–6932. https://doi.org/10.1523/JNEUROSCI.21-17-06917.2001

Josselyn, S. A., & Frankland, P. W. (2018). Memory Allocation: Mechanisms and Function. *Annual Review of Neuroscience*, *41*(1), 389–413. https://doi.org/10.1146/annurev-neuro-080317-061956

Josselyn, S. A., & Tonegawa, S. (2020). Memory engrams: Recalling the past and imagining the future. *Science*, *367*(6473), eaaw4325. https://doi.org/10.1126/science.aaw4325

Jung, F., & Carlén, M. (2021). Chapter Twelve—Neuronal oscillations and the mouse prefrontal cortex. In A. T. Brockett, L. M. Amarante, M. Laubach, & M. R. Roesch (Eds.), *International Review of Neurobiology* (Vol. 158, pp. 337–372). Academic Press. https://doi.org/10.1016/bs.irn.2020.11.005

Jurjuţ, O. F., Gheorghiu, M., Singer, W., Nikolić, D., & Mureşan, R. C. (2019). Hold Your Methods! How Multineuronal Firing Ensembles Can Be Studied Using Classical Spike-Train Analysis Techniques. *Frontiers in Systems Neuroscience*, *13*. https://doi.org/10.3389/fnsys.2019.00021

Kagan, J. (1972). Motives and development. *Journal of Personality and Social Psychology*, *22*(1), 51.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kaldy, Z., Guillory, S. B., & Blaser, E. (2015). Delayed Match Retrieval: A novel anticipation-based visual working memory paradigm. *Developmental Science*, *19*(6), 892–900. https://doi.org/10.1111/desc.12335

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The Wick in the Candle of Learning: Epistemic Curiosity Activates Reward Circuitry and Enhances Memory. *Psychological Science*, *20*(8), 963–973. https://doi.org/10.1111/j.1467-9280.2009.02402.x

Kaplan, F., & Oudeyer, P.-Y. (2007a). In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience*, *1*, 17. https://doi.org/10.3389/neuro.01.1.1.017.2007

Kaplan, F., & Oudeyer, P.-Y. (2007b). The progress-drive hypothesis: An interpretation of early imitation. In C. Nehaniv & K. Dautenhahn (Eds.), *Models and Mechanims of Imitation and Social Learning: Behavioural, Social and Communication Dimensions* (pp. 361–377). Cambridge University Press. https://doi.org/10.1017/CBO9780511489808.024

Kaufman, J., Csibra, G., & Johnson, M. H. (2003). Representing occluded objects in the human infant brain. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. https://doi.org/10.1098/rsbl.2003.0067

Kaufman, J., Csibra, G., & Johnson, M. H. (2005). Oscillatory activity in the infant brain reflects object maintenance. *Proceedings of the National Academy of Sciences*, *102*(42), 15271–15274. https://doi.org/10.1073/pnas.0507626102

Kessen, W., Salapatek, P., & Haith, M. (1972). The visual response of the human newborn to linear contour. *Journal of Experimental Child Psychology*, *13*(1), 9–20. https://doi.org/10.1016/0022-0965(72)90003-3

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks Effect in Infant Auditory Attention. *Child Development*, *85*(5), 1795–1804. https://doi.org/10.1111/cdev.12263

Kietzmann, T. C., & König, P. (2015). Effects of contextual information and stimulus ambiguity on overt visual sampling behavior. *Vision Research*, *110*, 76–86. https://doi.org/10.1016/j.visres.2015.02.023

Kim, H., Ährlund-Richter, S., Wang, X., Deisseroth, K., & Carlén, M. (2016). Prefrontal Parvalbumin Neurons in Control of Attention. *Cell*, *164*(1), 208–218. https://doi.org/10.1016/j.cell.2015.11.038

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776. https://doi.org/10.1016/j.visres.2005.10.002

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science*, *6*(5), 449–467. https://doi.org/10.1111/1467-7687.00300

Knudsen, E. I. (2007). Fundamental Components of Attention. *Annual Review of Neuroscience*, *30*(1), 57–78. https://doi.org/10.1146/annurev.neuro.30.051606.094256

Kobayashi, K., & Hsu, M. (2019). Common neural code for reward and information value. *Proceedings of the National Academy of Sciences*, *116*(26), 13061–13066. https://doi.org/10.1073/pnas.1820145116

Kopp, C. B., & Vaughn, B. E. (1982). Sustained Attention during Exploratory Manipulation as a Predictor of Cognitive Competence in Preterm Infants. *Child Development*, *53*(1), 174–182. https://doi.org/10.2307/1129650

Kubovy, M., & Wagemans, J. (1995). Grouping by Proximity and Multistability in Dot Lattices: A Quantitative Gestalt Theory. *Psychological Science*, *6*(4), 225–234. https://doi.org/10.1111/j.1467-9280.1995.tb00597.x

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science (New York, N.Y.)*, *320*(5872), 110–113. https://doi.org/10.1126/science.1154735

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology*, *94*(3), 1904–1911. https://doi.org/10.1152/jn.00263.2005

Land, M. F. (1999). Motion and vision: Why animals move their eyes. *Journal of Comparative Physiology A*, *185*(4), 341–352. https://doi.org/10.1007/s003590050393

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research*, *25*(3), 296–324. https://doi.org/10.1016/j.preteyeres.2006.01.002

Landau, A. N., Schreyer, H. M., van Pelt, S., & Fries, P. (2015). Distributed Attention Is Implemented through Theta-Rhythmic Gamma Modulation. *Current Biology*, *25*(17), 2332–2337. https://doi.org/10.1016/j.cub.2015.07.048

Lawson, K. R., & Ruff, H. A. (2004). Early Focused Attention Predicts Outcome for Children Born Prematurely. *Journal of Developmental & Behavioral Pediatrics*, *25*(6), 399–406.

Leckart, B. T. (1966). Looking time: The effects of stimulus complexity and familiarity. *Perception & Psychophysics*, *1*(3), 142–144. https://doi.org/10.3758/BF03210045

Leguire, L. E., & Rogers, G. L. (1985). Pattern electroretinogram: Use of noncorneal skin electrodes. *Vision Research*, *25*(6), 867–870. https://doi.org/10.1016/0042-6989(85)90195-6

Leopold, D. A., Murayama, Y., & Logothetis, N. K. (2003). Very Slow Activity Fluctuations in Monkey Visual Cortex: Implications for Functional Brain Imaging. *Cerebral Cortex*, *13*(4), 422–433. https://doi.org/10.1093/cercor/13.4.422

Liddle, E. B., Hollis, C., Batty, M. J., Groom, M. J., Totman, J. J., Liotti, M., Scerif, G., & Liddle, P. F. (2011). Task-related default mode network modulation and inhibitory control in ADHD: Effects of motivation and methylphenidate. *Journal of Child Psychology and Psychiatry*, *52*(7), 761–771. https://doi.org/10.1111/j.1469-7610.2010.02333.x

Lima, B., Singer, W., Chen, N.-H., & Neuenschwander, S. (2010). Synchronization Dynamics in Response to Plaid Stimuli in Monkey V1. *Cerebral Cortex*, *20*(7), 1556–1573. https://doi.org/10.1093/cercor/bhp218

Litman, J. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition & Emotion*, *19*(6), 793–814.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75. https://doi.org/10.1037/0033-2909.116.1.75

Loughlin, P., Pitton, J., & Atlas, L. (1992). An information-theoretic approach to positive time-frequency distributions. *[Proceedings] ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*. https://doi.org/10.1109/ICASSP.1992.226642

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., &

Wagenmakers, E.-J. (2019). JASP: Graphical Statistical Software for Common Statistical Designs. *Journal of Statistical Software*, *88*, 1–17. https://doi.org/10.18637/jss.v088.i02

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/10.1016/j.cognition.2013.12.010

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. *Journal of Neurophysiology*, *77*(1), 24–42. https://doi.org/10.1152/jn.1997.77.1.24

Mackay, D. G. (1973). Aspects of the Theory of Comprehension, Memory and Attention. *Quarterly Journal of Experimental Psychology*, *25*(1), 22–40. https://doi.org/10.1080/14640747308400320

Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2002). Dynamic Brain Sources of Visual Evoked Responses. *Science*, *295*(5555), 690–694. https://doi.org/10.1126/science.1066168

Mather, E. (2013). Novelty, attention, and challenges for developmental psychology. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00491

McCall, R. B., & McGhee, P. E. (1977). The Discrepancy Hypothesis of Attention and Affect in Infants. In I. Č. Užgiris & F. Weizmann (Eds.), *The Structuring of Experience* (pp. 179–210). Springer US. https://doi.org/10.1007/978-1-4615-8786-6_7

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

McIntosh, D. N., Miller, L. J., Shyu, V., & Hagerman, R. J. (1999). Sensory-modulation disruption, electrodermal responses, and functional behaviors. *Developmental Medicine and Child Neurology*, *41*(9), 608–615. https://doi.org/10.1017/S0012162299001267

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moca, V. V., Bârzan, H., Nagy-Dăbâcan, A., & Mureșan, R. C. (2021). Time-frequency super-resolution with superlets. *Nature Communications*, *12*(1), 337. https://doi.org/10.1038/s41467-020-20539-9

Moca, V. V., Nikolić, D., Singer, W., & Mureşan, R. C. (2014). Membrane Resonance Enables Stable and Robust Gamma Oscillations. *Cerebral Cortex*, *24*(1), 119–142. https://doi.org/10.1093/cercor/bhs293

Moca, V. V., Ţincaş, I., Melloni, L., & Mureşan, R. C. (2011). Visual Exploration and Object Recognition by Lattice Deformation. *PLOS ONE*, *6*(7), e22831. https://doi.org/10.1371/journal.pone.0022831

Montessori, M. (1912). A critical consideration of the new pedagogy in its relation to modern science. In *The Montessori method: Scientific pedagogy as applied child education in 'The Children's Houses', with additions and revisions by the author* (pp. 1–27). Frederick A Stokes Company. https://doi.org/10.1037/13054-001

Montgomery, K. C. (1954). The role of the exploratory drive in learning. *Journal of Comparative and Physiological Psychology*, *47*(1), 60–64. https://doi.org/10.1037/h0054833

Morrison, F. J., Ponitz, C. C., & McClelland, M. M. (2010). Self-regulation and academic achievement in the transition to school. In *Child development at the intersection of emotion and cognition* (pp. 203–224). American Psychological Association. https://doi.org/10.1037/12059-011

Muentener, P., Herrig, E., & Schulz, L. (2018). The Efficiency of Infants' Exploratory Play Is Related to Longer-Term Cognitive Development. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00635

Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., & Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(2), 241–268. https://doi.org/10.3758/s13415-011-0083-5

Nikolić, D., Mureşan, R. C., Feng, W., & Singer, W. (2012). Scaled correlation analysis: A better way to compute a cross-correlogram. *European Journal of Neuroscience*, *35*(5), 742–762. https://doi.org/10.1111/j.1460-9568.2011.07987.x

Nissen, H. W. (1930). A Study of Exploratory Behavior in the White Rat by Means of the Obstruction Method. *The Pedagogical Seminary and Journal of Genetic Psychology*, *37*(3), 361–376. https://doi.org/10.1080/08856559.1930.9944162

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological Review*, *75*(6), 522–536. https://doi.org/10.1037/h0026699

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, *42*(1), 188–204. https://doi.org/10.3758/BRM.42.1.188

Odom, J. V., Bach, M., Brigell, M., Holder, G. E., McCulloch, D. L., Tormene, A. P., & Vaegan. (2010). ISCEV standard for clinical visual evoked potentials (2009 update). *Documenta Ophthalmologica*, *120*(1), 111–119. https://doi.org/10.1007/s10633-009-9195-4

Orhan, A. E., & Jacobs, R. A. (2014). Toward ecologically realistic theories in visual short-term memory research. *Attention, Perception, & Psychophysics*, *76*(7), 2158–2170. https://doi.org/10.3758/s13414-014-0649-8

Oudeyer, P.-Y. (2017). Autonomous development and learning in artificial intelligence and robotics: Scaling up deep learning to human-like learning. *Behavioral and Brain Sciences*, *40*. https://doi.org/10.1017/S0140525X17000243

Oudeyer, P.-Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. In B. Studer & S. Knecht (Eds.), *Progress in Brain Research* (Vol. 229, pp. 257–284). Elsevier. https://doi.org/10.1016/bs.pbr.2016.05.005

Oudeyer, P.-Y., Kaplan, F., Hafner, V. V., & Whyte, A. (2005). The playground experiment: Task-independent development of a curious robot. *Proceedings of the AAAI Spring Symposium on Developmental Robotics*, 42–47.

Oudeyer, P.-Y., & Smith, L. B. (2016). How Evolution May Work Through Curiosity-Driven Developmental Process. *Topics in Cognitive Science*, *8*(2), 492–502. https://doi.org/10.1111/tops.12196

Palva, J. M., Palva, S., & Kaila, K. (2005). Phase Synchrony among Neuronal Oscillations in the Human Cortex. *Journal of Neuroscience*, *25*(15), 3962–3972. https://doi.org/10.1523/JNEUROSCI.4250-04.2005

Parr, T., & Friston, K. J. (2017). Working memory, attention, and salience in active inference. *Scientific Reports*, *7*(1), 14678. https://doi.org/10.1038/s41598-017-15249-0

Pavlov, I. P., & Thompson, W. H. (1910). *The work of the digestive glands;* London, C. Griffin. http://archive.org/details/workofdigestiveg00pavlrich

Peters, J. M., Taquet, M., Vega, C., Jeste, S. S., Fernández, I. S., Tan, J., Nelson, C. A., Sahin, M., & Warfield, S. K. (2013). Brain functional networks in syndromic and non-syndromic autism: A graph theoretical study of EEG connectivity. *BMC Medicine*, *11*(1), 54. https://doi.org/10.1186/1741-7015-11-54

Petersen, S. E., & Posner, M. I. (2012). The Attention System of the Human Brain: 20 Years After. *Annual Review of Neuroscience*, *35*(1), 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Peyrache, A., Battaglia, F. P., & Destexhe, A. (2011). Inhibition recruitment in prefrontal cortex during sleep spindles and gating of hippocampal inputs. *Proceedings of the National Academy of Sciences*, *108*(41), 17207–17212. https://doi.org/10.1073/pnas.1103612108

Piaget, J. (1952). *The origins of intelligence in children* (M. Cook, Trans.; Vol. 8). International Universities Press New York.

Piccardi, E. S., Johnson, M. H., & Gliga, T. (2020). Explaining individual differences in infant visual sensory seeking. *Infancy*, *25*(5), 677–698. https://doi.org/10.1111/infa.12356

Plenz, D., Ribeiro, T. L., Miller, S. R., Kells, P. A., Vakili, A., & Capek, E. L. (2021). Self-Organized Criticality in the Brain. *ArXiv:2102.09124 [q-Bio]*. http://arxiv.org/abs/2102.09124

Pomiechowska, B., & Gliga, T. (2021). Nonverbal category knowledge limits the amount of information encoded in object representations: EEG evidence from 12-month-old infants. *Royal Society Open Science*, *8*(3), 200782. https://doi.org/10.1098/rsos.200782

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Posner, M. I., & Raichle, M. E. (1994). *Images of mind* (pp. ix, 257). Scientific American Library/Scientific American Books.

Posner, M. I., & Rothbart, M. K. (1998). Summary and commentary: Developing attentional skills. In *Cognitive neuroscience of attention: A developmental perspective* (J. E. Richards, pp. 317–323). Lawrence Erlbaum Associates Publishers.

Posner, M. I., & Rothbart, M. K. (2007). Research on Attention Networks as a Model for the Integration of Psychological Science. *Annual Review of Psychology*, *58*(1), 1–23. https://doi.org/10.1146/annurev.psych.58.110405.085516

Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior and Development*, *29*(3), 386–401. https://doi.org/10.1016/j.infbeh.2006.01.004

Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., & Leerkes, E. (2014). Development and Assessment of Short and Very Short Forms of the Infant Behavior Questionnaire–Revised. *Journal of Personality Assessment*, *96*(4), 445–458. https://doi.org/10.1080/00223891.2013.841171

Putnam, S. P., & Rothbart, M. K. (2006). Development of Short and Very Short Forms of the Children's Behavior Questionnaire. *Journal of Personality Assessment*, *87*(1), 102–112. https://doi.org/10.1207/s15327752jpa8701_09

Radulescu, A., Shin, Y. S., & Niv, Y. (2021). Human Representation Learning. *Annual Review of Neuroscience*, *44*(1), 253–273. https://doi.org/10.1146/annurev-neuro-092920-120559

Ray, S., Miller, M., Karalunas, S., Robertson, C., Grayson, D. S., Cary, R. P., Hawkey, E., Painter, J. G., Kriz, D., Fombonne, E., Nigg, J. T., & Fair, D. A. (2014). Structural and functional connectivity of the human brain in autism spectrum disorders and attention-deficit/hyperactivity disorder: A rich club-organization study. *Human Brain Mapping*, *35*(12), 6032–6048. https://doi.org/10.1002/hbm.22603

Ren, P., Xiao, Y., Chang, X., Huang, P.-Y., Li, Z., Gupta, B. B., Chen, X., & Wang, X. (2021). A Survey of Deep Active Learning. *ACM Computing Surveys*, *54*(9), 180:1-180:40. https://doi.org/10.1145/3472291

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive Mechanisms Subserve Attention in Macaque Areas V2 and V4. *Journal of Neuroscience*, *19*(5), 1736–1753. https://doi.org/10.1523/JNEUROSCI.19-05-01736.1999

Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience*, *3*(11), 1199–1204. https://doi.org/10.1038/81479

Risko, E. F., Anderson, N. C., Lanthier, S., & Kingstone, A. (2012). Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. *Cognition*, *122*(1), 86–90. https://doi.org/10.1016/j.cognition.2011.08.014

Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal Expectation Improves the Quality of Sensory Information. *Journal of Neuroscience*, *32*(24), 8424–8428. https://doi.org/10.1523/JNEUROSCI.0804-12.2012

Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of Temperament at Three to Seven Years: The Children's Behavior Questionnaire. *Child Development*, *72*(5), 1394–1408. https://doi.org/10.1111/1467-8624.00355

Rothbart, M. K., & Derryberry, D. (1981). Development of individual differences in temperament. In *Advances in developmental psychology* (Vol. 1, pp. 37–86). Guilford.

Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychobiological approach to the development of temperament. In *Temperament: Individual differences at the interface of biology and behavior* (pp. 83–116). American Psychological Association. https://doi.org/10.1037/10149-003

Rothbart, M. K., Sheese, B. E., & Posner, M. I. (2007). Executive Attention and Effortful Control: Linking Temperament, Brain Networks, and Genes. *Child Development Perspectives*, *1*(1), 2–7. https://doi.org/10.1111/j.1750-8606.2007.00002.x

Ruddock, K. H., Wooding, D. S., & Mannan, S. K. (1996). The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images. *Spatial Vision*, *10*(3), 165–188. https://doi.org/10.1163/156856896X00123

Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The Development of Executive Attention: Contributions to the Emergence of Self-Regulation. *Developmental Neuropsychology*, *28*(2), 573–594. https://doi.org/10.1207/s15326942dn2802_2

Ruggeri, A., Swaboda, N., Sim, Z. L., & Gopnik, A. (2019). Shake it baby, but only when needed: Preschoolers adapt their exploratory strategies to the information structure of the task. *Cognition*, *193*, 104013. https://doi.org/10.1016/j.cognition.2019.104013

Sailer, U., Flanagan, J. R., & Johansson, R. S. (2005). Eye–Hand Coordination during Learning of a Novel Visuomotor Task. *Journal of Neuroscience*, *25*(39), 8833–8842. https://doi.org/10.1523/JNEUROSCI.2658-05.2005

Salarian, A. (2022). *Intraclass Correlation Coefficient (ICC)*. https://www.mathworks.com/matlabcentral/fileexchange/22099-intraclass-correlation-coefficient-icc

Schmidhuber, J. (1991). Curious model-building control systems. *Proceedings of the International Joint Conference on Neural Networks, Singapore, 1991*, *2*, 1458–1463.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for Rewards Like a Child Means Less Generalization and More Directed Exploration. *Psychological Science*, *30*(11), 1561–1572. https://doi.org/10.1177/0956797619863663

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*(4), 1045–1050. https://doi.org/10.1037/0012-1649.43.4.1045

Schunk, D. H. (2012). Learning theories an educational perspective sixth edition. Pearson.

Schwartz, B., & Reisberg, D. (1991). *Learning and memory* (pp. xix, 663). W W Norton & Co.

Schwiedrzik, C. M., Ruff, C. C., Lazar, A., Leitner, F. C., Singer, W., & Melloni, L. (2014). Untangling Perceptual Memory: Hysteresis and Adaptation Map into Separate Cortical Networks. *Cerebral Cortex*, *24*(5), 1152–1164. https://doi.org/10.1093/cercor/bhs396

Semon, R. (1904). *Die Mneme als erhaltendes Prinzip im Wechsel des organischen Geschehens*. W. Engelmann.

Shankar, V., Roelofs, R., Mania, H., Fang, A., Recht, B., & Schmidt, L. (2020). Evaluating Machine Accuracy on ImageNet. *Proceedings of the 37th International Conference on Machine Learning*, 8634–8644. https://proceedings.mlr.press/v119/shankar20c.html

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/10.1038/s41467-021-23431-2

Sim, Z. L., & Xu, F. (2017). Learning higher-order generalizations through free play: Evidence from 2- and 3-year-old children. *Developmental Psychology*, *53*(4), 642–651. https://doi.org/10.1037/dev0000278

Singer, W. (1999). Neuronal Synchrony: A Versatile Code for the Definition of Relations? *Neuron*, *24*(1), 49–65. https://doi.org/10.1016/S0896-6273(00)80821-1

Sirois, S., & Mareschal, D. (2002). Models of habituation in infancy. *Trends in Cognitive Sciences*, *6*(7), 293–298. https://doi.org/10.1016/S1364-6613(02)01926-5

Sirota, A., Montgomery, S., Fujisawa, S., Isomura, Y., Zugaro, M., & Buzsáki, G. (2008). Entrainment of Neocortical Neurons and Gamma Oscillations by the Hippocampal Theta Rhythm. *Neuron*, *60*(4), 683–697. https://doi.org/10.1016/j.neuron.2008.09.014

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. Appleton-Century.

Slater, A., Morison, V., & Rose, D. (1982). Visual memory at birth. *British Journal of Psychology*, *73*(4), 519–525. https://doi.org/10.1111/j.2044-8295.1982.tb01834.x

Slater, A., Morison, V., & Rose, D. (1984). Habituation in the newborn. *Infant Behavior and Development*, *7*(2), 183–200. https://doi.org/10.1016/S0163-6383(84)80057-0

Slater, A., Morison, V., Town, C., & Rose, D. (1985). Movement perception and identity constancy in the new-born baby. *British Journal of Developmental Psychology*, *3*(3), 211–220. https://doi.org/10.1111/j.2044-835X.1985.tb00974.x

Slone, L. K., Smith, L. B., & Yu, C. (2019). Self-generated variability in object images predicts vocabulary growth. *Developmental Science*, *22*(6), e12816. https://doi.org/10.1111/desc.12816

Snijders, T. M., Milivojevic, B., & Kemner, C. (2013). Atypical excitation–inhibition balance in autism captured by the gamma response to contextual modulation. *NeuroImage: Clinical*, *3*, 65–72. https://doi.org/10.1016/j.nicl.2013.06.015

Snyder, J. S., Schwiedrzik, C. M., Vitela, A. D., & Melloni, L. (2015). How previous experience shapes perception in different sensory modalities. *Frontiers in Human Neuroscience*, *9*, 594. https://doi.org/10.3389/fnhum.2015.00594

Snyder, K. A., & Keil, A. (2008). Repetition Suppression of Induced Gamma Activity Predicts Enhanced Orienting toward a Novel Stimulus in 6-month-old Infants. *Journal of Cognitive Neuroscience*, *20*(12), 2137–2152. https://doi.org/10.1162/jocn.2008.20149

Southwell, R., Baumann, A., Gal, C., Barascud, N., Friston, K., & Chait, M. (2017). Is predictability salient? A study of attentional capture by auditory patterns. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714), 20160105. https://doi.org/10.1098/rstb.2016.0105

Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires®*. Paul H. Brookes Publishing Co.

Squires, J., Potter, L., & Bricker, D. (1995). *The ASQ user's guide for the Ages & Stages Questionnaires: A parent-completed, child-monitoring system* (pp. xvi, 156). Paul H Brookes Publishing.

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*. https://doi.org/10.1126/science.aaa3799

Steriade, M. (1999). Coherent oscillations and short-term plasticity in corticothalamic networks. *Trends in Neurosciences*, *22*(8), 337–345. https://doi.org/10.1016/S0166-2236(99)01407-1

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press.

Suzuki, Y., Minami, T., & Nakauchi, S. (2018). Association between pupil dilation and implicit processing prior to object recognition via insight. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-25207-z

Tanaka, K. (1996). Inferotemporal Cortex and Object Vision. *Annual Review of Neuroscience*, *19*(1), 109–139. https://doi.org/10.1146/annurev.ne.19.030196.000545

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, *11*(5), 5. https://doi.org/10.1167/11.5.5

Taylor, K., Mandon, S., Freiwald, W. A., & Kreiter, A. K. (2005). Coherent Oscillatory Activity in Monkey Area V4 Predicts Successful Allocation of Attention. *Cerebral Cortex*, *15*(9), 1424–1437. https://doi.org/10.1093/cercor/bhi023

Ten, A., Kaushik, P., Oudeyer, P.-Y., & Gottlieb, J. (2021). Humans monitor learning progress in curiosity-driven exploration. *Nature Communications*, *12*(1), 5972. https://doi.org/10.1038/s41467-021-26196-w

Thelen, E. (1994). Three-month-old infants can learn task-specific patterns of interlimb coordination. *Psychological Science*, *5*(5), 280–285. https://doi.org/10.1111/j.1467-9280.1994.tb00626.x

Thorndike, E. L. (1898). *Animal intelligence: An experimental study of the associative processes in animals*. Macmillan,. https://www.biodiversitylibrary.org/bibliography/25848

Thorndike, E. L. (1927). The Law of Effect. *The American Journal of Psychology*, *39*(1/4), 212–222. https://doi.org/10.2307/1415413

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. https://doi.org/10.1038/381520a0

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934–1945. https://doi.org/10.1162/jocn.2009.21131

Uhlhaas, P. J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., & Singer, W. (2010). Neural synchrony and the development of cortical networks. *Trends in Cognitive Sciences*, *14*(2), 72–80. https://doi.org/10.1016/j.tics.2009.12.002

Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 892–903. https://doi.org/10.1037/0096-1523.22.4.892

van Lieshout, L. L. F., de Lange, F. P., & Cools, R. (2021). Uncertainty increases curiosity, but decreases happiness. *Scientific Reports*, *11*(1), 14014. https://doi.org/10.1038/s41598-021-93464-6

van Schijndel, T. J. P., Visser, I., van Bers, B. M. C. W., & Raijmakers, M. E. J. (2015). Preschoolers perform more informative experiments after observing theory-violating evidence. *Journal of Experimental Child Psychology*, *131*, 104–119. https://doi.org/10.1016/j.jecp.2014.11.008

Võ, M. L.-H., & Wolfe, J. M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences*, *1339*(1), 72–81. https://doi.org/10.1111/nyas.12667

von Stein, A., & Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: From local gamma to long range alpha/theta synchronization. *International Journal of Psychophysiology*, *38*(3), 301–313. https://doi.org/10.1016/S0167-8760(00)00172-0

Wass, S. V., Jones, E. J. H., Gliga, T., Smith, T. J., Charman, T., & Johnson, M. H. (2015). Shorter spontaneous fixation durations in infants with later emerging autism. *Scientific Reports*, *5*(1), 8284. https://doi.org/10.1038/srep08284

Weerahandi, S. (1995). ANOVA under Unequal Error Variances. *Biometrics*, *51*(2), 589–599. https://doi.org/10.2307/2532947

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297.

Wolfe, J. M. (2020). Visual Search: How Do We Find What We Are Looking For? *Annual Review of Vision Science*, *6*(1), 539–562. https://doi.org/10.1146/annurev-vision-091718-015048

Womelsdorf, T., Fries, P., Mitra, P. P., & Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature*, *439*(7077), 733–736. https://doi.org/10.1038/nature04258

Wu, R., Pruitt, Z., Runkle, M., Scerif, G., & Aslin, R. N. (2016). A neural signature of rapid category-based target selection as a function of intra-item perceptual similarity, despite inter-item dissimilarity. *Attention, Perception, & Psychophysics*, *78*(3), 749–760. https://doi.org/10.3758/s13414-015-1039-6

Wu, R., Scerif, G., Aslin, R. N., Smith, T. J., Nako, R., & Eimer, M. (2013). Searching for Something Familiar or Novel: Top–Down Attentional Selection of Specific Items or Object Categories. *Journal of Cognitive Neuroscience*, *25*(5), 719–729. https://doi.org/10.1162/jocn_a_00352

Wu, R., & Zhao, J. (2017). Prior Knowledge of Object Associations Shapes Attentional Templates and Information Acquisition. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/article/10.3389/fpsyg.2017.00843

Wyart, V., & Tallon-Baudry, C. (2008). Neural Dissociation between Visual Awareness and Spatial Attention. *Journal of Neuroscience*, *28*(10), 2667–2679. https://doi.org/10.1523/JNEUROSCI.4748-07.2008

Xie, W., Mallin, B. M., & Richards, J. E. (2018). Development of infant sustained attention and its relation to EEG oscillations: An EEG and cortical source analysis study. *Developmental Science*, *21*(3), e12562. https://doi.org/10.1111/desc.12562

Yarbus, A. L. (1956). The motion of the eye in the process of changing points of fixation. *Biofizika*, *1*, 76–78.

Yogman, M., Garner, A., Hutchinson, J., Hirsh-Pasek, K., Golinkoff, R. M., Health, C. on P. A. of C. and F., Media, C. on C. A., Baum, R., Gambon, T., Lavin, A., Mattson, G., Wissow, L., Hill, D. L., Ameenuddin, N., Chassiakos, Y. (Linda) R., Cross, C., Boyd, R., Mendelson, R., Moreno, M. A., … Smith, J. (2018). The Power of Play: A Pediatric Role in Enhancing Development in Young Children. *Pediatrics*, *142*(3). https://doi.org/10.1542/peds.2018-2058

Zetzsche, C., Schill, K., Krieger, G., Hauske, G., & Rentschler, I. (2000). Object and scene analysis by saccadic eye-movements: An investigation with higher-order statistics. *Spatial Vision*, *13*(2–3), 201–214. https://doi.org/10.1163/156856800741216

Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention Is Spontaneously Biased Toward Regularities. *Psychological Science*, *24*(5), 667–677. https://doi.org/10.1177/0956797612460407

Appendices

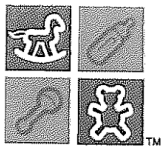
Measure	Test	Effect	Test statistic	Significance	Power value
Mean gamma (stages)	2*3 RM ANOVA: group * stage	Group	F(1,19) = 4.896 (F(1,28) = 10.266)	p = 0.039 * (p = 0.003 *)	η^2 = 0.088 (η^2 = 0.127)
		Stage	F(2,38) = 3.600 (F(2,56) = 3.621)	p = 0.037 * (p = 0.033 *)	η^2 = 0.043 (η^2 = 0.034)
		Stage * Group	F(2,38) = 0.118 (F(2,56) = 0.607)	p = 0.889 (p = 0.549)	η^2 = 0.006 (η^2 = 0.006)
	Correlation with time-on-task	*NA*	R = -0.523 (R = -0.586)	p = 0.015 * (p < 0.001 *)	*NA*
Stage difference in mean gamma (stages)	2*2 RM ANOVA: group * stage	Group	F(1,19) = 0.242 F(1, 28) = 1.170)	p = 0.629 (p = 0.289)	η^2 = 0.002 (η^2 = 0.008)
		Stage	F(1,19) = 6.806 (F(1, 28) = 7.322)	p = 0.017 * (p = 0.011 *)	η^2 = 0.264 (η^2 = 0.137)
		Stage * Group	F(1, 19) = 0.007 (F(1, 28) = 0.006)	p = 0.934 (p = 0.940)	η^2 < 0.001 (η^2 < 0.001)
	Comparison with zero	*Stopped* Stages 2-1	t(1,11) = 0.495 (t(1,15) = 0.750)	p = 0.631 (p = 0.465)	Cohen's d = 0.143 (Cohen's d = 0.188)
		Stopped Stages 3-2	t(1,11) = -0.261 (t(1,15) = 0.412)	p = 0.799 (p = 0.686)	Cohen's d = -0.075 (Cohen's d = 0.103)

		Did All Stages 2-1	t(1,9) = -0.467 (t(1,14) = -0.096)	p = 0.652 (p = 0.925)	Cohen's d = -0.148 (Cohen's d = -0.025)
		Did All Stages 3-2	t(1,9) = -0.233 (t(1,14) = -0.122)	p = 0.821 (p = 0.905)	Cohen's d = -0.074 (Cohen's d = -0.031)
Mean gamma (Vid. Rep. 1-2)	ANCOVA: group	Group	F(1,19) = 0.836 (F(1,26) = 0.801)	p = 0.372 (p = 0.378)	η^2 = 0.032 (η^2 = 0.014)
	Correlation with time-on-task	*NA*	R = -0.320 (R = -0.187)	p = 0.157 (p = 0.322)	*NA*
Gamma-ERP alignment: lag0 SCA (stages)	2*3 RM ANOVA: group * stage	Group	F(1,19) = 3.473 (F(1,28) = 2.950)	p = 0.078 **t** (p = 0.097 **t**)	η^2 = 0.046 (η^2 = 0.095)
		Stage	F(2,38) = 1.448 (F(2,56) = 2.093)	P = 0.248 (p = 0.133)	η^2 = 0.045 (η^2 = 0.045)
		Stage * Group	F(2,38) = 0.723 (F(2,56) = 2.894)	p = 0.492 (p = 0.064 **t**)	η^2 = 0.023 (η^2 = 0.062)
	Comparison with zero	*Stopped* Stage 1	t(1,11) = -1.019 (t(1,15) = -1.233)	p = 0.330 (p = 0.236)	Cohen's d = -0.294 (Cohen's d = -0.308)
		Stopped Stage 2	t(1,11) = -1.483 (t(1,15) = -1.456)	p = 0.166 (p = 0.166)	Cohen's d = -0.428 (Cohen's d = -0.364)
		Stopped Stage 3	t(1,11) = -0.614 (t(1,15) = -1.060)	p = 0.552 (p = 0.306)	Cohen's d = -0.177 (Cohen's d = -0.265)
		Did All Stage 1	t(1,9) = -0.296 (t(1,14) = -2.135)	p = 0.774 (p = 0.051[t])	Cohen's d = -0.094 (Cohen's d = -0.551)

		Did All Stage 2	t(1,9) = 1.007 (t(1,14) = 2.361)	p = 0.340 (p = 0.033*)	Cohen's d = 0.318 (Cohen's d = 0.601)
		Did All Stage 3	t(1,9) = 0.317 (t(1,14) = -0.274)	p = 0.758 (p = 0.788)	Cohen's d = 0.100 (Cohen's d = -0.071)
Stage difference in gamma-ERP alignment: lag0 SCA (stages)	2*2 RM ANOVA: group * stage	Group	F(1,19) = 0.111 F(1, 28) = 0.724)	p = 0.743 (p = 0.402)	η^2 = 0.002 (η^2 = 0.021)
		Stage	F(1,19) = 0.043 (F(1, 28) = 0.021)	p = 0.837 (p = 0.885)	η^2 = 0.002 (η^2 < 0.001)
		Stage * Group	F(1, 19) = 1.222 (F(1, 28) = 4.398)	p = 0.283 (p = 0.045 *)	η^2 = 0.046 (η^2 = 0.110)
	Comparison with zero	*Stopped* Stages 2-1	t(1,11) = -0.545 (t(1,15) = -0.419)	p = 0.597 (p = 0.681)	Cohen's d = -0.157 (Cohen's d = -0.105)
		Stopped Stages 3-2	t(1,11) = 1.128 (t(1,15) = 0.717)	p = 0.283 (p = 0.485)	Cohen's d = 0.326 (Cohen's d = 0.179)
		Did All Stages 2-1	t(1,9) = 1.037 (t(1,14) = 2.870)	p = 0.327 (p = 0.012*)	Cohen's d = 0.328 (Cohen's d = 0.741)
		Did All Stages 3-2	t(1,9) = -0.333 (t(1,14) = -1.592)	p = 0.747 (p = 0.134)	Cohen's d = -0.105 (Cohen's d = -0.411)
Gamma-ERP alignment: lag0 SCA (Vid. Rep. 1-2)	ANCOVA: group	Group	F(1,19) = 0.364 (F(1,26) < 0.001)	p = 0.554 (p = 0.986)	η^2 = 0.018 (η^2 < 0.001)
	Comparison with zero	*Stopped*	t(1,11) = 0.534 (t(1,15) = -0.641)	p = 0.604 (p = 0.531)	Cohen's d = 0.154 (Cohen's d = -0.160)
		Did All	t(1,9) = 0.780 (t(1,14) =-0.608)	p = 0.405 (p = 0.553)	Cohen's d = 0.247 (Cohen's d = -0.157)

	Correlation with time-on-task	*NA*	R = 0.284 (R = -0.280)	p = 0.211 (p = 0.134)	*NA*

Appendix 1: Summary statistics on the experimental brain measures, using the restricted follow-up sample only. Similarly to what was done on the whole sample of participants, the Ntrials variable was used as a covariate for ANCOVAs and RM ANOVAs, or as a conditioning variable for partial correlations. Statistics on the whole sample are reported again for reference in grey, under brackets and under the restricted follow-up sample values. Significance levels (p) for comparisons with zero are reported without Bonferroni corrections for 2 or 3 multiple comparisons. Significant tests are marked by a ***** ($p < 0.05$) or a **t** if they are trending towards significance ($0.05 < p < 0.1$). NA stands for non-applicable.



INFANT/TODDLER SENSORY PROFILE™

Winnie Dunn, Ph.D., OTR, FAOTA
with Debora B. Daniels, M.A., CCC-SLP

Caregiver Questionnaire

7 TO 36 MONTHS

Child's Name: _____ Birth Date: _____ Date: _____

Completed by: _____ Relationship to Child: _____

Service Provider's Name: _____ Discipline: _____

Circle the birth order of your child within the family 1st 2nd 3rd 4th 5th Other_____

Have there been more than 3 children, between the ages of birth–18 years, living in your household during the past 12 months?_____

INSTRUCTIONS

Please check the box that **best** describes the frequency with which your child does the following behaviors. Please answer all of the statements. If you are unable to comment because you have not observed the behavior or believe that it does not apply to your child, please draw an X through the number for that item. Write any comments at the end of each section.

Use the following key to mark your responses

ALMOST ALWAYS	When presented with the opportunity, your child **almost always** responds in this manner, 90% or more of the time.
FREQUENTLY	When presented with the opportunity, your child **frequently** responds in this manner, about 75% of the time.
OCCASIONALLY	When presented with the opportunity, your child **occasionally** responds in this manner, about 50% of the time.
SELDOM	When presented with the opportunity, your child **seldom** responds in this manner, about 25% of the time.
ALMOST NEVER	When presented with the opportunity, your child **almost never** responds in this manner, 10% or less of the time.

Therapy Skill Builders®
A Harcourt Assessment Company

THE PSYCHOLOGICAL CORPORATION®
A Harcourt Assessment Company

Printed in the United States of America
8 9 10 11 12 A B C D E

ISBN 076164957-3

Item		A. General Processing	ALMOST ALWAYS	FREQUENTLY	OCCASIONALLY	SELDOM	ALMOST NEVER
	1	My child's behavior deteriorates when the schedule changes.					
	2	My child avoids playing with others.					
	3	My child withdraws from situations.					

Note: You do not calculate a Raw Score Total for this section.

Comments

Item		B. Auditory Processing	ALMOST ALWAYS	FREQUENTLY	OCCASIONALLY	SELDOM	ALMOST NEVER
	4	I have to speak loudly to get my child's attention.					
	5	I have to touch my child to gain attention.					
	6	My child enjoys making sounds with his/her mouth.					
	7	My child takes a long time to respond, even to familiar voices.					
	8	My child startles easily at sound, compared to other children the same age.					
	9	My child is distracted and/or has difficulty eating in noisy environments.					
	10	My child ignores me when I am talking.					
	11	My child tries to escape from noisy environments.					
	12	My child finds ways to make noise with toys.					
	13	It takes a long time for my child to respond to his/her name when it is called.					
		Section Raw Score Total					

Comments

Item		C. Visual Processing	ALMOST ALWAYS	FREQUENTLY	OCCASIONALLY	SELDOM	ALMOST NEVER
	14	My child enjoys looking at moving or spinning objects (for example, ceiling fans, toys with wheels, floor fans).					
	15	My child enjoys looking at shiny objects.					
	16	My child avoids eye contact with me.					
	17	My child refuses to look at books with me.					
	18	My child does not recognize self in the mirror.					
	19	My child enjoys looking at own reflection in the mirror.					
	20	My child prefers fast-paced, brightly colored TV shows.					
		Section Raw Score Total					

Comments

Item		D. Tactile Processing	ALMOST ALWAYS	FREQUENTLY	OCCASIONALLY	SELDOM	ALMOST NEVER
	21	My child resists being held.					
	22	My child becomes agitated when having hair washed.					
	23	My child avoids getting face/nose wiped.					
	24	My child is distressed when having nails trimmed.					
	25	My child resists being cuddled.					
	26	My child is upset by changes in the bath water temperature, from one bath to the next.					
	27	My child avoids contact with rough or cold surfaces (for example, squirms, arches, cries).					
	28	My child becomes very upset if own clothing, hands, and/or face are messy.					
	29	My child gets upset with extreme differences in room temperature (for example, hotter, colder).					
	30	My child becomes anxious when walking or crawling on certain surfaces (for example, grass, sand, carpet, tile).					
	31	My child enjoys playing with food.					
	32	My child seeks opportunities to feel vibrations (for example, stereo speakers, washer, dryer).					
	33	My child bumps into things, seeming to not notice objects in the way.					
	34	My child enjoys splashing during bath time.					
	35	My child uses hands to explore food and other textures.					
		Section Raw Score Total					

Comments

Item		E. Vestibular Processing	ALMOST ALWAYS	FREQUENTLY	OCCASIONALLY	SELDOM	ALMOST NEVER
	36	My child requires more support for sitting than other children the same age (for example, infant seat, pillows, towel roll).					
	37	My child enjoys physical activity (for example, bouncing, being held up high in the air).					
	38	My child enjoys rhythmical activities (for example, swinging, rocking, car rides).					
	39	My child becomes upset when placed on back to change diapers.					
	40	My child resists having head tipped back during bathing.					
	41	My child cries or fusses whenever I try to move him/her.					
		Section Raw Score Total					

Comments

Item		F. Oral Sensory Processing	Almost Always	Frequently	Occasionally	Seldom	Almost Never
	42	My child licks/chews on nonfood objects.					
	43	My child mouths objects.					
	44	My child is unaware of food or liquid left on lips.					
	45	My child refuses all but a few food choices.					
	46	My child resists having teeth brushed.					
	47	My child refuses to drink from a cup.					
	48	My child refuses to try new foods.					
		Section Raw Score Total					

Comments

What do you see as your child's strengths? _____

What are your concerns? _____

STOP HERE IF YOUR CHILD IS 7 TO 36 MONTHS OLD.

ICON KEY	
—	Low Registration
	Sensation Seeking
	Sensory Sensitivity
|	Sensation Avoiding

SCORE KEY	
1	Almost Always
2	Frequently
3	Occasionally
4	Seldom
5	Almost Never

4

Appendix 2: Infant/Toddler Sensory Profile (ITSP) questionnaire for 7- to 36-month-old infants.

Early Childhood Behavior Questionnaire – Short Form

Child's name: _____ Child's birthdate: Mo:____ Day:____ Yr:____

Today's date: Month:____ Day:____ Yr:____ Child's age: _____Yrs, _____Months

Relation to child: _____ Sex of child (circle one): Male Female

INSTRUCTIONS: **Please read carefully before starting**.

As you read each description of the child's behavior below, please indicate how often the child did this during the last two weeks by circling one of the numbers in the right column. These numbers indicate how often you observed the behavior described during the last two weeks.

never	very rarely	less than half the time	about half the time	more than half the time	almost always	always	does not apply
1	2	3	4	5	6	7	NA

The "Does Not Apply" column (NA) is used when you did not see the child in the situation described during the last two weeks. For example, if the situation mentions the child going to the doctor and there was no time during the last two weeks when the child went to the doctor, circle the (NA) column. "Does Not Apply" (NA) is different from "NEVER" (1). "Never" is used when you saw the child in the situation but the child never engaged in the behavior mentioned in the last two weeks. Please be sure to circle a number or NA for every item.

When told that it was time for bed or a nap, how often did your child

1. get irritable? 1 2 3 4 5 6 7 NA

When approached by an unfamiliar person in a public place (for example, the grocery store), how often did your child

2. pull back and avoid the person? 1 2 3 4 5 6 7 NA
3. cling to a parent? 1 2 3 4 5 6 7 NA

During everyday activities, how often did your child

4. tap or drum with fingers on tables or other objects? 1 2 3 4 5 6 7 NA
5. become uncomfortable when his/her socks were not aligned properly on his/her feet? 1 2 3 4 5 6 7 NA
6. become distressed when his/her hands were dirty and/or sticky? 1 2 3 4 5 6 7 NA
7. notice low-pitched noises such as the air-conditioner, heater, or refrigerator running or starting up? 1 2 3 4 5 6 7 NA
8. blink a lot? 1 2 3 4 5 6 7 NA

While playing outdoors, how often did your child

9. enjoy sitting quietly in the sunshine? 1 2 3 4 5 6 7 NA
10. look immediately when you pointed at something? 1 2 3 4 5 6 7 NA
11. choose to take chances for the fun and excitement of it? 1 2 3 4 5 6 7 NA
12. seem to be one of the most active children? 1 2 3 4 5 6 7 NA

When s/he was carried, how often did your child

13. push against you until put down? 1 2 3 4 5 6 7 NA
14. snuggle up next to you? 1 2 3 4 5 6 7 NA

While having trouble completing a task (e.g., building, drawing, dressing), how often did your child

15. get easily irritated? 1 2 3 4 5 6 7 NA

When a familiar child came to your home, how often did your child

16. seek out the company of the child? 1 2 3 4 5 6 7 NA

When offered a choice of activities, how often did your child

17. stop and think before deciding?	1	2	3	4	5	6	7 NA
18. decide what to do very quickly and go after it?	1	2	3	4	5	6	7 NA

When asked NOT to, how often did your child

19. touch an attractive item (such as an ornament) anyway?	1	2	3	4	5	6	7 NA

During daily or evening quiet time with you and your child, how often did your child

20. enjoy just being quietly sung to?	1	2	3	4	5	6	7 NA
21. smile at the sound of words, as in nursery rhymes?	1	2	3	4	5	6	7 NA
22. enjoy just being talked to?	1	2	3	4	5	6	7 NA
23. enjoy rhythmic activities, such as rocking or swaying?	1	2	3	4	5	6	7 NA
24. want to be cuddled?	1	2	3	4	5	6	7 NA

While at home, how often did your child

25. show fear at a loud sound (blender, vacuum cleaner, etc.)?	1	2	3	4	5	6	7 NA
26. seem afraid of the dark?	1	2	3	4	5	6	7 NA

While bathing, how often did your child

27. sit quietly?	1	2	3	4	5	6	7 NA

When s/he was upset, how often did your child

28. change to feeling better within a few minutes?	1	2	3	4	5	6	7 NA

When engaged in play with his/her favorite toy, how often did your child

29. play for more than 10 minutes?	1	2	3	4	5	6	7 NA
30. continue to play while at the same time responding to your remarks or questions?	1	2	3	4	5	6	7 NA

When approaching unfamiliar children playing, how often did your child

31. watch rather than join in?	1	2	3	4	5	6	7 NA
32. seem uncomfortable?	1	2	3	4	5	6	7 NA

During everyday activities, how often did your child

33. move quickly from one place to another?	1	2	3	4	5	6	7 NA
34. notice the smoothness or roughness of objects s/he touched?	1	2	3	4	5	6	7 NA
35. become sad or blue for no apparent reason?	1	2	3	4	5	6	7 NA
36. pay attention to you right away when you called to him/her?	1	2	3	4	5	6	7 NA
37. seem to be disturbed by loud sounds?	1	2	3	4	5	6	7 NA
38. seem frightened for no apparent reason?	1	2	3	4	5	6	7 NA
39. seem to be irritated by tags in his/her clothes?	1	2	3	4	5	6	7 NA

After having been interrupted, how often did your child

40. return to a previous activity?	1	2	3	4	5	6	7 NA
41. have difficulty returning to the previous activity?	1	2	3	4	5	6	7 NA

When told that loved adults would visit, how often did your child

42. get very excited?	1	2	3	4	5	6	7 NA
43. become very happy?	1	2	3	4	5	6	7 NA

During quiet activities, such as reading a story, how often did your child

44. swing or tap his/her foot?	1	2	3	4	5	6	7 NA
45. fiddle with his/her hair, clothing, etc.?	1	2	3	4	5	6	7 NA
46. show repeated movements like squinting, hunching up the shoulders, or twitching the facial muscles?	1	2	3	4	5	6	7 NA

While playing indoors, how often did your child							
47. like rough and rowdy games?	1	2	3	4	5	6	7 NA
48. enjoy playing boisterous games like 'chase'?	1	2	3	4	5	6	7 NA
49. enjoy vigorously jumping on the couch or bed?	1	2	3	4	5	6	7 NA
In situations where s/he is meeting new people, how often did your child							
50. turn away?	1	2	3	4	5	6	7 NA
When being gently rocked or hugged, how often did your child							
51. seem eager to get away?	1	2	3	4	5	6	7 NA
When encountering a new activity, how often did your child							
52. sit on the sidelines and observe before joining in?	1	2	3	4	5	6	7 NA
53. get involved immediately?	1	2	3	4	5	6	7 NA
When visiting the home of a familiar child, how often did your child							
54. engage in an activity with the child?	1	2	3	4	5	6	7 NA
When engaged in an activity requiring attention, such as building with blocks, how often did your child							
55. move quickly to another activity?	1	2	3	4	5	6	7 NA
56. tire of the activity relatively quickly?	1	2	3	4	5	6	7 NA
While in a public place, how often did your child							
57. seem uneasy about approaching an elevator or escalator?	1	2	3	4	5	6	7 NA
58. cry or show distress when approached by an unfamiliar animal?	1	2	3	4	5	6	7 NA
59. seem afraid of large, noisy vehicles?	1	2	3	4	5	6	7 NA
60. show fear when the caregiver stepped out of sight?	1	2	3	4	5	6	7 NA
When being dressed or undressed, how often did your child							
61. squirm and try to get away?	1	2	3	4	5	6	7 NA
62. stay still?	1	2	3	4	5	6	7 NA
When told "no", how often did your child							
63. stop the forbidden activity?	1	2	3	4	5	6	7 NA
64. become sadly tearful?	1	2	3	4	5	6	7 NA
Following an exciting activity or event, how often did your child							
65. calm down quickly?	1	2	3	4	5	6	7 NA
66. have a hard time settling down?	1	2	3	4	5	6	7 NA
67. seem to feel down or blue?	1	2	3	4	5	6	7 NA
During everyday activities, how often did your child							
68. easily shift attention from one activity to another?	1	2	3	4	5	6	7 NA
69. become bothered by sounds while in noisy environments?	1	2	3	4	5	6	7 NA
70. become bothered by scratchy materials like wool?	1	2	3	4	5	6	7 NA
71. notice changes in your appearance (such as wet hair, a hat, or jewelry)?	1	2	3	4	5	6	7 NA
72. appear to listen to even very quiet sounds?	1	2	3	4	5	6	7 NA
73. seem full of energy, even in the evening?	1	2	3	4	5	6	7 NA
74. become irritated when his/her clothes were tight?	1	2	3	4	5	6	7 NA
While playing indoors, how often did your child							
75. run through the house?	1	2	3	4	5	6	7 NA
76. climb over furniture?	1	2	3	4	5	6	7 NA
77. enjoy activities such as being spun, etc.?	1	2	3	4	5	6	7 NA

When playing alone, how often did your child

78. become easily distracted?	1	2	3	4	5	6	7 NA
79. play with a set of objects for 5 minutes or longer at a time?	1	2	3	4	5	6	7 NA
80. tear materials close at hand?	1	2	3	4	5	6	7 NA

Before an exciting event (such as receiving a new toy), how often did your child

81. get very excited about getting it?	1	2	3	4	5	6	7 NA
82. remain pretty calm?	1	2	3	4	5	6	7 NA

When s/he asked for something and you said "no", how often did your child

83. become frustrated?	1	2	3	4	5	6	7 NA
84. protest with anger?	1	2	3	4	5	6	7 NA
85. have a temper tantrum?	1	2	3	4	5	6	7 NA
86. become sad?	1	2	3	4	5	6	7 NA

While playing or walking outdoors, how often did your child

87. notice sights or sounds (for example, wind chimes or water sprinklers)?	1	2	3	4	5	6	7 NA

When asked to wait for a desirable item (such as ice cream), how often did your child

88. go after it anyway?	1	2	3	4	5	6	7 NA
89. wait patiently?	1	2	3	4	5	6	7 NA

When being gently rocked, how often did your child

90. smile?	1	2	3	4	5	6	7 NA

When you removed something s/he should not have been playing with, how often did your child

91. become sad?	1	2	3	4	5	6	7 NA

While being held on your lap, how often did your child

92. seem to enjoy him/herself?	1	2	3	4	5	6	7 NA
93. mold to your body?	1	2	3	4	5	6	7 NA

When hearing about a future family outing (such as a trip to the playground), how often did your child

94. look forward to it?	1	2	3	4	5	6	7 NA

While looking at picture books on his/her own, how often did your child

95. become easily distracted?	1	2	3	4	5	6	7 NA

When a familiar adult, such as a relative or friend, visited your home, how often did your child

96. want to interact with the adult?	1	2	3	4	5	6	7 NA

When asked to do so, how often was your child able to

97. stop an ongoing activity?	1	2	3	4	5	6	7 NA
98. be careful with something breakable?	1	2	3	4	5	6	7 NA

When visiting a new place, how often did your child

99. not want to enter?	1	2	3	4	5	6	7 NA

While you were talking with someone else, how often did your child

100. easily switch attention from speaker to speaker?	1	2	3	4	5	6	7 NA

When you mildly criticized or corrected her/his behavior, how often did your child

101. get mad?	1	2	3	4	5	6	7 NA

When s/he was upset, how often did your child							
102. cry for more than 3 minutes, even when being comforted?	1	2	3	4	5	6	7 NA
103. become easily soothed?	1	2	3	4	5	6	7 NA
When you were busy, how often did your child							
104. find another activity to do when asked?	1	2	3	4	5	6	7 NA
While playing outdoors, how often did your child							
105. want to jump from heights?	1	2	3	4	5	6	7 NA
When around large gatherings of familiar adults or children, how often did your child							
106. enjoy playing with a number of different people?	1	2	3	4	5	6	7 NA
When s/he was asked to share his/her toys, how often did your child							
107. become sad?	1	2	3	4	5	6	7 NA

Appendix 3: Early Childhood Behavior Questionnaire (ECBQ), short form, for 18- to 36-month-old children.

SCORING PROCEDURE

Early Childhood Behavior Questionnaire (ECBQ)- Short Form

Scale scores for the eighteen dimensions represent the mean score of all scale items applicable to the child, as judged by the caregiver. If a caregiver omitted an item, or if the caregiver checked the "Does not apply" response option for an item, the item receives no numerical score and is not factored into the scale score.

Scores are to be computed by the following method:

1) Items indicated with an R on the items-by-scale list below are reverse-scored. Before using them to calculate the scale score, they must be reversed. This is done by subtracting the numerical response given by the caregiver from 8. Thus, a caregiver response of 7 becomes 1, 6 becomes 2, 5 becomes 3, 4 remains 4, 3 becomes 5, 2 becomes 6, and 1 becomes 7.

2) Sum the scores for items receiving a numerical response (do not include items marked "does not apply" or items receiving no response). For example, given a sum of 50 for a scale of 12 items, with one item receiving no response, two items marked "does not apply," and 9 items receiving a numerical response, the sum of 50 would be divided by 9 to yield a mean of 5.56 for the scale score.

Note: Most statistics programs will carry out these steps for you. Users of SPSS can copy the following commands into a syntax file to reverse items and calculate scale scores. The syntax assumes that items are titled "ECBQ1", "ECBQ2", "ECBQ3", etc. It is also assumed that no score was entered when caregivers omitted an item or checked "Does not apply".

```
COMPUTE ecbq27r = (8-ecbq27).
COMPUTE ecbq62r = (8-ecbq62).
COMPUTE ecbq55r = (8-ecbq55).
COMPUTE ecbq56r = (8-ecbq56).
COMPUTE ecbq78r = (8-ecbq78).
COMPUTE ecbq95r = (8-ecbq95).
COMPUTE ecbq41r = (8-ecbq41).
COMPUTE ecbq13r = (8-ecbq13).
COMPUTE ecbq51r = (8-ecbq51).
COMPUTE ecbq17r = (8-ecbq17).
COMPUTE ecbq52r = (8-ecbq52).
COMPUTE ecbq19r = (8-ecbq19).
COMPUTE ecbq88r = (8-ecbq88).
COMPUTE ecbq82r = (8-ecbq82).
COMPUTE ecbq102r = (8-ecbq102).
COMPUTE ecbq66r = (8-ecbq66).


COMPUTE shortact = mean (ecbq27R, ecbq12, ecbq73, ecbq61, ecbq75, ecbq76, ecbq33, ecbq62R).
COMPUTE shortatf = mean (ecbq29, ecbq55R, ecbq78R, ecbq56R, ecbq79, ecbq95R).
COMPUTE shortats = mean (ecbq30, ecbq40, ecbq41R, ecbq36, ecbq100, ecbq104, ecbq10, ecbq68).
COMPUTE shortcud = mean (ecbq13R, ecbq14, ecbq51R, ecbq92, ecbq93, ecbq24).
COMPUTE shortdis = mean (ecbq5, ecbq6, ecbq37, ecbq39, ecbq69, ecbq74, ecbq70).
COMPUTE shortfea = mean (ecbq38, ecbq25, ecbq58, ecbq59, ecbq60, ecbq99, ecbq26, ecbq57).
```

```
COMPUTE shortfru = mean (ecbq15, ecbq83, ecbq84, ecbq85, ecbq1, ecbq101).
COMPUTE shorthip = mean (ecbq77, ecbq11, ecbq47, ecbq48, ecbq49, ecbq105).
COMPUTE shortimp = mean (ecbq18, ecbq52R, ecbq53, ecbq17R).
COMPUTE shortinh = mean (ecbq88R, ecbq19R, ecbq63, ecbq89, ecbq97, ecbq98).
COMPUTE shortlip = mean (ecbq20, ecbq22, ecbq9, ecbq90, ecbq21, ecbq23).
COMPUTE shortmot = mean (ecbq8, ecbq4, ecbq44, ecbq45, ecbq46, ecbq80).
COMPUTE shortpsn = mean (ecbq7, ecbq71, ecbq72, ecbq34, ecbq87).
COMPUTE shortapp = mean (ecbq42, ecbq43, ecbq81, ecbq82R, ecbq94).
COMPUTE shortsad = mean (ecbq67, ecbq35, ecbq86, ecbq91, ecbq107, ecbq64).
COMPUTE shortshy = mean (ecbq2, ecbq3, ecbq50, ecbq31, ecbq32).
COMPUTE shortsoc = mean (ecbq16, ecbq106, ecbq54, ecbq96).
COMPUTE shortsth = mean (ecbq28, ecbq66R, ecbq65, ecbq102R, ecbq103).

COMPUTE SHNEGAF = MEAN (shortdis, shortfea, shortmot, shortsad, shortpsn, shortshy, (8
-shortsth), shortfru) .
COMPUTE SHSURGE = MEAN (shortimp, shortact, shorthip, shortsoc, shortapp).
COMPUTE SHEFFCO = MEAN (shortinh, shortats, shortlip, shortcud, shortatf).
```

Activity Level/Energy (8 items)
Level (rate and intensity) of gross motor activity, including rate and extent of locomotion.

While bathing, how often did your child
27.R sit quietly?

During everyday activities, how often did your child
33. move quickly from one place to another?
73. seem full of energy, even in the evening?

While playing outdoors, how often did your child
12. seem to be one of the most active children?

When being dressed or undressed, how often did your child
61. squirm and try to get away?
62.R stay still?

While playing indoors, how often did your child
75. run through the house?
76. climb over furniture?

Attentional Focusing (6 items)
Sustained duration of orienting on an object of attention; resisting distraction.

When engaged in play with his/her favorite toy, how often did your child
29. play for more than 10 minutes?

When engaged in an activity requiring attention, such as building with blocks, how often did your child
55.R move quickly to another activity?
56.R tire of the activity relatively quickly?

When playing alone, how often did your child
78.R become easily distracted?
79. play with a set of objects for 5 minutes or longer at a time?

While looking at picture books on his/her own, how often did your child
95.R become easily distracted?

Attentional Shifting (8 items)
The ability to transfer attentional focus from one activity/task to another.

While playing outdoors, how often did your child
10. look immediately when you pointed at something?

When engaged in play with his/her favorite toy, how often did your child
30. continue to play while at the same time responding to your remarks or questions?

After having been interrupted, how often did your child
40. return to a previous activity?
41.R have difficulty returning to the previous activity?

During everyday activities, how often did your child
36. pay attention to you right away when you called to him/her?
_68. easily shift attention from one activity to another?

While you were talking with someone else, how often did your child
100. easily switch attention from speaker to speaker?

When you were busy, how often did your child
104. find another activity to do when asked?

Cuddliness (6 items)
Child's expression of enjoyment in and molding of the body to being held by a caregiver.

When your child was carried, how often did s/he
13.R push against you until put down?
14. snuggle up next to you?

When being gently rocked or hugged, how often did your child
51.R seem eager to get away?

During daily or evening quiet time with you and your child, how often did your child
24. want to be cuddled?

While being held on your lap, how often did your child
92. seem to enjoy him/herself?
93. mold to your body?

Discomfort (7 items)
Amount of negative affect related to sensory qualities of stimulation, including intensity, rate or complexity of light, sound, texture.

During everyday activities, how often did your child
5. become uncomfortable when his/her socks were not aligned properly on his/her feet?
6. become distressed when his/her hands were dirty and/or sticky?
37. seem to be disturbed by loud sounds?
39. seem to be irritated by tags in his/her clothes?
69. become bothered by sounds while in noisy environments?
70. become bothered by scratchy materials like wool?
74. become irritated when his/her clothes were tight?

Fear (8 items)
Negative affect, including unease, worry, or nervousness related to anticipated pain or distress and/or potentially threatening situations; startle to sudden events.

During everyday activities, how often did your child
38. seem frightened for no apparent reason?

While at home, how often did your child
25. show fear at a loud sound (blender, vacuum cleaner, etc.)?
26. seem afraid of the dark?

While in a public place, how often did your child
57. seem uneasy about approaching an elevator or escalator?
58. cry or show distress when approached by an unfamiliar animal?
59. seem afraid of large, noisy vehicles?
60. show fear when the caregiver stepped out of sight?

When visiting a new place, how often did your child
99. not want to enter?

Frustration (6 items)
Negative affect related to interruption of ongoing tasks or goal blocking.

When told that it is time for bed or a nap, how often did your child
1. get irritable?

While having trouble completing a task (e.g., building, drawing, dressing), how often did your child
15. get easily irritated?

When s/he asked for something and you said "no", how often did your child
83. become frustrated?
84. protest with anger?
85. have a temper tantrum?

When you mildly criticized or corrected her/his behavior, how often did your child
101. get mad?

High Intensity Pleasure (6 items)
Pleasure or enjoyment related to situations involving high stimulus intensity, rate, complexity, novelty and incongruity.

While playing outdoors, how often did your child
11. choose to take chances for the fun and excitement of it?
105. want to jump from heights?

While playing indoors, how often did s/he:
47. like rough and rowdy games?
48. enjoy playing boisterous games like 'chase'?
49. enjoy vigorously jumping on the couch or bed?
77. enjoy activities such as being spun, etc.?

Impulsivity (4 items)
Speed of response initiation.

When offered a choice of activities, how often did your child
17.R stop and think before deciding?
18. decide what to do very quickly and go after it?

When encountering a new activity, how often did your child
52.R sit on the sidelines and observe before joining in?
53. get involved immediately?

Inhibitory Control (6 items)
The capacity to stop, moderate, or refrain from a behavior under instruction.

When asked NOT to, how often did your child
19.R touch an attractive item (such as an ornament) anyway?

When told "no", how often did your child
63. stop the forbidden activity?

When asked to wait for a desirable item (such as ice cream), how often did your child
88.R go after it anyway?
89. wait patiently?

When asked to do so, how often was your child able to
97. stop an ongoing activity?
98. be careful with something breakable?

Low Intensity Pleasure (6 items)
Pleasure or enjoyment related to situations involving low stimulus intensity, rate, complexity, novelty and incongruity.

While playing outdoors, how often did your child
9. enjoy sitting quietly in the sunshine?

During daily or evening quiet time with you and your child, how often did your child
20. enjoy just being quietly sung to?
21. smile at the sound of words, as in nursery rhymes?
22. enjoy just being talked to?
23. enjoy rhythmic activities, such as rocking or swaying?

When being gently rocked, how often did your child
90. smile?

Motor Activation (6 items)
Definition: Repetitive small-motor movements; fidgeting.

During everyday activities, how often did your child
4. tap or drum with fingers on tables or other objects?
8. blink a lot?

During quiet activities, such as reading a story, how often did your child
44. swing or tap his/her foot?
45. fiddle with his/her hair, clothing, etc.?
46. show repeated movements like squinting, hunching up the shoulders, or twitching the facial muscles?

When playing alone, how often did your child
80. tear materials close at hand?

Perceptual Sensitivity (5 items)
Detection of slight, low intensity stimuli from the external environment.

During everyday activities, how often did your child
7. notice low-pitched noises such as the air-conditioner, heater, or refrigerator running or starting up?
34. notice the smoothness or roughness of objects s/he touched?
71. notice changes in your appearance (such as wet hair, a hat, or jewelry)?
72. appear to listen to even very quiet sounds?

While playing or walking outdoors, how often did your child
87. notice sights or sounds (for example, wind chimes or water sprinklers)?

Positive Anticipation (5 items)
Excitement about expected pleasurable activities.

When told that loved adults would visit, how often did your child
42. get very excited?
43. become very happy?

Before an exciting event (such as receiving a new toy), how often did your child
81. get very excited about getting it?
82.R remain pretty calm?

When hearing about a future family outing (such as a trip to the playground), how often did your child
94. look forward to it?

Sadness (6 items)
Tearfulness or lowered mood related to exposure to personal suffering, disappointment, object loss, loss of approval, or response to other's suffering.

During everyday activities, how often did your child
35. become sad or blue for no apparent reason?

When told "no", how often did your child
64. become sadly tearful?

Following an exciting activity or event, how often did your child
67. seem to feel down or blue?

When s/he asks for something, and you say, "no", how often did your child
86. become sad?

When you removed something s/he should not have been playing with, how often did your child
91. become sad?

When your child was asked to share his/her toys, how often did your child
107. become sad?

Shyness (5 items)
Slow or inhibited approach and/or discomfort in social situations involving novelty or uncertainty.

When approached by an unfamiliar person in a public place (for example, the grocery store), how often did your child
2. pull back and avoid the person?
3. cling to a parent?

When approaching unfamiliar children playing, how often did your child
31. watch rather than join in?
32. seem uncomfortable?

In situations where s/he is meeting new people, how often did your child
50. turn away?

Sociability (4 items)
Seeking and taking pleasure in interactions with others.

When a familiar child came to your home, how often did your child
16. seek out the company of the child?

When visiting the home of a familiar child, how often did your child
54. engage in an activity with the child?

When a familiar adult, such as a relative or friend, visited your home, how often did your child
96. want to interact with the adult?

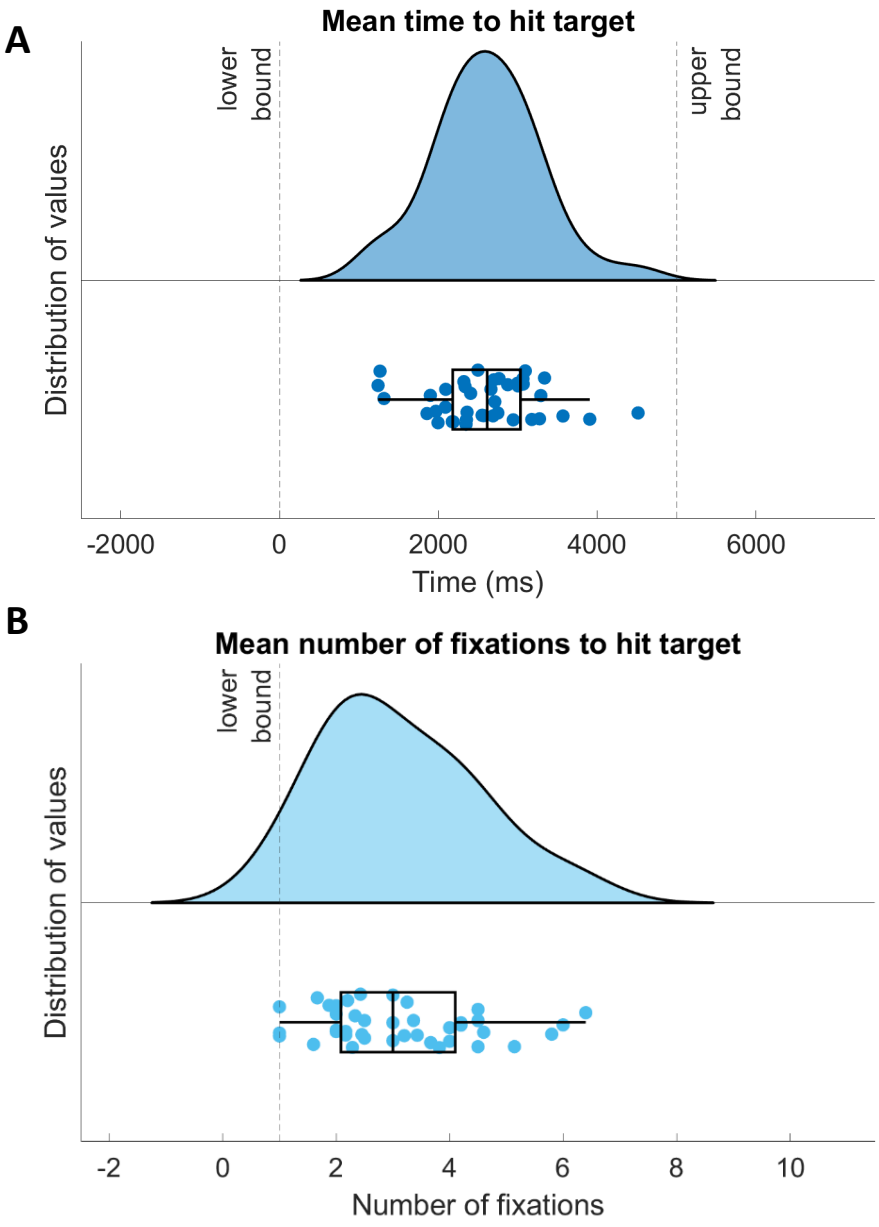
When around large gatherings of familiar adults or children, how often did your child
106. enjoy playing with a number of different people?

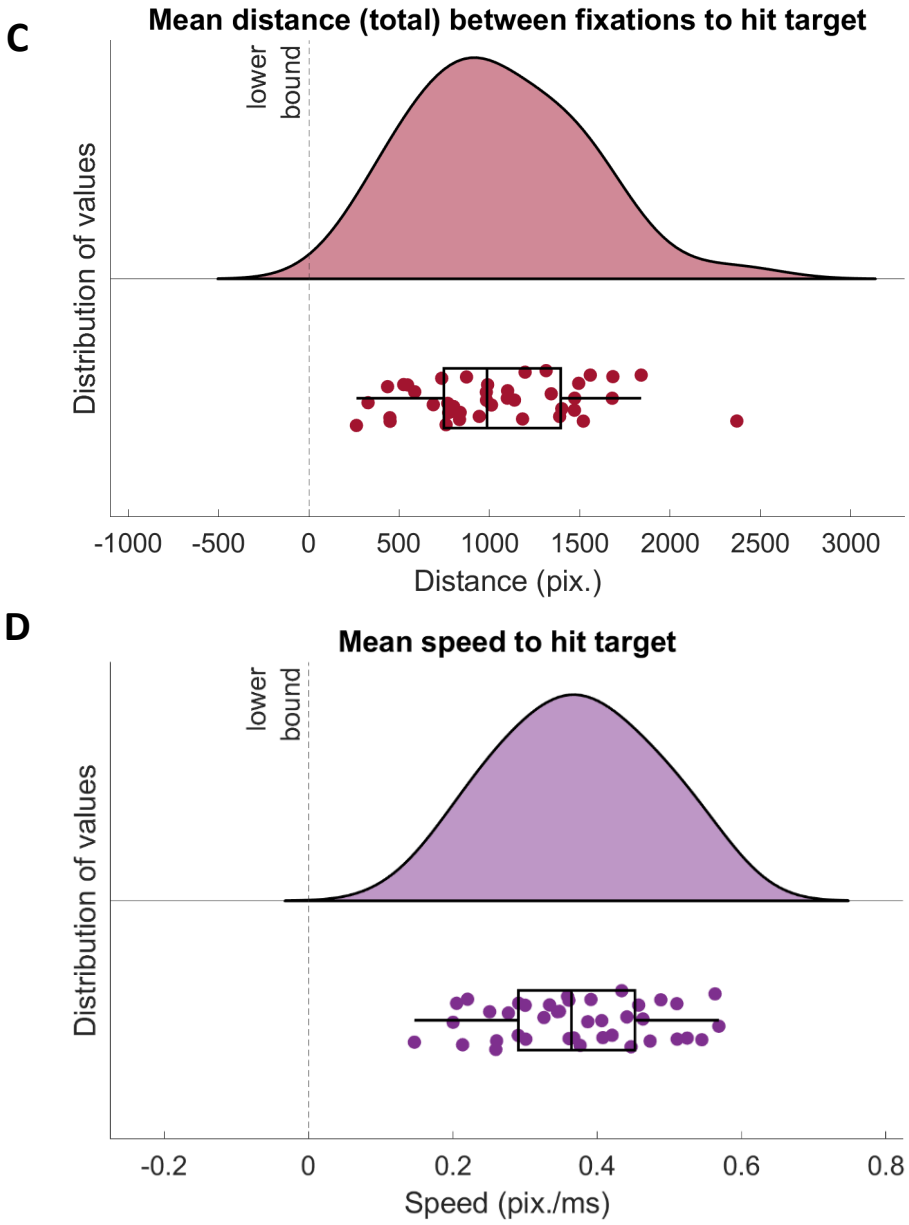
Soothability (5 items)
Rate of recovery from peak distress, excitement, or general arousal.

When s/he was upset, how often did your child
28. change to feeling better within a few minutes?
102.R cry for more than 3 minutes, even when being comforted?
103. become easily soothed?

Following an exciting activity or event, how often did your child
65. calm down quickly?
66.R have a hard time settling down?

Appendix 4: ECBQ scoring sheet.


A
Mean time to hit target
lower bound
upper bound
Distribution of values
-2000
0
2000
4000
6000
Time (ms)
B
Mean number of fixations to hit target
lower bound
Distribution of values
-2
0
2
4
6
8
10
Number of fixations



Appendix 5: Raincloud plots of the four continuous performance measures. **A.** *Time* (ms) to hit the target AOI.

B. *Number of fixations* made to hit the target AOI.

C. Total *distance* (pix.) travelled by gaze from fixation to fixation to hit the target AOI.

D. *Speed* (distance divided by time, pix./ms) to hit the target AOI.

Each dot (rain) represents a participant's data point. Their average score over all triggered trials is plotted along the x-axis, and y-axis positions are randomised to avoid overlapping points. The group's values are depicted as a distribution curve above the points (cloud) as well as a boxplot, in which the box signifies the 25% to 75% percentiles, the line inside the

box represents the median, and the whiskers depict the range, discarding any potential outlier (points outside of 1.5 times the interquartile range; non identified for this measure). Vertical dotted lines are drawn at the lower and upper bound values that constrain the measure where bounds exist.

Variable	**Stage 1**	**Stage 2**	**Stage 3**
Time to hit target	0.088 (0.599)	-0.021 (0.618)	-0.147 (0.814)
Number of fixations to hit target	0.085 (0.644)	0.045 (0.572)	-0.198 (0.791)
Distance to hit target	0.052 (0.599)	0.064 (0.473)	-0.257 (0.759)
Speed to hit target	-0.005 (0.544)	0.112 (0.358)	-0.327 (0.710)
First AOI hit is target	-0.039 (0.578)	-0.058 (0.491)	-0.045 (0.539)
Target was hit	0.094 (0.566)	-0.003 (0.521)	-0.215 (0.491)

Appendix 6: Descriptive variables – mean (S.D.) – for each performance variable at each stage.

Variable	Stage	Mean	Standard Error (S.E.)	Statistical test (t)	Statistical significance (p)
Time to hit target	1	0.088	0.157	t(40.607) = 0.564	0.576
	2	-0.021	0.157	t(40.607) = -0.134	0.894
	3	-0.147	0.157	t(40.607) = -0.939	0.353
Number of fixations to hit target	1	0.085	0.155	t(41.866) = 0.546	0.588
	2	0.045	0.155	t(41.866) =.293	0.771
	3	-0.198	0.155	t(41.866) = -1.278	0.208
Distance to hit target	1	0.052	0.143	t(40.891) = 0.367	0.716
	2	0.064	0.143	t(40.891) = 0.448	0.656
	3	-0.257	0.143	t(40.891) = -1.804	0.079
First AOI hit is target	1	-0.039	0.098	t(64.290) = -0.402	0.689
	2	-0.058	0.098	t(64.290) = -0.595	0.554
	3	-0.045	0.098	t(64.290) = -0.461	0.646

Appendix 7: Statistical significance of the marginal means of performance variables at each stage compared to zero. The two variables for which a stage value was found to significantly differ from zero are reported in text in Table 5.4.

ASQ-3 Ages & Stages Questionnaires®

15 months 0 days to 16 months 30 days (inclusive)

16 Month Questionnaire

Child's name: _____

Child's date of birth: _____ Boy ❑ Girl ❑

If child was born 3 or more weeks prematurely, please indicate the number of weeks premature: _____

Date ASQ-3 completed by parent/caregiver: _____

Date of review with health professional: _____

Child's home address: _____

_____

Town: _____ Postcode: _____

Person completing the questionnaire: _____

Relationship to child: _____

Home tel: _____ Mobile no: _____

Email address: _____

All children develop at different rates and in different ways. Please do not worry if your child is not doing all or any of the activities mentioned in the questionnaire. It is not a test. *The activities are simply one way of understanding how your child is progressing.*

Possible answers:

Yes =	your child does this activity (or has done it and has now progressed, e.g., crawling, but is now walking)
Sometimes =	your child is just beginning to do this activity (but does not do it regularly)
Not Yet =	your child has not yet started doing this

Please leave **blank** any activities your child has not been able to try with you.

BE Av. 1 1.15

16 Month Questionnaire

15 months 0 days to 16 months 30 days (inclusive)

On the following pages are questions about activities children may do. Your child may have already done some of the activities described here, and there may be some your child has not begun doing yet. For each item, please fill in the circle that indicates whether your child is doing the activity regularly (yes), sometimes, or not yet.

Important Points to Remember:

- ☑ Try each activity with your child before marking a response.
- ☑ Make completing this questionnaire a game that is fun for you and your child.
- ☑ Make sure your child is not tired or hungry.
- ☑ Please bring this questionnaire with you to your child's health and development review.

Notes:

At this age, many toddlers may not be cooperative when asked to do things. You may need to try the following activities with your child more than one time. If possible, try the activities when your child is cooperative. If your child can do the activity but refuses, mark "yes" for the item.

COMMUNICATION

	YES	SOMETIMES	NOT YET	
1. Does your child point to, pat, or try to pick up pictures in a book?	○	○	○	___
2. Does your child say four or more words in addition to "Mama" and "Dada"?	○	○	○	___
3. When your child wants something, does she tell you by *pointing* to it?	○	○	○	___
4. When you ask your child to, does he go into another room to find a familiar toy or object? (*You might ask, "Where is your ball?" or say, "Bring me your coat," or "Go and get your blanket."*)	○	○	○	___
5. Does your child imitate a two-word sentence? For example, when you say a two-word phrase, such as "Mama eat," "Dada play," "Go home," or "What's this?" does your child say both words back to you? (*Mark "yes" even if her words are difficult to understand.*)	○	○	○	___
6. Does your child say eight or more words in addition to "Mama" and "Dada"?	○	○	○	___
			COMMUNICATION TOTAL	___

GROSS MOTOR

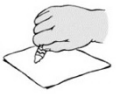
	YES	SOMETIMES	NOT YET	
1. Does your child stand up in the middle of the floor by himself and take several steps forward?	○	○	○	___
2. Does your child climb onto furniture or other large objects, such as large climbing blocks?	○	○	○	___
3. Does your child bend over or squat to pick up an object from the floor and then stand up again without any support?	○	○	○	___

GROSS MOTOR *(continued)*

	YES	SOMETIMES	NOT YET	
4. Does your child move around by walking, rather than crawling on her hands and knees?	○	○	○	___
5. Does your child walk well and seldom fall over?	○	○	○	___
6. Does your child climb on an object such as a chair to reach something he wants (for example, to reach a toy on a table or worktop or to "help" you in the kitchen)?	○	○	○	___

GROSS MOTOR TOTAL ___

FINE MOTOR

	YES	SOMETIMES	NOT YET	
1. Does your child help turn the pages of a book? (*You may lift a page for her to grasp.*)	○	○	○	___
2. Does your child throw a small ball with a forward arm motion? (*If he simply drops the ball, mark "not yet" for this item.*) 	○	○	○	___
3. Does your child stack a small block or toy on top of another one? (*You could also use cotton reels, small boxes, or toys that are about 1 inch in size.*)	○	○	○	___
4. Does your child stack three small blocks or toys on top of each other by herself?	○	○	○	___
5. Does your child make a mark on the paper with the *tip* of a crayon (or pencil or pen) when trying to draw? 	○	○	○	___
6. Does your child turn the pages of a book by himself? (*He may turn more than one page at a time.*)	○	○	○	___

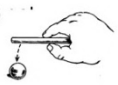
FINE MOTOR TOTAL ___

PROBLEM SOLVING

	YES	SOMETIMES	NOT YET	
1. After you scribble back and forth on paper with a crayon (or a pencil or pen), does your child copy you by scribbling? (*If she already scribbles on her own, mark "yes" for this item.*)	○	○	○	___
2. Can your child drop a raisin into a clear plastic bottle (such as a small water bottle or baby bottle)?	○	○	○	___
3. Does your child drop several small toys, one after another, into a container like a bowl or a box? (*You may show him how to do it.*)	○	○	○	___

BE Av. 1 1.15

PROBLEM SOLVING *(continued)*

	YES	SOMETIMES	NOT YET	
4. After you have shown your child how, does she try to  a small toy that is slightly out of reach by using a spoon, stick, or similar tool?	○	○	○	___
5. Without showing him how, does your child scribble back and forth when you give him a crayon (or pencil or pen)?	○	○	○	___ *
6. After a raisin is dropped into a clear plastic bottle, does your child turn the bottle upside down to tip it out? (*You can show her how.*) (*You can use a small water bottle or baby bottle.*)	○	○	○	___

PROBLEM SOLVING TOTAL ___

* If Problem Solving Item 5 is marked "yes," mark Problem Solving Item 1 as "yes."

PERSONAL-SOCIAL

	YES	SOMETIMES	NOT YET	
1. Does your child feed himself with a spoon and fork even though he may spill some food?	○	○	○	___
2. Does your child help undress herself by taking off clothes like socks, hat, shoes, or mittens?	○	○	○	___
3. Does your child play with a doll or soft toy by hugging it?	○	○	○	___
4. While looking at himself in the mirror, does your child offer a toy to his own image?	○	○	○	___
5. Does your child get your attention or try to show you something by pulling on your hand or clothes?	○	○	○	___
6. Does your child come to you when she needs help, such as with winding up a toy or unscrewing a lid from a jar?	○	○	○	___

PERSONAL-SOCIAL TOTAL ___

OVERALL

Parents and providers may use the space below for additional comments.

1. Do you think your child hears well? If no, explain: ○ YES ○ NO

BE Av. 1 1.15

OVERALL *(continued)*

2. Do you think your child talks like other toddlers his age? If no, explain: ◯ YES ◯ NO

3. Can you understand most of what your child says? If no, explain: ◯ YES ◯ NO

4. Do you think your child walks, runs, and climbs like other toddlers her age? If no, explain: ◯ YES ◯ NO

5. Does either parent have a family history of childhood deafness or hearing problems? If yes, explain: ◯ YES ◯ NO

6. Do you have concerns about your child's eyesight? If yes, explain: ◯ YES ◯ NO

7. Has your child had any medical or health-related problems in the last few months? If yes, explain: ◯ YES ◯ NO

BE Av. 1 1.15

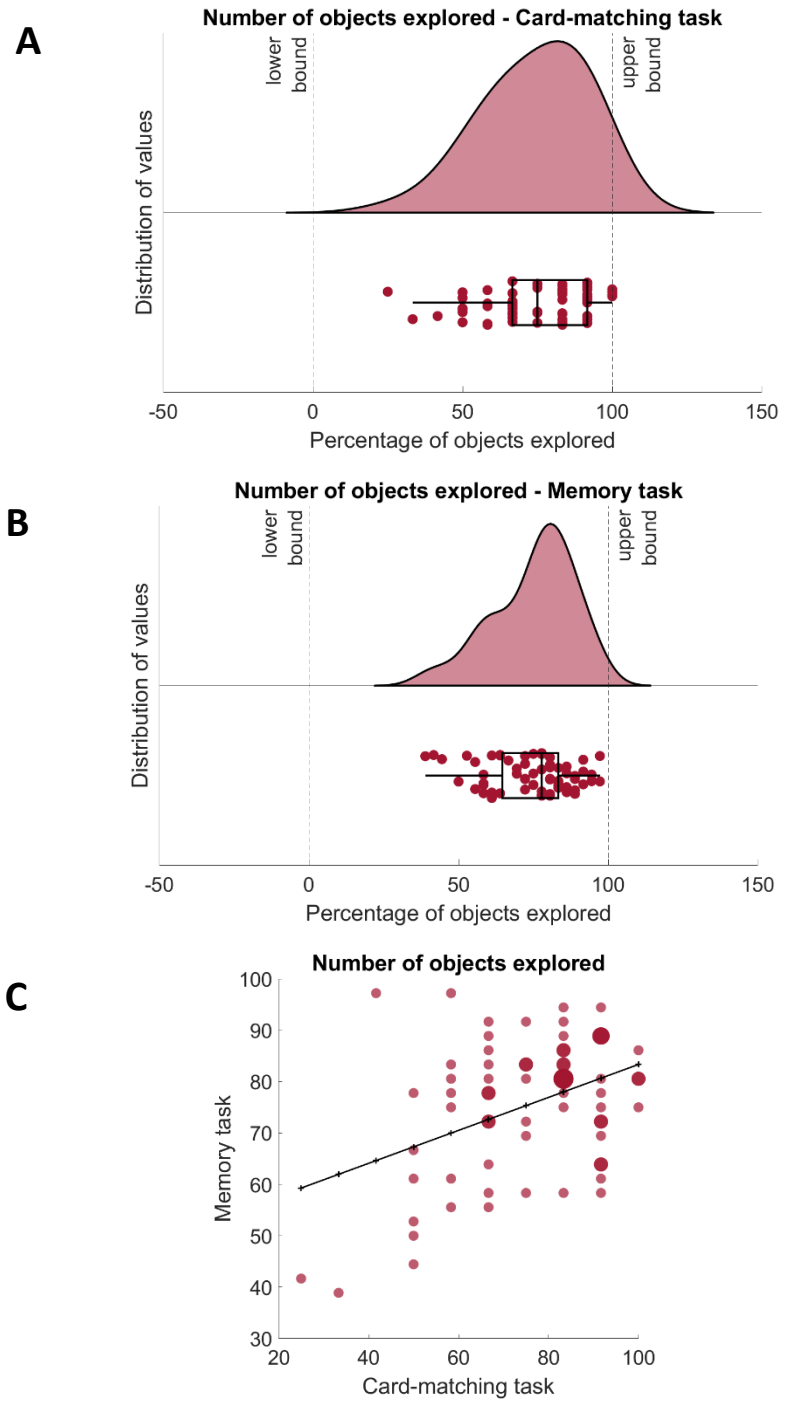
OVERALL *(continued)*

8. Do you have any concerns about your child's behaviour? If yes, explain: ◯ YES ◯ NO

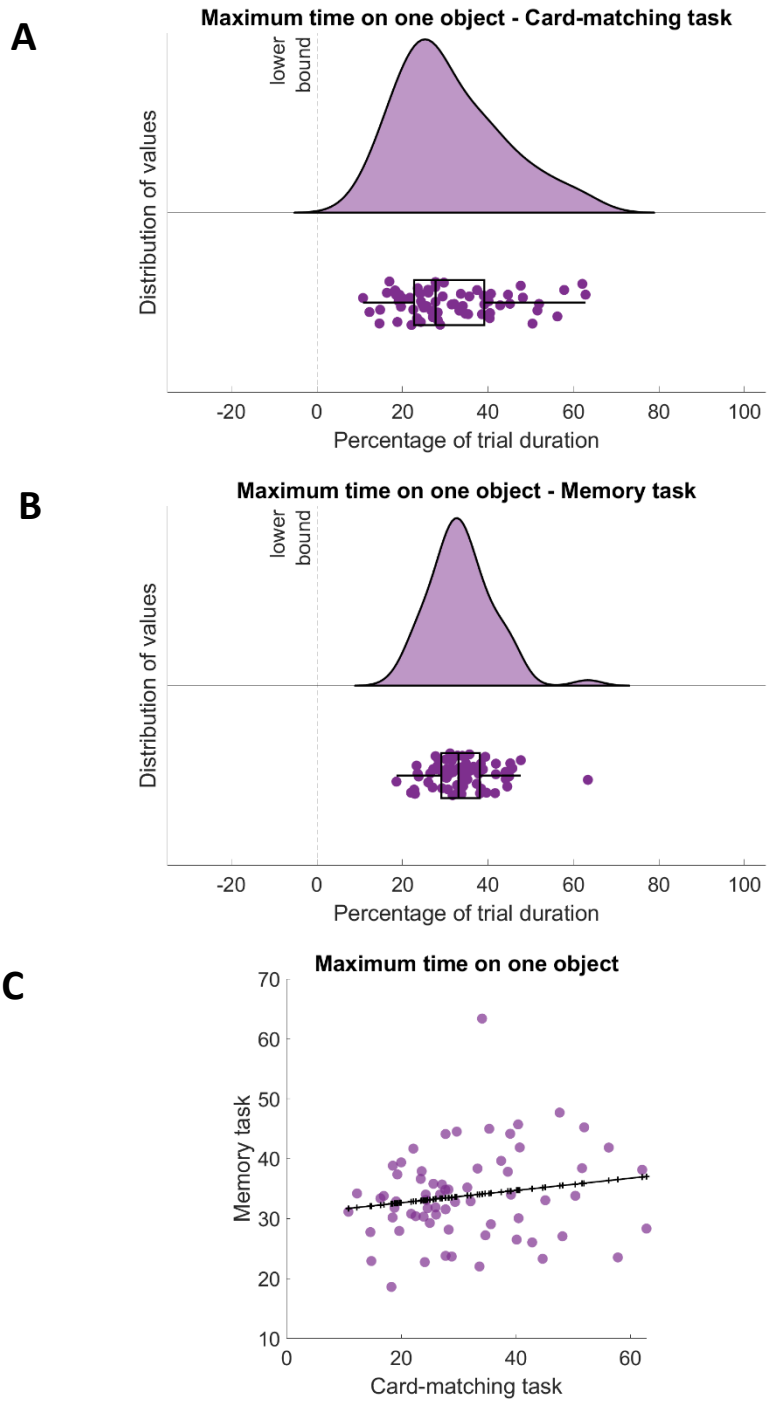
9. Does anything about your child worry you? If yes, explain: ◯ YES ◯ NO

BE Av. 1 1.15

Appendix 8: Ages and Stages Questionnaire (ASQ) for 16 month old infants.



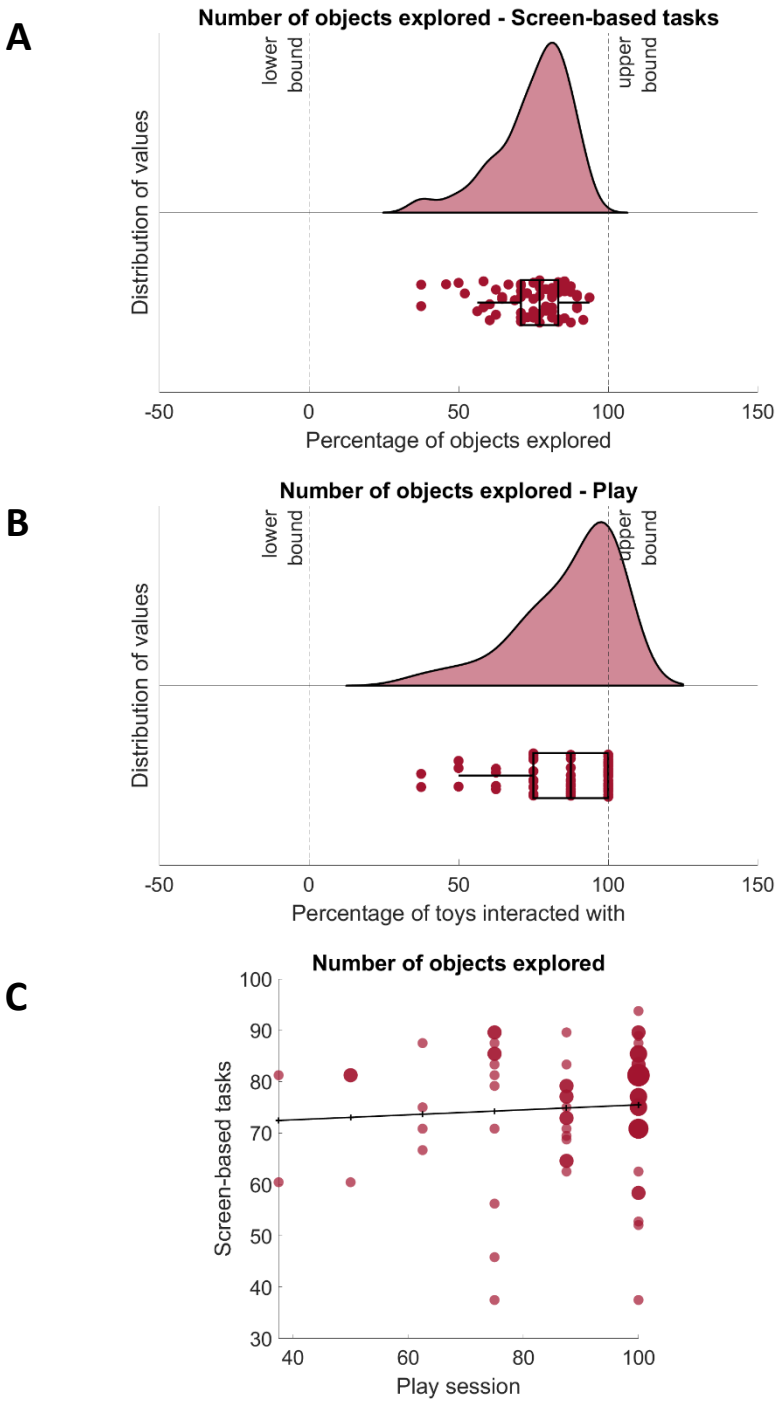
Appendix 9: Raincloud plot for the *number of objects* exploratory scores of the card-matching (**A**) and the memory (**B**) tasks, as well as the relationship between the two (**C**). All plots include potential outliers. See Figure 6.10 and Figure 6.8 for a detailed description of the elements of the rain-plots and correlation plots, respectively.



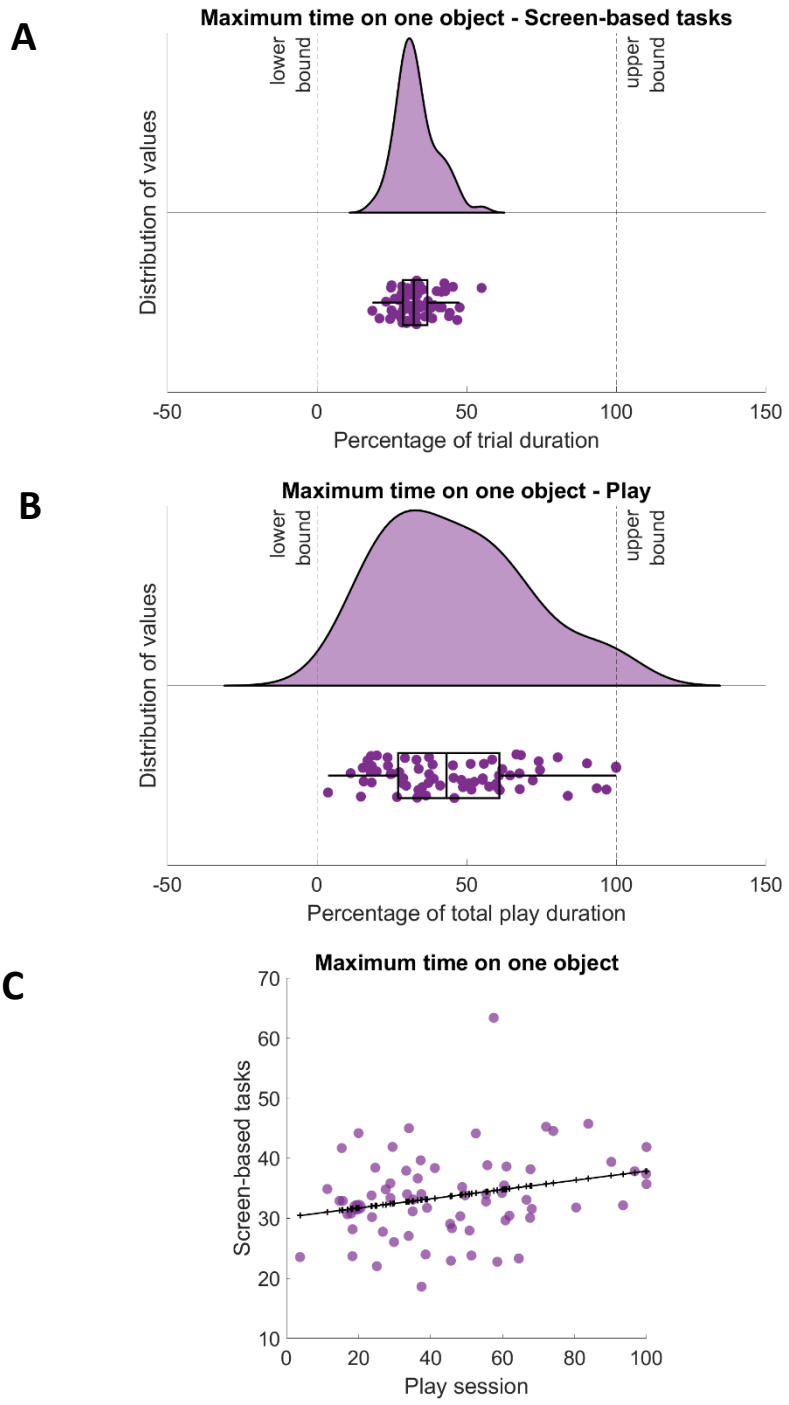
Appendix 10: Raincloud plot for the *maximum time* exploratory scores of the card-matching (**A**) and the memory (**B**) tasks, as well as the relationship between the two (**C**). All plots include potential outliers. See Figure 6.10 and Figure 6.8 for a detailed description of the elements of the rain-plots and correlation plots, respectively.

Variable	Mean (S.D.) card-matching task	Mean (S.D.) memory task	Correlation (Pearson's R)	Statistical significance (p)
Number of objects (%)	73.9 (17.0)	75.1 (13.5)	**0.402**	**< 0.001**
Maximum time (%)	31.7 (12.5)	33.6 (7.31)	0.168	0.175

Appendix 11: Mean (Standard Deviation, S.D.), Pearson's R value, and statistical significance of the correlations between the card-matching and the memory task exploration measures *keeping all outliers*; each row presents the results for one exploratory variables. Significant results are highlighted in grey cells and bold font.



Appendix 12: Raincloud plot for the *number of objects* exploratory scores of the card-matching (**A**) and the memory (**B**) tasks, as well as the relationship between the two (**C**). All plots include potential outliers. See Figure 6.10 and Figure 6.8 for a detailed description of the elements of the rain-plots and correlation plots, respectively.



Appendix 13: Raincloud plot for the *maximum time* exploratory scores of the card-matching (**A**) and the memory (**B**) tasks, as well as the relationship between the two (**C**). All plots include potential outliers. See Figure 6.10 and Figure 6.8 for a detailed description of the elements of the rain-plots and correlation plots, respectively.

Variable	Mean (S.D.) screen-based tasks	Mean (S.D.) play	Correlation (Pearson's R)	Statistical significance (p)
Number of objects (%)	69.9 (19.7)	90.0 (16.4)	-0.064	0.593
Maximum time (%)	31.0 (9.15)	46.2 (24.1)	**0.888**	**< 0.001**

Appendix 14: Mean (Standard Deviation, S.D.), Pearson's R value, and statistical significance of the correlations between the card-matching and the memory task exploration measures *keeping all outliers*; each row presents the results for one exploratory variable. Significant results are highlighted in grey cells and bold font.