Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Communication-induced and Spontaneous Object Representations in Infancy

DISSERTATION

zur Erlangung des akademischen Grades
DOCTOR RERUM NATURALIUM

an der Universität Hamburg
Fakultät für Psychologie und Bewegungswissenschaft
Institut für Psychologie

vorgelegt von
Wiebke Pätzold

Hamburg, Mai 2018

Promotionsprüfungsausschuss:

Vorsitzende:	Prof. Dr. phil. Juliane Degner
1. Dissertationsgutachter:	Prof. Dr. rer. nat. Ulf Liszkowski
2. Dissertationsgutachter:	Prof. Dr. rer. nat. Jan Wacker
1. Disputationsgutachter:	Jun.-Prof. Dr. rer. nat. Robert Hepach
2. Disputationsgutachterin:	PD Dr. rer. nat. Kirsten Hötting

Tag der Disputation: 3. September 2018

Acknowledgements

I would like to thank my supervisor Prof. Ulf Liszkowski, for his vision for my project and for always having "a minute" (or a hundred) when I needed discussion. Also, thank you for encouraging – and funding – my travels to many conferences, summer schools and workshops. The contacts and friendships cultivated through those meetings really broadened my horizon. Among those contacts, I would like to especially thank Robert Hepach, Eugenio Parise, Sylvain Sirois and Iain Jackson, who generously answered my probing questions during poster sessions, Skype calls and follow-up e-mail exchanges.

Thank you to my research group in the Developmental Psychology department. Whether it was during long days in the EEG lab, teaching together, organizing our "Stammtisch" or on adventures at various conference sites – last-minute print jobs in Budapest or the lost posters of Vienna – it was great to have you as companions. Johanna, Marianna, Sebastian and Ranjani, thanks for the ride! Also, thank you to all of the students who worked with us over the years, be it as research assistants, master students or interns. Our work would not be possible without your commitment.

Thank you also, in completely random order, to: Jessica, for playing the role of the "rubber duck"; Claudi, for her pure honesty; Andrea, for believing not only that I could do it, but that I could do it well; Susi, for showing up when I needed it; Lena, for long-distance therapy; Margarita, for giving me a glimpse of the glamorous side of academia.

Thank you to Susanna, for always having two open ears and one wide open heart from my very first day, and, through life changes and job changes, still being there for me.

Thank you to Nina, for giving me "a roof over my head" to finish writing in a productive environment – that is, in the office next to hers.

I would also like to thank my parents for supporting my endeavor in every possible way and always remaining positive and reassuring when things didn't go as planned.

And last but not least, thank you to Johannes for his steadfast support throughout it all. From encouraging me to pursue research, to coding scripts for me in any environment I could wish for (but always trying to get me into Python), to delivering ice cream after exhausting days, I couldn't have asked for anyone better to share this time with.

Contents

List of Figures

List of Abbreviations

in alphabetical order

AMS	analog magnitude system	31
ANS	autonomic nervous system	43
EEG	Electroencephalography	56
EW	Edinger-Westphal nucleus	44
FFT	Fast Fourier transform	57
LC	locus coeruleus	48
OTS	object tracking system	31
PDR	pupil dilation response	51
PLR	pupillary light reflex	44
ROI	regions of interest	136
SC	superior colliculus	48
ToM	Theory of mind	9
VOE	violation-of-expectation	22
WT	wavelet transform	42

Abstract

At the end of their first year, infants start to engage in meaningful, if nonverbal communication with their caregivers. At the same time, they appear to show sophisticated understanding of physical objects and their continued existence during occlusion. Many studies have brought forward evidence of this early conceptual understanding of referentiality and object permanence, but have remained vulnerable to the critique of supporters of leaner, non-mentalistic explanations of infant behavior. The goal of this thesis was to investigate how infants at the end of their first year process the referential content of social-communicative cues, and represent objects during occlusion, using (neuro-)physiological measures that are more resistant to low-level perceptual accounts than traditional behavioral measures.

Therefore, two different methodological approaches were taken: On the one hand, pupillometry was used to measure cognitive load during the presentation of socially meaningful scenes and surprising occlusion-related events. On the other hand, EEG was used to find neuro-correlates of object representation in response to social and nonsocial cues. In particular, increase in gamma band activity was interpreted as a marker for object maintenance.

Three studies explored infants' comprehension of social-communicative cues and object representation. In the first two studies, pupil dilation was measured to investigate expectation elicited by pointing (Study 1) and expectation elicited by an occlusion event (Study 2) in violation-of-expectation paradigms. In Study 1 (Chapter 2), I found that infants expected an object to appear after they had seen an agent point towards the occluder at 12 months, but not at 8 months, and not after a non-social control cue. In Study 2 (Chapter 3), I found that 18-month-olds, but not 10-month-olds, expected an object in a nonsocial occlusion experiment. In Study 3 (Chapter 4), I measured activity in the EEG gamma band to investigate pointing comprehension and spontaneous object expectation in two experiments. Infants saw an occlusion event followed by a cue which was either social-communicative or nonsocial (Experiment 1) or social-communicative or social-noncommunicative (Experiment 2). In the first experiment, I was able to establish the previously reported object maintenance effect and a new response pertaining to the communicative cue in 12-month-olds. In the second experiment, I found the object maintenance effect only in the social-communicative, but not in the social-noncommunicative control condition, in 10-month-olds.

The findings of Study 1 support the hypothesis that infants understand the referential content of communicative cues, like declarative pointing, around their first birthday. The divergence of the results between Study 1 and Study 2 led me to suspect that object representation may not be independent from social cues. The findings of Study 3 further emphasize the idea that cognitive processing of object occlusion events may be influenced by the communicative context in which they occur.

1 Introduction

> *"There is a gap between the mind and the world, and (as far as anybody knows) you need to posit internal representations if you are to have a hope of getting across it. Mind the gap. You'll regret it if you don't."*
>
> – Jerry A. Fodor

We are, as human adults, capable of remembering, thinking ahead and interacting with other people in a meaningful way. To do this, we must be able to have mental representations of all sorts of things: where we put that folder with the participant data, what time would be best to schedule the dentist appointment next week, what our colleague already knows about our new experiment.

But how do these cognitive functions arise in the human brain during development? When do infants develop mental representations? One theme of this thesis is the development of two aspects of mental representations: The representation of objects induced by referential communication and the spontaneous representation of objects following occlusion events. In the following introduction, I will present theories and empirical data for both branches separately. In the studies, referential cues and object cognition will be considered both separately and together, and in the discussion, I will address their unique interplay during development.

The other theme is the age-old question of nature versus nurture: Is the ability to have mental representations ingrained in us from birth, through our genetic make-up? Or are mental representations constructed, maybe socially, through our upbringing? Nativist theories posit that cognitive abilities are inherent to our nature. Social-constructivist theories however presume very little hereditary functionality, but rather suggest that what makes us thinking humans is our experience within human culture. Both positions, as well as current evidence, will be reviewed in order to devise specific hypotheses about infants' mental representations at the end of their first year.

1.1 The Development of Referential Understanding

One of the cornerstones of human identity is our ability to communicate with each other. While many other species use forms of communication to warn, distract, or attract mates (Bradbury & Vehrencamp, 1998), humans are the only species to use communication to mutually share interest in something and engage in a conversation about something that is not necessarily goal-directed (Tomasello, 2014). Language is our most precious vehicle for communication – an enormous phylogenetic and ontogenetic cultural accomplishment. Thousands of distinctly evolved languages exist in the world (Katzner, 2011) and for a good reason: Languages enable us to communicate about absent entities by means of having words that stand in for actual things in the world. They allow us to directly transplant our own mental representations into others. From a developmental point of view, therefore, it appears that we need to start with language acquisition when we want to learn about how infants learn to communicate with others.

However, language is not where communication starts, and language by itself is not sufficient to allow for meaningful conversation. Words are meaningful only when used in a meaningful context. In writing, the context arises from the redundancy in the company of other words to transport the intended meaning. But it is also entirely possible to communicate in a meaningful way about something while using completely ambiguous language. Say, we are having a conversation about trees falling in a storm, and the risks of having certain types of trees close to houses. During this conversation, I step to the window and point outside, at the tree next to the house, and say "like this one". These three words are meaningless without context. Only by being having paid attention to the whole situation can a conversational partner understand what is meant. This phenomenon is called the "utterer's meaning" (Grice, 1957) which arises from the richness of context the speaker, or utterer, has created. It is something he or she says to convey *intention*, without directly stating what their intention is. But within the context of the preceding conversation and the pointing gesture directed at the tree, it is clear that this tree right there is at risk of causing damage to the house when the next storm comes. This is because the context, the words and the gesture created a meaningful *referentiality* shared between the participants of the conversation.

1.1.1 Origins of Communication

How do humans learn to understand the referentiality of words and gestures in a communicative interaction?

On the one hand, we can look at the phylogenetic development of communication as a part of the evolution of *Homo sapiens*. The oldest evidence of modern humans was recently found near Jebel Irhoud, Morocco, and dates back 300,000 years (Hublin et al., 2017). Based on the analysis of the development of phonemic diversity, verbal language is estimated to have evolved 350,000 – 150,000 years ago (Perreault & Mathew, 2012). According to Perreault and Mathew, this means the emergence of language happened early in the history of *Homo sapiens*, or may even precede it, appearing first in ancestors of modern humans. We may assume that the ability to convey meaningful information precedes the rise of actual language (Tomasello, 2008). The development of humankind, therefore, can hardly be described without the development of human communication. Verbal language, as shown above, is just one part of having a meaningful conversation. Humans have also used their surroundings to transport meaningful content for a long time. The first evidence of symbol use on physical objects can be traced back to western Asian and northern African regions 10,000 years ago through archaeological findings in the shape of geometric tokens, while the first evidence of written language dates back a mere 5,400 years ago to discoveries in Sumer, where markings on clay were first used to represent conventions for the sounds of the human voice (Powell, 2009, p. 11).

On the other hand, we can look at the ontogenetic development of communication. Newborn infants imitate their social partner (Meltzoff & Moore, 1977), prefer visual stimuli that are arranged to resemble facial features (Goren, Sarty, & Wu, 1975), recognize their mother's voice (DeCasper & Fifer, 1980), prefer to listen to the language they have been exposed to in utero (Moon, Cooper, & Fifer, 1993) and preferentially look at faces that show direct eye gaze compared to averted eye gaze (Farroni, Csibra, Simion, & Johnson, 2002). By three months, they recognize pictures of their mothers compared to that of strange women (Barrera & Maurer, 1981). In an interplay between a developing motor system, neuronal pruning and input from attentive caregivers, infants soon begin to show signs of active social engagement. In their seminal work on *joint attention*, Scaife and Bruner (1975) were the first to show that as early as two months of age, infants follow the gaze of their mother to an object in a controlled setting. By six months, they spontaneously follow adults' gaze in an effort to join in shared visual attention (Adamson & Bakeman, 1991; Butterworth & Cochran, 1980; D'Entremont, Hains, & Muir, 1997), and, around the first birthday, infants start to participate in what is sometimes called joint engagement, or triadic interaction: The mutual engagement of infant and caregiver with an object (Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004; Liszkowski, Carpenter, Striano, & Tomasello, 2006; Tomasello, Carpenter, Call, Behne, & Moll, 2005).

This *joint engagement* may be the holy grail of communication: At the very least two parties, jointly attending to the same topic (Tomasello, 2008). In the example above, there would be a joint engagement between two people, both attending to each other as well as the topic, the tree. Here it is where the *common ground* arises (Clark, 1996): I know, that you know, that I am attending to this object. And you know, that I know, that you are attending to this object. Therefore, we are both together attending to the same object (Tomasello, 2014).

1.1.2 Theories of Social Development

What is it that allows infants to start interacting and communicating with others so readily? Philosophers and psychologists alike have tried to come up with answers for a long time. In the following, I will introduce a few theories that have shaped the way psychologists see infants' social-cognitive development.

Drawing from anecdotal evidence, one of the first psychologists to write about infants' understanding of self and other, Baldwin (1884) formulated an early social-cognitive approach to development, in that the understanding of others' minds was fundamental to achieving common goals. His work made an impression on a young Piaget, who would go on to publish influential, but also rigid views on child development focusing on the egocentrism of the infant, which is gradually overcome through socialization (Piaget, 1954, pp. 352). In contrast to Baldwin (1884), Piaget (1954)'s infant did not start off with understanding of the other's mind.

Vygotsky (1981, first published posthumously in Russia in 1960) argued that higher mental functions develop through social interactions. By observing and practicing actions that are beyond the child's behavioral scope, the actions are internalized. Therefore, it is the world around the child, specifically the people and culture in which he or she grows up, that allows mental functions to develop. Vygotsky explained infants' early gestural communication as follows: Having seen how other people grasp objects, infants perform unsuccessful grasps towards an object that is out of reach. The caregiver then hands over the object in an effort to help the child reach their goal. Having received reinforcement for the movement, the child uses the gesture again to get other objects. This purely imperative use of the pointing gesture has been refuted in several studies (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Franco & Butterworth, 1996), which proved that infants' grasping and pointing gestures tend to be distinct and not be confused with each other. However, Vygotsky's account was one of the first *constructivist* accounts of development: Interaction with others, and input from caregivers, enables infants to gradually construct knowledge of the world around them. Today, advocates of the social-constructivist approach (e.g. Carpendale & Lewis, 2004) still argue in favor of social construction of mental representations. They and Vygotsky, therefore, can be positioned on the far end of the nurture side of social-cognitive development.

Werner and Kaplan (1963) introduced an approach that allowed for some innate referentiality, but still relies on social interaction for pre-lingustic communicative behavior to become fully meaningful. They suggest that pointing initially arises as a communicative gesture for the self, helping the child hold attention to the object of interest. Through social interaction, pointing then becomes ritualized to extend to the sharing of interest with others. Today, their approach on symbol formation is still being applied to questions of development, for example in Levy and McNeill (2013)'s theory on the development of narratives and imagery (see also Glick, 2013).

It was not until Bruner (1975) that the cognitive revolution arrived for theories on infant development. Bruner's objective was to find an alternative for the nativist account of language acquisition proposed by Chomsky (1957, 1966) and Fodor (1983), and by formulating his social constructivist theory of language development, he developed a theory on what underlies language in terms of referentiality and meaning in the tradition of Grice (1957) and Searle (1969). In contrast to earlier theorists, he presented empirical evidence to demonstrate these underlying elements in preverbal children, thus suggesting that language builds upon an understanding that is already there.

Bruner built on Grice's definition of the "utterer's meaning" in relation to its occasion of use, his or her intention. He compared the process of channeling visual attention to the way language is structured: As we visually scan a scene, we alternate between paying attention to details, and seeing the big picture. Language is structured in a way that allows description of exactly this process because of its *topic and comment* structure, with the topic being the whole or context and the comment being the features. Bruner used this hypothesis to make the case that infants learn to make conceptual distinctions embodied pre-linguistically in what he called *case grammar*. This proto-grammatical structure contains, among others, categories of agent, action, object of action, recipient of action, location and possession, not unlike the core knowledge system later proposed by Spelke (1990). This case grammar is innate to infants and enables them to start using signals to steer their social partner's attention. According to Bruner, pointing is one of the first forms of that precursor of language, and allows infants to build simple "sentences" using the point to stand in for actions, objects of actions, or other elements of case grammar. At first, only the caregiver will have enough context information to understand the child's intention, but as infants learn to express themselves more and more, their signals become more conventional, until they start using verbal language.

Bruner's objective was that of explaining how infants learn language. But by doing so, he entered uncharted waters in terms of infant cognition: He ascribed *intentions*, and more importantly, *comprehension of other's intentions*, to preverbal infants. Therefore, in many ways his work is the starting point of modern cognitive theories on infant development.

Today, far from the tabula rasa idea of earlier times, most theorists agree that infants possess an early specialized adaptation to social cues (Meltzoff & Moore, 1998; Spelke & Kinzler, 2007), putting them on the *nativist* side of the spectrum. In the following, I will focus on two theorists in particular who have posited humans' attunement to cues from social partners as the crib of social-cognitive development.

Tomasello (2008)'s theory builds on the notion first described by Grice (1975) that the main motive of communication for humans is to collaborate with conspecifics. He proposes that humans, in contrast to all other great apes, are collaborators first, and aspire to share with and inform others (Tomasello, 2008, p. 5). With the first collaborative activities, joint attention evolved in our ancestors, beginning with a first common ground in which two social partners can be sure to have the other's attention in order to perform a task together (Tomasello, 2014, p. 44). In order to go from a simple joint activity to successful, possibly verbal, communication, a common *conceptual* ground is needed: To know that, not only does the social partner have knowledge about the same thing that I do, but that he knows the same about my knowledge, and we are aware of our mutual understanding (Tomasello, 2008, p. 5). From this common conceptual ground arises what may be called *shared intentionality* (Searle, 1995).

Tomasello (2008) argues that this is why, from a young age, infants are adapted to seek and thrive in joint attentional activities. The first actions of human cooperative communication came in the form of pointing and iconic gestures to direct attention to something perceptually available to the recipient, through attention-direction gestures like head– or finger points. Then, the next step was to direct the imagination of the recipient to something not immediately available through iconic gestures; behaviorally simulating an action (Tomasello, 2008, p. 61). The aim of both types of referential gestures was to cause the recipient to infer the communicator's intention. Therefore, for communication to be successful, both social partners need to be able to attribute mental states to the other (Tomasello, Carpenter, & Liszkowski, 2007).

A second influential current theory is Csibra and Gergely's *Natural Pedagogy* (Csibra & Gergely, 2009; Csibra, 2010; Csibra & Gergely, 2011). Similar to Tomasello, they also describe joint attention as the root of communication. In their view, infants are efficient social learners, enabling them to acquire cultural knowledge, such as language, at a rapid pace. The reason for this rapid enculturation are *ostensive cues* provided by the adult communicative partner: Direct eye contact, the speech contour found in infant-directed speech, and contingent responsivity. While infants have biologically adapted to attend to and learn from them; adults have adapted to naturally communicate ostensively towards infants (Csibra & Gergely, 2009; Csibra, 2010). However, there is no need for the infant to grasp the other's communicative intent in order to understand and use the communicative cue. Instead, Csibra and Gergely posit that infants develop referential expectations towards objects whenever they are in an ostensive context. That means infants expect ostensive cues to be "about" something. Adults' shifts in eye-gaze, gesturing to or handling of objects mark this "aboutness" for the infant toward the referent, and infants' tendency to follow their eye. Hand movement signals a readiness to process an interesting artifact. Studies in support of this view will be discussed in Section "Current Findings from the Study of Referential Understanding".

1.1.3 Lean versus Rich Interpretations of Social Cognition

The two cognitive theories of development described above can be summarized as rich views on infant development because they ascribe early referential understanding and therefore mental representations to the infant. These theories can be further subdivided in how much of an understanding they ascribe to the infant. The richest interpretation would be that of understanding the infant's brain as much the same as the adult's, being able to juggle different representations of mental states of self and other at the same time, as implied by Bruner (1975). Bretherton (1991) also expressed this view and ascribed an "understanding of mind" to the preverbal infant who communicates through gestures, guiding the adult's attention to achieve a goal. A form of an early Theory of mind (ToM) is also implied in Tomasello's theory of infant's understanding of the other's intention, in which infants use pre-linguistic communication in an attempt to influence the intentional and mental states of others (Tomasello, 2008; Tomasello et al., 2007). Others have been even more explicit about the mind-reading abilities of infants, ascribing them with innate modules for automatically interpreting another's perspective (Kampis, Parise, Csibra, & Kovács, 2015; Kovács, Téglás, & Endress, 2010; Onishi & Baillargeon, 2005), or being able to discriminate self from other from birth (Meltzoff, 2007).

Other theories can be described as moderately rich, because while they ascribe mental representations to infants, they describe them as being less complex and more dependent on perceptual processes than adults' (Csibra & Gergely, 2009; Perner, 1991). They mostly explain joint interaction without the infant needing to attribute a mental state or intention to their social partner. Instead, these theories focus on the mental representations evoked in the child during such an interaction.

On the opposite side of the spectrum are critics of any mental ascriptions to infants from birth (Butterworth & Jarrett, 1991; C. Moore & Corkum, 1994; Heyes, 2014b). They refuse to describe behaviors, physical and neurological reactions as evidence of mental processes when much simpler explanations, like learned behavior, suffice to account for them (C. Moore, 1996; C. Moore & D'Entremont, 2001; Shatz & O'Reilly, 1990). Their reasoning can be demonstrated using the example of comprehension of an infant watching an adult point towards a box. In the cognitive interpretation, the infant immediately understands the communicative intent of the pointer, and uses the cue to find an interesting thing that he or she has come to expect. However, the leaner interpretation of the same scenario is that the infant has been conditioned, through interaction, to see the facial expression and gesture as an attentional cue, which will lead to a reward if followed. The infant does not, however, understand the pointer's intention, and he or she does not form an expectation of a referent. By following the cue, she discovers the object, which is designed by the experimenter to be interesting. This is why he or she picks it up. There is no need for the infant to mentally represent something beforehand.

Butterworth and Jarrett (1991) suggested an account for gaze following that does not require mental representations on part of the infant. They identified three mechanisms of increasing complexity that enable infants to follow an adult's gaze from infancy to the second year of life. At first, an *ecological mechanism* allows the infant to follow the gaze because both adult and infant possess similar visual systems whose attention is likely to be captured by the same objects or events. This allows gaze following within the infant's field of vision, but not the precise localization of a target. Around 12 months of age, this changes as infants employ a *geometric mechanism* which enables them to identify the adult's line of sight in order to find the target. Only later, towards 18 months of age, are infants able to follow gaze outside their current visual field, using a "representational" mechanism which allows them to infer a target that they do not see. However, even with this most advanced mechanism, there is no need to attribute intentions to the other person. In Butterworth and Jarrett's view, therefore, mental representations are not needed for communication in the first two years of life. However, they highlight the importance of preverbal communication between caregiver and child has for later language acquisition ("Pointing is the royal road to language for babies", Butterworth, 2003).

C. Moore and Corkum (1994) disagreed with the early gaze following evidence presented by (Bates, Camaioni, & Volterra, 1975; Butterworth & Cochran, 1980, and others) and postulated that robust evidence of gaze following appears relatively late in infancy (around 18 – 24 months of age, see C. Moore & Corkum, 1994; C. Moore & D'Entremont, 2001). They therefore argued in favor of an interactionist account. In their view, joint attention is learned largely through operant conditioning (C. Moore & Corkum, 1998). The caregiver turns their head, which distracts the infant from whatever he or she is doing. If the infant turns to look at any other direction (or does not react), nothing happens. However, if the infant turns in the caregiver's direction, he or she is rewarded by the sight of an interesting object. Over time, the infant becomes much more likely to turn in the same direction as the caregiver. This is the simplest scenario, in which the caregiver does not explicitly foster a reaction by the infant. In real-life settings, it is much more likely that the caregiver's response will be further rewarding for the child and form a kind of "scaffolding" that allows for more complex responses (C. Moore & Corkum, 1994, p. 354). They criticized the "richer" cognitive interpretation, that in order to follow an adult's communicative cue, the infant has to have a representation of the adult's goal or intention, or even their point of view (C. Moore & Corkum, 1994, p. 351). Attributing this kind of secondary representations to the infant, they claimed, poses problems when explaining deficiencies in ToM much later in childhood. Instead, they argued in line with the cognitive theory introduced by Perner (1991), which allows the infant to entertain one representation, or "model" at a time. Whenever there is an update to the model, the previous one gets overwritten. It is not until the second or third year of life, that a "multiple modeling capacity" arises, which allows the child to have several representations at the same time. According to C. Moore and Corkum, therefore, an infant at the end of the first year, while capable of following an eye gaze or a pointing gesture, does not possess the ability to attribute referential intention to the gesturer.

C. Moore and Corkum did not oppose the existence of innate, evolved mechanisms for the orienting response towards the head in the first place (1994, p. 355), but emphasize the role interaction plays in acquiring the skill. In a training study, they showed that infants naturally oriented their head starting only around 10 months of age, but were able to be "trained" to do so starting from around 8 months of age (Corkum & Moore, 1998, Experiment 1). However, while a head-turn response in the direction of an adult's eye gaze could be conditioned in 8– 9 month-old-infants, it was impossible to train infants to look in the opposite direction, even with rewards. Rather, even infants who were never rewarded to follow the model's eye gaze did so (Corkum & Moore, 1998, Experiment 2). The authors offered two possible explanations for this behavior: The head orientation contains powerful informational characteristics which cannot easily be overridden, which would be a nativist account, or the infants have understood there to be a correct and incorrect side at which to look for the reward, and use the head turn as an attentional cue. Whatever it is that drives the effect of the cue, they stressed the interactionist nature of their model: Interaction with others and objects allows infants to develop a self-other-discrimination, and representations of other's intentions and goals. Weakening the empirical evidence for this account, a recent longitudinal study on pointing production and gaze following showed no effect of additional pointing experience on pointing production in 9– to 11-month-olds (Matthews, Behne, Lieven, & Tomasello, 2012).

Carpenter (Carpenter & Call, 2013; Carpenter et al., 1998) used evidence from a longitudinal study and several behavioral studies to make the claim that infants are representing their social partner's knowledge or preference for an object in joint attention tasks beginning around their first birthday. Infants selectively repair adult misunderstandings about the referent of their declarative points at 12-month-old (Liszkowski, Carpenter, & Tomasello, 2007b), use pointing to refer to objects that are no longer there, but that were relevant to an earlier interaction with a social partner (Liszkowski, Carpenter, & Tomasello, 2007a) and, by 18 months of age, point selectively to objects that are specific to a previously shared experience with a particular person (Liebal, Carpenter, & Tomasello, 2010). However, Carpenter and Call cited mostly broad, apparent behavior in the form of gaze checks, smiling, insistence or complaints in response to a more or less complex social situation as evidence for engaging in joint attention, and rejected evidence from implicit measures showing evidence of joint attention in younger infants, who are not yet capable of producing the necessary social behavior. Their account, therefore, is also interactionist.

Heyes (2014b) took on an extreme point of view, claiming that not only do infants fail to represent mental states of others, but for the most part, adults do as well. Contrary to accounts of "automatic perspective taking" in adults (Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, Sarah E., 2010), Heyes (2014b) argued that adults do not automatically take into account another's perspective, but that simple, attention-directing processes can explain results of reaction-time studies (Santiesteban, Catmur, Coughlan Hopkins, Bird, & Heyes, 2013). In turn, she rejected evidence from looking-time studies (Kovács et al., 2010; Onishi & Baillargeon, 2005) that showed that infants are born with innate modules for interpreting another's perspective, offering low-level perceptual alternatives for the authors' cognitive interpretations of the results (Heyes, 2014a).

The problem with the lean view is that it does not offer a developmental account of cognitive development. If infants do not have mental representations while they execute a goal-directed action, when do they start having them? As adults, we ascribe mental representations and intentions to others all the time. More importantly, we ascribe mental representation to *ourselves*, and we should be the ones most entitled to judge whether they truly exist.[1] In the following sections, I am going to introduce current evidence on infants' comprehension of communicative cues as referential and intentional.

1.1.4 The Pointing Gesture

Before language, caregivers and infants are able to establish common ground through eye contact, proximity to the other or the object, body posture and gestures. One early gesture in particular has been an important marker for the understanding of reference and other's intentions: the index-finger point. This is because it not only directs the communicative partner's attention, but it also creates a reference towards the topic of the conversation, and therefore stands in for the object of interest.

---

[1] Even though not in the scope of the current thesis, one can ask the question whether what mental representations actually *are* and whether our feeling that we have them proves their existence. Chomsky (1983) tried to answer this question for cognitive psychology, and Dennett (1996) gave a philosophical account of this "intentional stance", building on the notion of *intentional in-existence* introduced by Brentano (1973/1924).

One of the earliest theoretical accounts of infant pointing goes back to Bates et al. (1975), who examined preverbal communication in the first and second year of life. They were the first to make the conceptual difference between *imperative* pointing, which infant use to get others to give them something they cannot reach, and *declarative* pointing, which they describe as the infant's attempt to share attention to an object. Human infants start pointing declaratively and comprehend declarative pointing around the age of 11 – 12 months (Behne, Liszkowski, Carpenter, & Tomasello, 2012; Butterworth, 2003; Schaffer, 1984; Liszkowski et al., 2004; Liszkowski, 2006; Liszkowski, Schäfer, Carpenter, & Tomasello, 2009; E. H. Leung & Rheingold, 1981; Tomasello et al., 2007). Their pointing behavior tends to be both spontaneous and ubiquitous, meaning it does not take much to create a situation that provokes infants to point, and to respond to adult points (Liszkowski, 2011). Performance in pointing comprehension tasks does not correlate to them having received explicit training in pointing (Matthews et al., 2012) and infants from many cultures point (Liszkowski, Brown, Callaghan, Takada, & de Vos, 2012) which implies that it is not primarily the input that drives the development.

Is pointing unique to humans? While some great apes in captivity may learn, with great effort, to use pointing gestures to interact with humans (Call & Tomasello, 1994; Leavens, 2004; Zimmermann, Zemke, Call, & Gómez, 2009) and even conspecifics (R. Moore, Call, & Tomasello, 2015; Pelé, Dufour, Thierry, & Call, 2009; Tomasello, Call, & Hare, 1998), they perform poorly at comprehending human pointing gestures (Hare & Tomasello, 2004; E. Herrmann & Tomasello, 2006; Tomasello, Call, & Gluckman, 1997). Views differ on whether great apes point for each other in the wild. Some theorists (Leavens, 2012; Lœvenbruck, Dohen, & Vilain, 2009) claim that they do, even if this behavior has only been observed on rare occasions (in chimpanzees: de Waal, 1982 and Pika & Mitani, 2006, 2010 and in bonobos: Veà & Sabater-Pi, 1998), while several authors (Tomasello, 1998; Gontier, 2013) argue that pointing behavior consistent with communicative intent has never been observed in apes in their natural habitat.

One explanation for this discrepancy in human and great ape performance may be that humans evolved in cooperative social groups (Tomasello, 2014), while conspecific relationships in apes tend to be naturally competitive (Hamann, Warneken, Greenberg, & Tomasello, 2011; Hare & Tomasello, 2004). Indeed, chimpanzees perform better at pointing tasks that are embedded in an ecologically valid context; e.g. hindering a competitor from getting resources (E. Herrmann & Tomasello, 2006). However, R. Moore (2013) suggests that apes also perform poorly in pointing tasks that cater to their competitive disposition (Tempelmann, Kaminski, & Liebal, 2013) and in tasks where the referent of a gesture has to be deduced from two choices (Kirchhofer, Zimmermann, Kaminski, & Tomasello, 2012). R. Moore concludes that what they fail to accomplish is not only comprehending cooperative motives, but, on a more basic level, comprehending referential gestural communication. The difference, then, between human point comprehension and ape point comprehension, is the referential intention that is evoked by the human pointer and picked up by his human social partner, but not by a non-human hominid partner. R. Moore (2013) explains that the point itself is a weak, ambiguous signal (just as language can be), so the recipient has to form an idea what is in the mind of the pointer in order to decipher his intent, and consequently, the referent of the point. Apes fail to do what infants apparently readily do before they even walk (Behne et al., 2012).

1.1.5 Current Findings from the Study of Referential Understanding

Trying to shed light on the lively discussion between lean and rich interpretations of the 1990s (Bretherton, 1991; Carpenter et al., 1998; C. Moore, 1999; C. Moore & Corkum, 1994), Woodward and Guajardo (2002) investigated 9– and 12-month-olds' understanding of a pointing gesture as referential in a behavioral study. They argued that if infants interpreted the pointing as object-directed, and therefore referential, they should react to a violation of this reference, i.e. when the agent suddenly pointed at another object. They found that 12-month-olds, but not 9-month-olds, looked longer at the point to a novel object than to the same object in a habituation-dishabituation paradigm. In this study, however, reference to the object was not only established through the point, but also through the proximity of finger and target object and its continuous perceptual availability. Therefore, the results do not warrant the interpretation that infants formed the expectation of the referent, as it was already there.

Moll and Tomasello (2004) first showed that from 12 months of age, infants are able to follow communicative cues outside of their own visual field, contradicting Butterworth and Jarrett's *geometric mechanism.* They tested 12– and 18-month-old infants in an interactive gaze following paradigm. The experimental set-up included a barrier, and an experimenter who was positioned so that she could see behind it. The infant however, could see the experimenter, but not what was behind the barrier. The experimenter turned her head to look behind the barrier at a target object, made a sound of surprise and showed a facial expression of excitement, and then turned back to the infant. In the control condition, the target object was visible to the child, but either out of reach (Moll & Tomasello, 2004, Experiment 1) or in front of the barrier (Moll & Tomasello, 2004, Experiment 2). In both experiments, infants crawled behind the barrier significantly more often when the adult looked behind it, then when she looked at the visible target.

Behne et al. (2012) were the first to present evidence that 12-month-olds do not only follow points, but are able to infer what the referent of the point is. In a tabletop set-up, the infant faced the experimenter who presented a board with two locations which could both be occluded by cloth. The infant was familiarized with the set-up to establish their capability to search under the cloth. For this, the experimenter visibly put a small item under a location and covered it, prompting the infant to look for it. In the test phase, the experimenter hid the item non-visibly and used a pointing gesture to indicate the correct location. Infants searched at the correct location significantly above chance. Behne et al. also found a correlation between pointing comprehension and pointing production in their sample, suggesting that infants who reached for the pointed-at location were aware of the communicative function of pointing. However, the behavioral nature of the paradigm still made it difficult to fully exclude leaner, attentional explanations of the grasping of the cloth and toy.

Investigating the referential expectation elicited by direct gaze using implicit measures, Csibra and Volein (2008) developed a preferential looking paradigm. In line with Csibra and Gergely's proposal on the importance of ostensive cues in establishing reference, 8– and 12-month-old infants watched videos of an agent looking directly into the camera and greeting the infant, then looking down towards one of two occluded locations. A curtain then came down to conceal the agent, and the occluders moved to the sides to reveal a toy either on the referred side (consistent outcome) or on the opposite side (inconsistent outcome). Infants were only included if they followed the gaze on at least two out of four trials and most infants did so. Infants of both age groups looked longer at the empty side in inconsistent trials, even when excluding the first look due to this side also being the cued side. They also switched sides more often on inconsistent trials. Csibra and Volein took these results to mean that infants understood that the agent meant to communicate information about a referent's location to them and therefore expected an object to appear where it had been indicated.

Similarly, a study by Gliga and Csibra (2009) demonstrated that infants appreciated the reference towards an objects using both a pointing gesture and its label. In a design similar to the one used by Csibra and Volein (2008), infants saw an agent greet and then point and label (e.g. "Look, it's a duck!") an occluder on one side. When the occluder moved away, the labeled object could be on the labeled, or on the opposite side. 13-month-old infants looked longer at pointed-at site when it contained the unlabeled, inconsistent object than when it contained the labeled, consistent object. Interestingly, the authors also included a dual-source condition in which the point came from the visible agent, but the sound came from an opposite-gender voice. In this condition, the infants did not show the effect. The authors interpreted the results in line with the *Natural Pedagogy* hypothesis (Csibra & Gergely, 2009) as evidence that infants expect communication to be referential, and that signals in different modalities refer to the same thing only when they originate from the same source.

Both the study by Csibra and Volein (2008) and the one by Gliga and Csibra (2009) come closer to providing direct evidence of referential expectation. However, their paradigms are influenced by the attentional, directional quality of the communicative cue, as evidenced by the tendency of infants to look first at the cued side. This introduced a confound between conditions in that infants had more reason to re-orient their attention to the other side on inconsistent trials than on consistent trials, at a point in time where the referent was already fully visible and therefore did not have to be anticipated.

A recent meta-analysis on communicative cues investigated the connection between pointing behavior and language acquisition (Colonnesi, Stams, Koster, & Noom, 2010). The objective of the meta-analysis was to find out whether pointing only precedes language development or whether it also contributes to it. Twenty-five studies were included in the analysis. In twelve studies that tested concurrent relations between pointing and language, Colonnesi et al. found a combined effect size of r = .52, giving strong evidence that pointing and language are correlated. In 18 longitudinal studies, a combined effect size r of .35 was found, which indicates a medium-to-large effect size. Out of the pointing motives, declarative pointing, but not imperative pointing, was correlated with language. This provides evidence for the intentional, referent-sharing nature of the pointing gesture. Age also played an important role as a moderator. Colonnesi et al. found that the older the infant, the more pointing and language correlated, with the highest correlation between 15 and 20 months of age. The earliest significant correlations between pointing - declarative, not imperative - and language were found around 10 – 11 months of age (Camaioni, Perucchini, Bellagamba, & Colonnesi, 2004). The two modalities for pointing, production and comprehension, were both equally correlated with language outcomes, which aligns with the findings that pointing production and pointing comprehension arise in conjunction with each other (Behne et al., 2012). This correlation could be interpreted as evidence for the hypothesis that children who have an advanced communicative ability show it first in their pointing before they also show advanced linguistic skills later in childhood. However, the authors suggested that another plausible explanation was that infants who point more frequently also may get more, and more specialized linguistic feedback by adults, which may enhance their language development. Interestingly, the authors found that environmental factors (country of origin and socioeconomic background) did not moderate the effect, which may be interpreted as evidence against an interactionist account. However, the selection of studies had a bias to be from a *WEIRD* country (Western, Educated, Industrialized, Rich, and Democratic, Henrich, Heine, & Norenzayan, 2010), which makes generalization harder. A recent longitudinal study (McGillion, Pine, Herbert, & Matthews, 2017) found a combination of pointing onset and maternal education to be the best predictor for receptive, but not expressive, vocabulary at 18 months, in a British sample with low socio-economic status.

In a somewhat removed line of studies, researchers reported evidence of simple point following at earlier and earlier ages. In these spatial cueing paradigms, reaction time to the referred-to side of a pointing gesture (or control stimulus) is measured (Bertenthal, Boyer, & Harding, 2014; Gredebäck, Melinder, & Daum, 2010; Rohlfing, Longo, & Bertenthal, 2012). Infants as young as 4 or 4.5 months of age switch their visual attention to the cued side when the cue is a pointing finger, and less often when the cue is a perceptual control. While of interest for the study of the characteristic properties of the pointing gesture, and the innate mechanisms they may ignite, studies on point following should not be confused with studies on point comprehension. Whereas point comprehension is a communicative feat in which cue and context are integrated to execute a meaningful response, point following is solely concerned with the automatic processes take place to orient towards a target.

In conclusion, a broad spectrum of research on joint engagement, from early philosophical accounts of establishing reference, to concrete aspects of human behavior that seem to be important in some way (like gaze following and pointing), to extensive developmental accounts of these behaviors, has been put forward to explain what is so special about human social interaction. Views differ on the depths of the infants' interpretation of social-communicative cues, but not on their significance for development, and infants' inherent and persistent ways of engaging in them. Behavioral studies as well as looking time studies have investigated referential understanding of communicative cues in infancy, but evidence remains susceptible to lean interpretations. Therefore, new implicit measures are needed to provide a better picture on the nature of referential understanding in infancy.

1.2 The Development of Object Cognition

When talking about referentiality in the context of preverbal infants, the referents usually are objects in the physical world. In real-life scenarios, visible objects often are the first referents of infants' points (Tomasello et al., 2007). However, in later communication, the referents become invisible: Words do not exist in a place in the room. Therefore, in order to exclude low-level explanations of point comprehension, the referent must be inferred by the infant, not seen. Essentially, it must be made invisible, as in the occlusion experiments described above (Moll & Tomasello, 2004; Behne et al., 2012). Otherwise, simple attention-directing accounts, and the salience of the object itself dilute the explanatory power of the experiment.

While pointing to absent referents puts the experimenter in a clever position when explaining the results, it poses another problem. For the infant to be able to comprehend the communicative cue, he or she needs to be able to represent an invisible object in the first place. We cannot make the claim that an infant is able to represent an object to be the referent of a communicative cue, if they are not able to represent an object per se. Here comes in our second strand of mental representation: The study of representations of objects.

The study of object representation can be approached from two paths of research: The more general notion of *object permanence*, or knowing that an object continues to exist even when it is occluded or contained (Piaget, 1954) and that of *object knowledge* (Spelke, 1990, 2000; Spelke & Kinzler, 2007) which mostly concerns itself with how infants learn what an object is, how one object is different from another object and how they behave in the physical world.

The evidence delivered by the accounts discussed here provide a rich picture of an early, sophisticated object understanding in the first year of life. Critical accounts of the early findings will be discussed subsequently.

1.2.1 Object Permanence

Object permanence can be traced back to Piaget (1954), but experienced a renaissance through violation-of-expectation (VOE) paradigms in the 1980's (Baillargeon, Spelke, & Wasserman, 1985) and has been the subject of many studies since. In principle, object permanence has been tested in two ways: In behavioral studies, a graspable object is hidden under a cloth or similar easily removed occluder, and the infant is given the chance to retrieve it. In VOE paradigms, an object is hidden behind an occluder or within a container, and a subsequent process is presented which is inconsistent with the continued existence of the object; e.g. the occluder moves "through" the object or the occluder is removed to reveal an empty space where the object should have been. Various behavioral and psychophysical measures may be applied to record the infant's reaction to the violation of expectation, but looking time is the most common (Oakes, 2010).

Piaget (1954) put the beginning of object permanence around nine months, at which age he observed his infant son lifting a cloth and grasping the toy hidden beneath. According to Piaget's theory, this happens in the sensorimotor stage of infancy, which precedes the ability to form inner representations of the outside world. Rather than having a mental representation of the object, Piaget believed that the infant appreciates something to be there, but does not know what it is. True object permanence in the stricter sense of representing a specific object during occlusion, according to Piaget, does not arise until 18 – 24 months of age, as infants move from the sensorimotor stage to the preoperational stage. Behavioral studies modelled after this first grasping experiment require the coordination of action and knowledge, and therefore reliably find infants reaching for occluded objects at the earliest around 8 – 10 months of age (Miller, Cohen, & Hill, 1970; Willatts, 1984). These findings, though, do not support or oppose a mental representation of the object.

Assuming the motor response might be partially responsible for the failure at a younger age, numerous alternative behavioral tasks were created in an effort to make the task more accessible for infants. Reaching in the dark was one of the tasks where infants did not need to remove a cloth, but occlusion was achieved by turning off the lights (Hood & Willatts, 1986, see also Babinsky, Braddick, & Atkinson, 2012. Earlier studies had found that infants reliably reached for sounding objects in the dark that they could not see (Bower & Wishart, 1972; Wishart, Bower, & Dunkeld, 1978). A longitundial study also showed that reaching in the dark arose simultaneously with reaching for visible objects in light around 2.5 months of age (Clifton, Muir, Ashmead, & Clarkson, 1993). Using a search paradigm with two locations and one toy, infants as young as 5 months of age reached for the previous location of the toy after the light had been turned off, compared to the control location. The authors viewed this data as evidence not only for object permanence, but for mental representation. However, Mandler (1998) argued that reaching towards the same location the object was found at previously does not necessitate a mental representation, but can be explained by sensorimotor conditioning. This especially because most studies did not have a delay between turning off the light and allowing a motor response. To fully infer a representation of the object, Mandler (1998) required that a) infants should not have previous experience with reaching in the direction of the object and b) a delay must be introduced between the perceptual representation (lights on) and the motor response (lights off). Clifton, Perris, and McCall (1999) took the paradigm of Hood and Willatts (1986) and made changes to comply with the suggestions made by Mandler. They argued that if reaching was based on a primitive sensorimotor response, only infants who had been familiarized with the reaching location in the dark would succeed in reaching, whereas if there was a mental representation of the object, the infants who were only familiarized with reaching in the light, or not familiarized at all would also succeed in reaching in the dark. Infants reached for the sounding object in the dark in all conditions, leading the authors to conclude that infants are able to represent invisible objects.

It was during the height of the habituation task era in the 1980's that Baillargeon et al. (1985) created a new, visual experiment to assess implicit object permanence in even earlier infancy, with no need for grasping, the drawbridge paradigm.

In this live-action task, infants were first habituated to a screen flapping forward and backward on a plane, covering a 180° arc. In the test event, a box was placed behind the screen, being completely occluded from view when the screen was upright, but also hindering the screen from fully folding backwards. Two events were now shown to the infants: In the possible event, the screen flapped upwards, only to be stopped by the box, reducing the arc of movement to 120°, before coming down again, uncovering the box. In the impossible event, the screen folded up, covered the box from view and continued to fully fold down on the other side, as if the box did not exist. Crucially, when the screen flipped back, it uncovered the box as if it had been there the entire duration of the test trial.

Five-month-old infants looked longer at the impossible than at the possible test event, and a control experiment demonstrated that it was not the preference for one of the events driving the difference. (Baillargeon et al., 1985) argued that these findings are evidence for early knowledge that objects continue to exist even when they are occluded and further, that objects are solid and cannot pass through other objects. A later study confirmed longer looking times at the impossible event for infants as young as 3.5 months (Baillargeon, 1987). Using not the drawbridge paradigm, but other variations of the violation of object solidity, later looking time studies pushed the age of understanding the occlusion of objects to 2 months of age (Aguiar & Baillargeon, 1999; Hespos & Baillargeon, 2001).

Critics claimed that the looking time results could be explained by low level perceptual mechanisms other than object permanence, e.g. perceptual salience of the 180° arc event and novelty of the test event (e.g. Bogartz, Shinskey, & Speaker, 1997; Haith, 1998; Meltzoff & Moore, 1998). In a recent study comparing different looking measures, Dunn and Bremner (2017) found that 6-month-old infants looked longer both at test trials with a violation of expectation and at test trials with a novel object, thus making it hard to draw conclusions from experiments in which the two are confounded, such as the classic drawbridge paradigm. Another stimuli-driven explanation is the inherent preference for the (impossible) test event because of its wider motion range, which has been found in several replication attempts of the drawbridge study (Rivera, Wakeley, & Langer, 1999; Sirois & Jackson, 2012). Others argued that the outcome of the experiment depended on the fragile balance between familiarity vs. novelty preference, which can easily be manipulated by varying the number of familiarization trials (Hunter & Ames, 1988; Schilling, 2000). These findings emphasize the notion that looking time in general and the drawbridge paradigm in particular may be not be suitable tools to investigate object understanding in infants.

1.2.2 Object Knowledge

Concerning object knowledge in infancy, diverse accounts exist to approach the topic, coming from adult cognitive science (Leslie, Xu, Tremoulet, & Scholl, 1998), infant physical reasoning (Baillargeon et al., 2012) and even philosophy (Xu, 2005). In general, these accounts go beyond the question of whether an infant knows about the continued existence of an object, and tap into the question of what the infant knows about an object, and how such knowledge comes into existence. Paradigms may use principles of object permanence, and manipulate the target object or objects in different ways in order to violate the infant's expectation. In the following, I will briefly touch on three aspects of investigating object knowledge, namely principles of object boundaries, or knowing what constitutes an object (Spelke, 2000), individuation, or knowing that distinct objects exist over space and time (Carey & Xu, 2001; Xu, 2005) and numerosity, or knowing that one doll cannot suddenly turn into two dolls (Wynn, 1992; Wynn & Chiang, 1998; Xu, 2003).

Among the first researchers to note infants' ability to anticipate the path of movement of an object, thus tracking the object as a solid unit, were M. K. Moore, Borton, and Darby (1978, see also Bower, Broughton, & Moore, 1971). They found that as young as 5 months of age, infants look to the opposite of a screen behind which an object has disappeared, anticipating its reappearance. Contrary to Piaget (1954), who attributed this kind of gaze behavior to an accommodation of an *action scheme*, unrelated to object representation, M. K. Moore et al. (1978) argues in favor of not only object permanence, but also object identity. Later, Leslie (1984); Leslie and Keeble (1987) demonstrated in a series of experiments using habituation paradigms and carefully constructed control conditions that seven month old infants are sensitive to the spatiotemporal continuity of objects and the relationship of cause and effect.

Therefore, Leslie and Keeble (1987) argue in favor of early, possibly innate capabilities of the visual system that form the basis of later cognitive functions. Pulling these and other findings together, Spelke (2000, see also Spelke & Kinzler, 2007) developed the core knowledge systems account of cognitive development, in which domain-specific core knowledge is part of the genetic makeup, giving infants innate tools to make sense of the world around them and build up their cognitive skills.

1.2.2.0.1 Defining Object Boundaries According to Spelke (2000), infants hold domain-specific knowledge in four systems to represent particular classes of entities from early on in ontogeny: Objects, actions, number and space.[2] Later cognitive skills arise from this core knowledge, and all of human knowledge can be broken down to fit into one of the core knowledge systems. The core system of object representation has received the most attention in research (Spelke & Kinzler, 2007). Objects are defined as persisting bodies with internal unity and stable boundaries (Spelke, 1990). Over the decades, several spatio-temporal principles of what constitutes an object were devised (Spelke, 1990; Spelke & Kinzler, 2007). Most recently, the three principles of cohesion (all points on an object are connected and move as one bounded whole), continuity (objects move on connected, unobstructed paths over space and time), and contact (distinct objects move together only if they touch and do not interact at a distance) have come to describe the phenomena associated with early object representation (Spelke & Kinzler, 2007). Spelke (2000) emphasized that the domain-specificity of object cognition applies only to objects, but not parts of objects or object stuff, such as sand.

Another theory concerning object representation uses visual attention to explain how

---

[2] In addition to the four core knowledge systems mentioned above, Spelke and Kinzler (2007) also introduced a fifth system of social cognition, which is more fuzzily defined. This raises the question of how many more possible "core knowledge" systems we are currently ignoring.

objects are held in our short-term memory while they are within our visual field. Leslie et al. (1998) introduced an indexing system for object-based visual attention that allows tracking objects continuously through space and time, including short periods of occlusion of the object. The index is a mechanism enabling selective attention to an object in the visual field. This index is not a full mental representation of the object, but is described as a sort of "placeholder" for such a representation later on. This can also be described as an "object file", with the possibility of adding featural information about the object. This developmental theory is in line with the theories of adult visual attention brought forward by Kahneman, Treisman, and Gibbs (1992) and Scholl (2001, see also Perner, Huemer, and Leahy's (2015) theory on mental files in infancy and early childhood). Leslie et al. (1998) argued that, over time, infants learn to bind featural information to object indexes once they have been formed, essentially succeeding in feature binding. Therefore, they move from attending to the mostly spatiotemporal information of the index (dorsal "where" path) to simultaneously attending to the featural information of the object (ventral "what" path). This theory is useful in explaining the gradual development from simple object permanence to object individuation and later object identification.

1.2.2.0.2 Object Individuation Being able to individuate objects means an infant is capable of tracking numerically distinct objects through time and space (Xu, 2005). If an infant is able to individuate, he or she should be surprised when object A is hidden, but object B turns up behind the occluder, because the infant was keeping track of object A, which mysteriously disappeared. This capability is described by Xu as "one level up" from object permanence, in that it requires the infant to establish and represent a world with "multiple distinct objects, all of which are permanent" (Xu, 2005, p. 8).

Spelke, Kestenbaum, Simons, and Wein (1995) used a split screen design to test infants' tracking of objects as distinct individuals. Fourteen-month-old infants were presented with a stage with two screens, separated by a space. The infants were habituated to one toy disappearing behind one of the screens, and an identical toy reappearing behind the other screen, without being visible in the space in between. When the infants reached a habituation criterion, the screens were dropped to reveal either one object or two objects. Infants looked longer at the outcome of one object, overriding their initial baseline preference for a display containing two objects instead of one. The authors attributed this prolonged looking to the representation of two distinct objects, which was violated when only one appeared in the end.

In an extension of Spelke et al.'s findings by Xu and Carey (1996), these results were replicated with 10-month-olds. In a second experiment, however, the authors used one continuous screen and two perceptually distinct objects which appeared on either side of the screen, but never at the same time (spatiotemporal condition). When the screen came down, either both objects were present, or only one of them. Surprisingly, infants did not show a violation of expectation towards the single toy event, thus providing evidence that they did not encode the two objects as separate entities. In several control experiments, the authors demonstrated that infants did encode the featural information about the objects: They were able to individuate the objects when both objects were presented simultaneously for a few seconds at the beginning of the experiment (spatiotemporal condition, Experiments 2 and 4). Only around 12 months do infants succeed in the property condition (Experiment 5, Xu & Carey, 1996).

Wilcox (1999) explained the failure in the spatiotemporal condition by differentiating between event-mapping and event-monitoring tasks. They counted the paradigm introduced by Xu and Carey (1996) to the event-mapping tasks: In order to succeed, the infant has to maintain a representation of one state of a display before the reveal and "map" it onto the end state. This task is more demanding than the event-monitoring task, in which a continuous event is displayed. Wilcox used evidence from two previous studies (Wilcox

& Baillargeon, 1998a, 1998b) to make a point that infants as young as 4.5 months can use featural information to individuate objects if memory demands are reduced. They showed that infants looked longer at events where one object went behind an occluder and another one appeared on the other side only when the occluder was not wide enough to hide two objects at the same time.

Xu (2005) argued that individuation also depends on what we know about the objects: If we see a red triangle move behind an occluder, and green square emerge from the other side, our perception of whether there are two distinct objects depends mostly on spatiotemporal information. If timing and speed are right, it might even appear to us that the triangle turns into the square (*Burke effect* Burke, 1952). On the contrary, if the movement is distinct, we easily differentiate between the objects. This is called the *object-based individuation system.* Xu explained that in adults, this effect is the result of an absence of a category. If we know the objects – say, a cup and a plate –, it would be far from our imagination to think one had turned into the other, even if they moved as one. Rather, we would assume that the cup stopped behind the occluder, whereas the plate was already there before emerging on the other side. This is because adults' perception of objects is largely influenced by a *kind-based individuation system.* Infants fail to individuate perceptually different objects when spatiotemporal information is strong, because their kind-based individuation is not fully developed. This explains the failure of 10-month-olds in the property condition, and their success in the spatiotemporal condition Xu and Carey (1996). The authors attributed the success of 4.5-month-olds in the adapted paradigm (Wilcox & Baillargeon, 1998a, 1998b) to the reduced number of alternations, which inhibits the infants from picking up a spatiotemporal contingency between the objects, in addition to being less attentionally demanding. The success of 12-month-olds in tasks with both property and spatiotemporal information is line with her developmental stance on the move to a kind-based individuation system, facilitated by language, starting at around 9 – 12 months of age.

In an attempt to link the study of object permanence and the study of object representation in the adult literature, Carey and Xu (2001) tested infants' ability to track specific, individuated objects. Their objective was the issue of *what* infants represent during typical object permanence tasks: a general idea of "something" or an object and its features. Using evidence from studies of both adults and infants, they point out the similarities between the way objects are processed, and propose the theory that object representation underlies one and the same system in both cases. This means that infants encode objects in a conceptual, not simply a perceptual way, from the beginning. This view has received some support (Feigenson & Carey, 2003; Pylyshyn, 2001) but has not gone unchallenged (Marino & Scholl, 2005; Noles, Scholl, & Mitroff, 2005).

1.2.2.0.3 Numerosity We can keep track not only of one, but of several objects. Converging evidence from infancy research, adult neuroscience and animal models points towards a "fuzzy" system representing discrete (e.g. item count) and continuous (e.g. time passing) quantities (Mou & vanMarle, 2014). This is described as the analog magnitude system (AMS). However, there has long been a debate whether small positive integers are processed in a separate system from large numbers. Early reaction time studies showed that adults were fast and accurate when estimating the number in a small set of up to six items, whereas they became slower and made more mistakes when this number went up (E. L. Kaufman, Lord, Reese, & Volkmann, 1949). In the index/object tracking theory, tracking is limited to three to four objects (Scholl, Pylyshyn, & Feldman, 2001). In higher numbers, the discrimination of two quantities appears to be determined by their ratio, in accordance with Weber's law (e.g. Laming, 2008). Therefore, some researchers argue that the object tracking system (OTS), a mechanism for monitoring visual attention, is used for tracking the number of a small set of items. In the opposing view, the AMS tracks any number of objects, but is more precise in the small-number range.

In contrast, Wynn (1992) argued that neither object tracking nor perceptual discrimination explains infants' early number sensitivity. According to their view, infants possess innate numerical concepts that allow them to operate on small numbers. They provide evidence for their accounts through a number of experiments. In one group, five-month-olds were presented with a puppet theater representing the addition $1 + 1 = 2$, which consisted of a single doll being visible on stage, and, after an occluder moved up, a second doll being placed next to it. When the occluder came down, it revealed two dolls (consistent condition) or one doll (inconsistent condition). In a subtraction condition ($2 - 1 = 1$), infants saw an equivalent movement and the same two reveals, but the reveal of two dolls was now inconsistent, whereas one doll was consistent with the operation. Infants looked longer towards the end reveal when it was inconsistent with the previously presented sequence. The authors make the claim that these results suggest that mathematical abilities are innate.

In a critical response to the strong claims made by Wynn (1992), Haith (1998) explained the findings without using mathematical concepts. According to him, infants simply react to the events that are more unusual, because they do not fit with what they know about objects being constant (Haith, 1998, p. 175). Going even further and criticizing the validity of the VOE paradigm used in the study, L. B. Cohen and Marks (2002) explained the results by pointing out a simple familiarity preference. The infants did not have enough time to examine the display at first when the toy or toys are placed on the stage. As a result, when the occluder came down, the infants tended to look longer at the final display that matched the initial display, so they could continue processing it.

In their review, Mou and vanMarle (2014) investigated whether small number representations in infancy can be attributed to the AMS or the OTS. They come to the conclusion that early in development, the two developing systems interfere with each other. It appears that the OTS may be mostly used for small number sets, whereas the AMS is activated by larger, or continuous items. This argument stems from evidence that the AMS is far less accurate in childhood than in adulthood. In infancy, lack of experience with both small and big number sets makes it hard to reliably estimate the number of items. However, if infants succeeded in small number sets but not in larger number sets, this points towards OTS activation. This explains why there is plenty of empirical evidence for both accounts, depending on the questions asked.

1.2.3 Object Representation and Absence Representation

When discussing object representation, it is also important to consider the conceptual opposite: The representation of nothing, or the representation of empty sets.[3] Imagine a scenario where an object is taken out of a container, and the container remains visibly empty when closed. Now, upon opening it again, the object is suddenly there again. This would be a violation of a "nothingness" expectation, and, at least to adults, a surprising event. While it has been shown that infants as young as five months are able to discriminate the correct solution of a subtraction when the answer is a small positive integer (McCrink & Wynn, 2004; Simon, Hespos, & Rochat, 1995; Wynn, 1992; Xu, 2003), they do not correctly discriminate when the correct answer is an empty set (Wynn & Chiang, 1998). It therefore appears that infants process the representation of objects quite differently from the representation of emptiness. This phenomenon may help us understand what object representation means in infancy. This discrepancy has to be taken into consideration when balancing out violation-of-expectation (VOE) paradigms, because it is important to separate the outcome (full or empty) from the violation of expectation (possible or impossible).

---

[3] In the context of infant development, it makes the most sense to discuss absence representation in terms of empty sets, and not in terms of the mathematical concept of zero. While zero is a number denoting "nothing", its definition is not equal to that of an empty set as used in a behavioral or visual experimental paradigm: A container, or scene, may hold items. Even if it holds zero items, it is still a set and not zero. In the progression of zero-like concepts, representations of empty sets emerge before the mathematical concept of zero (Nieder, 2016).

In their first experiment, Wynn and Chiang (1998) tested 8-month-old infants for their sensitivity for both magical appearances and disappearances using a looking time paradigm. They presented infants with two screens which could be raised to hide one object each. They then manipulated whether one of the objects was actually occluded or removed; and whether the outcome showed only one object or both, representing either expected or unexpected appearance or disappearance of the second object. They found that infants looked longer in the unexpected disappearance condition, but showed no preference in the unexpected appearance condition. To further validate this results, Wynn and Chiang conducted a second experiment where they presented infants with a display containing only one object. In the magical appearance condition, a hand came into the scene to visibly remove the object from behind the screen. In the expected appearance condition, the hand came in to push the object from view behind the occluder. In both cases, the screen was lowered to reveal the object. If the infants had formed an understanding of the space behind the screen to be empty, they should have looked longer in the magical appearance condition. However, the authors found no difference between the two conditions in looking time over three trials. Translated into mathematical terms, this means that 8-month-olds fail to recognize a violation in the set equivalent of the operation of 1 - 1 = 0, whereas it has been shown that even 5-month-olds are sensitive to violations to the similar equation of 2 - 1 = 1, which has the same subtractive magnitude (Wynn, 1992).

As described above, Mou and vanMarle (2014) assume a competing activation of the AMS and the OTS for numerical representation in infancy. They also state that neither of the systems is able to represent empty sets: The AMS indicates that nothing has been counted yet. Therefore, no output or representation is produced. The OTS indicates no object to be tracked, so again, no representation is produced. This may explain why infants have such a difficulty succeeding in the magical appearance condition, but not in the magical disappearance condition (Wynn & Chiang, 1998). They also argue, that from an evolutionary point of view, it makes sense that the representation of empty sets comes late in development: At any given time, there is an infinite amount of empty sets that we do not represent – we do not notice zero bears preparing breakfast when we walk into the kitchen in the morning.

In an experiment performed by Feigenson and Carey (2005), 10- to 12-month-old infants were either given the choice between 1 cracker and 4, or 0 crackers and 4. The authors found significantly more reaches into the container with more crackers in the 0 vs 4 group, than in the 1 vs 4 group. This result is interesting, but it is questionable whether infants had to represent the empty set at all, because there was only one container with crackers to go to in the condition containing an empty set.

While infancy research on the representation of empty sets is sparse, there is some evidence in preschoolers. Bialystok and Codd (2000) asked 3– to 7-year-old children to distribute a number of cookies into containers, including small positive integers and "none". They were then asked to write on a note how many cookies were in each container, to help them remember later. Three-year-olds had trouble putting no cookies into a container, whereas the majority of 4-year-olds were able to do so. In addition, 3- and 4-year-olds tended to denote the number of cookies in an analogue, as opposed to a symbolic way, e.g. drawing two circles on the note to mark two cookies. In case of "no cookies", they left the post-it completely blank, as a representation of the empty container. By 6 years of age, children had completely made the switch to using symbolic representations, including writing the number zero. Merritt and Brannon (2013) tested whether 4-year-olds already represented empty sets as having a numerical value before developing a concept of zero. They presented children with paired stimuli depicting a rectangle containing a number of dots on a touch screen. Children were trained with feedback to touch the rectangle containing fewer dots first and were given 60 test trials. Most of them were able to correctly indicate the ascending numerical order if the numbers of items were small positive integers

(2,4,8), but not if one of the containers was empty. This results suggests that they did not fully represent the empty set as being smaller than the smallest positive integer 1. These results are in line with an earlier finding by Wellman and Miller (1986), which showed that preschoolers (3.5 – 6.5 years) seem have difficulty incorporating zero into the cardinal sequence of numbers, even when they correctly operate on small positive integers. However, Merritt and Brannon (2013) showed that the children who were only able to order small positive numbers, but not empty sets, nonetheless exhibited a distance effect for the empty set, in that they performed better at putting zero as the smaller number the farther away the positive integer was from zero. This distance effect was similar to the one they exhibited with the numeral 1. According to the authors, this means preschoolers without a concept of a symbolic zero already have a mental representation of the empty set as a (small) numerical value which can be directly compared to other numerical values.

Taken together, the evidence on the development of empty sets, or precursors of zero, indicate that infants and toddlers have trouble representing something as being empty. Only around the beginning of the preschool years from 3 to 4 years of age are children able to "give no cookies" to someone, denote emptiness, albeit in an analogue way, and show a distance effect when placing empty sets in a numerical scale. It appears that it is easier for toddlers and children to recognize empty sets as containing nothing when they are able to directly compare it to a set containing something (see Feigenson & Carey, 2005; Merritt & Brannon, 2013). Whenever they are faced with only one empty set and no immediate reference to what it may contain, however, they fail to represent it as empty. Therefore, context information seems to be crucial to denote an empty set as such, and the sheer presentation of a container, which is something, after all, may not be enough to help establish a reference to emptiness.

The reason for this absence of absence representation may lie in the way both the AMS and the OTS allow infants to make sense of the world: Something that is not there does not get encoded. In relation to the concept of expected appearance/unexpected appearance (or violation of nothing expectation, expectation of "nothingness"), we therefore assume that infants in the first year do not form an expectation of "nothingness".

1.2.4 Critical Views on Early Competence

As discussed above, the tendency to attribute cognitive abilities to younger and younger infants using the same looking-time paradigm has drawn critics (Bogartz et al., 1997; Haith, 1998; Meltzoff & Moore, 1998) and may not be helpful in determining the actual abilities in infancy. One specific problem for the proponents of early competence is the *A-not-B error*, also called perseverative search error. Piaget (1954) noted that, after having repeatedly uncovered an object under one cloth A, his 9.5 month-old son failed to uncover the object from cloth B, even though he had hidden it under the new location in plain view. Instead, he again reached for cloth A. This error has been widely studied over the decades and replicated in different settings and circumstances, as researchers have tried to pry apart the connection between thinking and acting (Diamond, 1990; Marcovitch & Zelazo, 1999; Wellman & Miller, 1986).

In line with the early competence stance, Baillargeon, Graber, DeVos, and Black (1990) argued that infants have a complete object representation when they pass the VOE task, and their failure to complete the reaching task is to be attributed to the dissociation between knowing and acting (Bertenthal, 1996). This view poses a problem as to how this disunity arises in the infant brain and which cognitive process, if any, ultimately leads the infant to move his hand back to A despite the knowledge of B. One proposed solution for this issue is the *response inhibition account* (Diamond, 1990) which suggests that infants need to overcome a strong motor habit to reach to a new location, after having formed a behavior that, up until this point, had always been rewarded.

In contrast, Munakata (1997) proposed an account of a gradual achievement of object representation: Infants who pass the VOE task may only have a weak, incomplete object representation, which would not suffice to reach to location B, whereas infants passing the reaching task have a robust object representation. They used neural networks to simulate such a gradual increase in knowledge. Interestingly, their networks start out with object representation and add action to it, which is in line with "rich" interpretations of infant cognition that attribute mental representation to the infant (see Section "Lean versus Rich Interpretations of Social Cognition").

Mareschal and Johnson (2003) suggested that initially, young infants are unable to coordinate and integrate information from the dorsal ("where") and ventral ("what") routes during occlusion, which explains their performance in looking time paradigms as well as the A-not-B-error. They provided evidence with 4-month-olds, who were able to detect changes in either location or identity, but not both, depending on whether the stimulus material included faces or objects. They concluded that maturation of the frontal lobe probably attributes to the inability to effectively bind information. Taking into account findings from animal models, Káldy and Sigala (2004) suggested that structures in the temporal lobe (ventral pathway) and parietal lobe (dorsal pathway) are central to the development of the working memory (see also Cowan, 2016).

Hespos, Gredebäck, von Hofsten, and Spelke (2009) investigated both infants' and adult's predictive reaches to invisible objects during occlusion events and darkness, and concluded that the act of occlusion is especially difficult to comprehend for both adults and infants, compared with the maintenance during darkness. They found that infants' and adults' performance was similarly affected by increasing the difficulty of the reaching task and therefore argue that the basic mechanisms of object representation are constant over development, and therefore present at birth.

Challenging the notion of knowledge and behavior as separate entities, Smith, Thelen, Titzer, and McLin (1999) proposed the *dynamic systems account* that depicts the visual input, the reach goal and the movement as unified parts of a dynamic system in which each parts influences the other over time. Over the course of the trials, each reach emerges in part out of the memory of the previous reaches. The A-not-B error arises as the reach goal suddenly opposes the recent directional bias of looking and reaching, and both the lacking salience of the visual input and the inexperience of infants in performing a reach lead to their failure in overcoming this bias (see also Smith & Thelen, 2003).

However, more recently, new evidence has led researchers to question the dynamics systems account. Topál, Gergely, Miklósi, Erdőhegyi, and Csibra (2008) were able to manipulate the occurrence of the A-not-B error in 10-month-olds simply by changing the context. Instead of embedding the reach task in a social context, as has previously been done, they reduced the communicative input to a minimum. Surprisingly, this change helped 10-month-olds infants reach for the correct location B more often than in the standard ostensive-communicative condition. The authors argue that this improvement is due to an interpretive bias that normally helps infants learn from demonstrations but in this context leads them to the misinterpretation that the game teaches generalizable information. A later study (Topál, Gergely, Erdőhegyi, Csibra, & Miklósi, 2009) comparing dogs and wolves substantiates their account that a genetic predisposition for social-communicative learning may explain the A-not-B error in human infants.

When examining the evidence presented above, it appears that both the manual search-task (see Section "Object Permanence") and the looking time paradigm have serious limitations when it comes to drawing conclusions concerning object representation in infancy. It seems curious that so much research exists to explain what infants know about physical objects and their number (see Section "Object Knowledge" above), yet there is no consensus when exactly infants represent absent entities.

On the one hand, even studies that achieve a reversal of the A-not-B error do not report success rates of grasping the toy at the correct location as different from chance at 10 months of age (Topál et al., 2008) and a recent longitudinal study showed that the perseverance error even extends to after 12 months in many infants (Rüther & Liszkowski, unpublished data). On the other hand, new and exploratory measures and paradigms have provided evidence in favor of early competence. For instance, Ruffman, Slade, and Redman (2005) found that infants as young as 4 months correctly predicted the location of a hidden toy through anticipatory looks in an interactive study with rich communicative context, but no behavioral outcome measures. Kaufman, Csibra, and Johnson (2005) used EEG to provide an implicit measure of object maintenance consistent with the continued representation of occluded objects in 6-month-olds. Both of this accounts add evidence to the "early" camp without using VOE paradigms or global looking time measures.

Continuing this new approach to use non-traditional measures of infant cognition, I am going to explore object cognition at the end of the first year, using novel measures and paradigms away from the pitfall of global looking time.

1.3 The Use of Pupillometry and EEG Power Band Analysis in Infancy Research

The methodology I have chosen for this investigation is twofold: For the questions concerning object expectation in a communicative – or non-communicative – context, I employ eye tracking, specifically pupillometry. Pupillometry lends itself beautifully to finding out about the attentional processing of a situation and is suitable for VOE paradigms, which enable us to test violations of object expectation. For the question concerning object representation, I employ EEG, and there specifically band power analysis, namely of the gamma frequency band, which has been associated with the maintenance of object representations in short-term memory. This is of interest when asking what kind of representation the communicative cue elicits in the infant's brain.

Employing eye tracking and EEG for infant studies is not without its caveats. Although widely used at this point (Aslin, 2012; Gredebäck, Johnson, & von Hofsten, 2010; Hoehl & Wahl, 2012), implicit physiological and neurophysiological measures differ from traditional measures of infant cognition through behavior in substantial aspects. One, they were not designed with infant participants in mind. Adapting methodology to use with infants can be tricky, and differences in the behavior and biology between adults and infants have to be taken into account. Using methods commonly applied to adults also poses the risk of wanting to compare the two populations. This should only be done very cautiously.

Two, compared to the myriad of evidence on the specific capacities of paradigms like preferential looking or looking while listening (L. B. Cohen & Cashon, 2006; Fernald, Zangl, Portillo, & Marchman, 2008; Golinkoff, Ma, Song, & Hirsh-Pasek, 2013), paradigms employing pupillometry and EEG power band analysis are relatively new and the exact characteristics of each experiment differ enormously. As we know from the study of novelty and familiarity preference in infants (Schilling, 2000), small differences in timing, contrast, salience, or any other factor within the stimulus material can change the way infants process information. Therefore, it is especially important to have fully balanced designs and create control conditions that are adequate for direct comparison both conceptually and perceptually.

In the following sections, I am going to first introduce pupillometry as a measure of sympathetic arousal (see Section "Pupillometry"), than turn to the cognitive substrates of this correlation before presenting current uses of pupillometric paradigms in infancy research. Addressing the EEG measures (see Section "EEG Band Power Analysis"), I am going to discuss the significance of oscillations in different EEG frequency bands, briefly introduce wavelet transforms (WTs) as a tool to access time-specific changes in the frequency band power, before turning to the gamma band in particular and discussing current applications both in the adult and the infant literature.

1.3.1 Pupillometry

The human eye works like any other optical apparatus: There is a photoreceptive material (the retina), a lens to focus the light and a mechanism to control how much light enters the eye (Loewenfeld & Lowenstein, 1993). In cameras, this is called the aperture setting. In humans, it is the iris.

The front layer of the iris consists of pigmented fibrovascular tissue, called the stroma. This membrane has an adjustable circular opening, the pupil. The stroma is connected to a sphincter muscle (*sphincter pupillae*) which is responsible for contracting the tissue, making the pupil smaller, and to a set of dilator muscles (*dilator pupillae*) which pull the iris radially to enlarge the pupil (Beatty & Lucero-Wagoner, 2000, see Figure 1.1).

The visible back layer of the iris is covered in epithelial cells. The high pigmentation of the epithelial layer aids the light-restricting function of the iris by effectively blocking light from entering the eye by any other way than the pupil. This layer is also responsible for

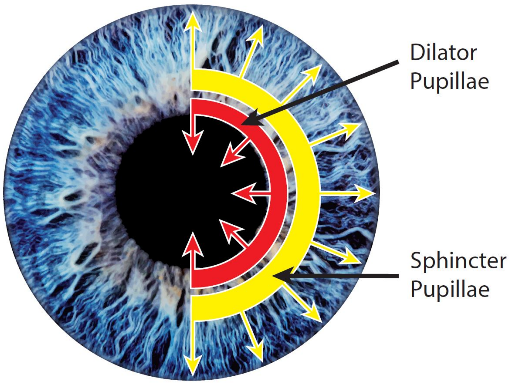


Figure 1.1: Schematic illustration of the muscles in the iris in the author's left eye.

the eye color.

The sphincter pupillae and the dilator pupillae are innervated by the autonomic nervous system (ANS, Beatty & Lucero-Wagoner, 2000). The autonomic nervous system (ANS) controls the function of smooth muscles and glands, including internal organs and bodily functions such as heart rate and respiratory rate. It is regulated by the hypothalamus, an almond-sized structure which is highly interconnected with the brainstem. The ANS can be classified into two branches: the sympathetic nervous system, responsible for fast responses ("fight or flight") and the parasympathetic nervous system, associated with rest and digestion. The dilation of the pupil is primarily a byproduct of sympathetic activation and occurs as the dilator pupillae contracts while the sphincter pupillae relaxes. Constriction occurs as the sphincter pupillae contracts while the dilator pupillae relaxes and derives primarily from parasympathetic activation (McDougal & Gamlin, 2015). Note that dilation can also be a result of the inhibition of parasympathetic activation; and constriction may also be a result of the inhibition of parasympathetic arousal (Einhäuser, 2017). It is the interplay between sympathetic and parasympathetic arousal that results in a net increase or decrease in pupil size.

1.3.1.1 Primary Optical Functions of the Pupil

There are two primary optical functions of the pupil: One, the pupil adjusts the exposure to light to the retina by constricting in bright light and dilating in dark surroundings, called the pupillary light reflex (PLR) (Ellis, 1981). Two, the pupil plays a role in detecting the depth of field by changing the curvature of the lens. This adaptation is also called accommodation response, or near reflex. Both of these functions have to be taken into consideration when examining at the pupil size changes caused by cognitive processes, because they make up the majority of variance in pupil diameter and cause the biggest magnitudes in diameter change (from less than 1 mm to more than 9 mm, see Beatty & Lucero-Wagoner, 2000). Any changes in pupil size, optical or psychological, can be traced through sympathetic and parasympathetic pathways to super-ordinate structures in the brain that control or modulate them.

1.3.1.1.1 Parasympathetic Innervation of the Eye When a person enters a brightly lit room, the first neurons to signal the change in light are the rods and cones on the retina. They activate retinal ganglion cells of the W-type which are sensitive to change in luminance. Cells of the same hemifields in the retinas of both eyes project to the olivary, medial, and posterior prerectal nuclei, which are part of the subcortical visual system and are located at the juncture of the diencephalon and the midbrain (Barlow & Levick, 1969). The olivary pretectal cells then project bilaterally to the parasympathetic, preganglionic, pupilloconstriction neurons of the Edinger-Westphal nucleus (EW). The EW is located in the brainstem posterior to the main motor nucleus (Figure 1.2). It is the primary preganglionic source of parasympathetic innervation of the iris and the ciliary body (Kozicz et al., 2011; McDougal & Gamlin, 2015).

The axons of these preganglionic fibers project to the *ciliary ganglion* via the third cranial nerve, the oculomotor nerve (Kourouyan & Horton, 1997). The ciliary ganglion is about

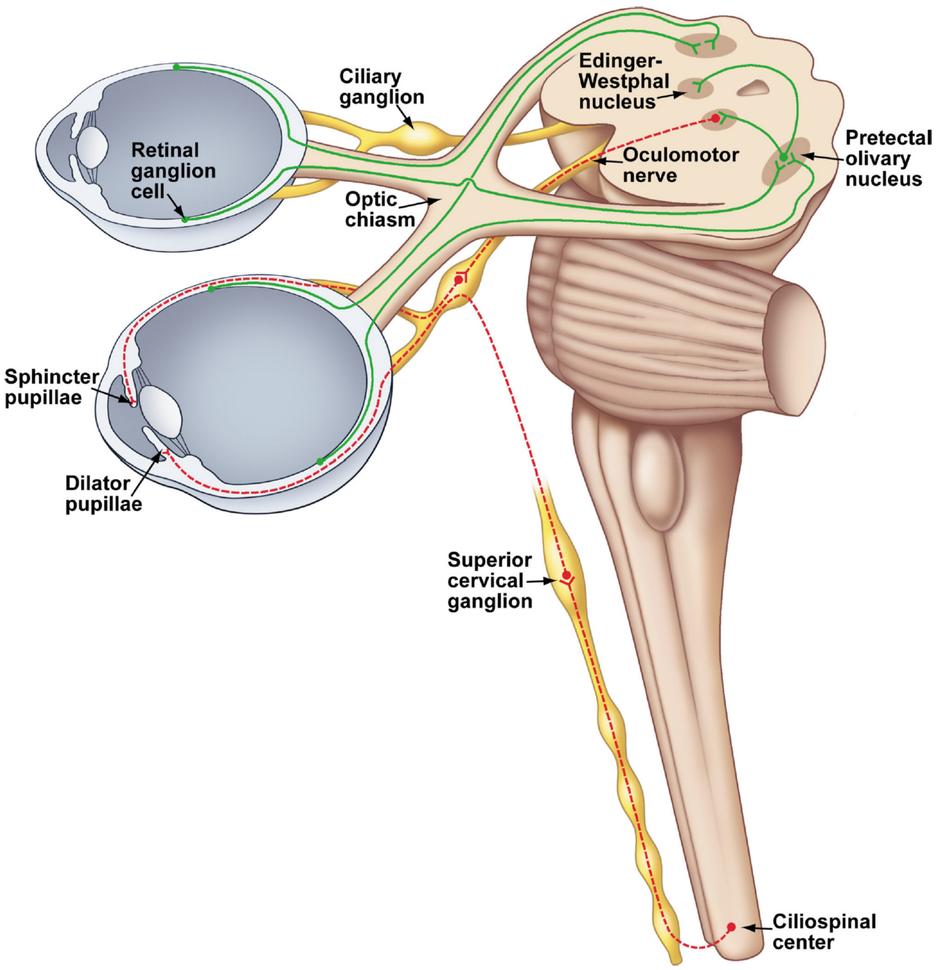


Figure 1.2: Afferent (green) and efferent (red) sympathetic and parasympathetic innervation of the iris. Reprinted from "Autonomic control of the eye" (p. 58), by D.H. McDougal and P.D. Gamlin, 2015, *Comprehensive Physiology*, *5*(1). © 2015 American Physiological Society. Reprinted with permission.

3mm in size and is located in the posterior orbit close to the optic nerve. It processes input from the EW, but also receives input from other sources that may modulate its signal. This means that the ciliary ganglion may not just be a point of passage for neural activation but also function as a point of neural integration (McDougal & Gamlin, 2015). Postganglionic fibers enter the eye via the short ciliary nerves and are then distributed to the sphincter muscles via the short to produce pupillary constriction (Loewenfeld & Lowenstein, 1993).

Interestingly, it is possible to elicit the PLR even in the absence of cones and rods. Recent findings suggest that the PLR is driven by retinal ganglion cells which are intrinsically photosensitive (Gooley et al., 2012; McDougal & Gamlin, 2010). McDougal and Gamlin (2015) reason that photoresponses from both photoreceptor cells and photosensitive retinal ganglion cells are necessary for the full PLR. The PLR can also be influenced by psychological manipulation. Images of the sun, for example, can elicit a PLR even if the image is not brighter than a control image (Naber & Nakayama, 2013). Even reading words associated with brightness can cause the pupil to constrict, though the effect is much smaller than a true PLR (Mathôt, Grainger, & Strijkers, 2017). Dilation naturally occurs in dark settings and is incited by darkness-sensitive cells in the posterior pretectal nucleus, which project to a similar pathway as described above (Clarke & Ikeda, 1985).

There is a secondary parasympathetic pathway from preganglionic neurons in the superior salivatory nucleus projecting through the seventh cranial nerve to postganglionic cells in the nasal ganglion, located halfway between the ear and the malar bone. The nasal ganglion is connected to structures in the eye which are involved in the regulation of blood flow and intraocular pressure (Reiner, Fitzgerald, & Li, 2012).

1.3.1.1.2 Sympathetic Innervation of the Eye The sympathetic innervation of the eye arises from the area called ciliospinal center of Budge (and Waller) in the intermediolateral cell column in the C8-T2 segments of the spinal cord. It connects via the sympathetic chain ganglia and the sympathetic trunk to the superior cervical ganglion, which is situated immediately anterior to the common carotid artery bifurcation on the C1-C3 vertebral level. From here, the axons of the postganglionic neurons project to the orbit, where they enter the eye through the short and long ciliary nerves and through the optic canal (Ruskell, 2003).

While it is generally assumed that pupillary constriction associated with the PLR relies almost entirely on parasympathetic control (Beatty & Lucero-Wagoner, 2000; Clarke &

Ikeda, 1985), there is evidence that light also causes a slower reduction in the tone of the dilator muscle of the iris via the sympathetic pathway, which may enhance a sustained pupillary light reflex (McDougal & Gamlin, 2015).

Because neurons in the Edinger-Westphal nucleus and intermediolateral cell column receive input from superordinate structures of the central nervous system (see Section "Activity in the Locus Coeruleus and the Pupil Response"), sympathethic arousal may lead to pupil dilation (McDougal & Gamlin, 2015).

1.3.1.2 The Cognitive Influence on the Pupil

In addition to the visual functions in adjusting for light and depth perception, the pupil also changes in size in response to cues which are not related to brightness (Loewenfeld & Lowenstein, 1993). From as early as the 1850s, scientists noticed that the pupil changed in size in response to stimuli that were not visual in nature, e.g. tactile, auditory, gustatory, olfactory, or noxious, or even to mental processes that did not include external stimulation at all. Interestingly, this response is always a dilation of the pupil, never a restriction (Beatty & Lucero-Wagoner, 2000). Therefore, some authors call this effect the pupil dilation response (PDR, Wetzel, Buttelmann, Schieler, & Widmann, 2016), others simply call it pupil dilation (Hepach, Vaish, & Tomasello, 2012; Jackson & Sirois, 2009; Sirois & Jackson, 2012), psychosensory pupil response (PPR), reflex dilation, arousal-related dilation, or effort-related dilation (Mathôt, 2018). In the psychological literature, even the superordinate term pupillometry is often used synonymously with the measurement of the psychophysiological widening of the pupil (Hepach & Westermann, 2016; Laeng, Sirois, & Gredebäck, 2012; Sirois & Brisson, 2014).

Hess and Polt (1960) were the first to publish a paper of pupil dilation as a measure of internal arousal. They found that adults' pupils reacted differently to images of different levels of sexual stimulation. A little later they showed that pupils also changed in size when multiplication tasks of varying difficulty had to be solved: the more difficult the task, the bigger the pupil (Hess & Polt, 1964).

Several authors, more notably Beatty and Kahneman (Kahneman & Beatty, 1966, 1967; Kahneman, Beatty, & Pollack, 1967; Kahneman & Wright, 1971) picked up the thread and published a series of experimental studies on a variety of tasks including short-term memory load, pitch discrimination, digit transformation and visual search tasks, establishing pupil dilation as a measure of cognitive load.

Since then, pupil dilation has been used in many different settings, operationalized as a measure of a vast number of varying internal states – including, but not limited to, arousal, intrinsic motivation, perception, memory, decision making, and emotion (for recent reviews of the adult literature, see Einhäuser, 2017; Laeng et al., 2012; Mathôt, 2018; Sirois & Brisson, 2014)

But how and why does this apparent correlation between mental processing and pupil size arise? Part of the answer lies in the innervation of the iris and the interconnection of components of the ANS and the central nervous system. In the next section, I am going to discuss two structures in particular that seem to play a role in the PDR, the locus coeruleus (LC) and the superior colliculus (SC).

The question as to why may be harder to answer. Some authors hypothesize that the pupil's responsiveness to cognitive load is an evolutionary byproduct that is inconsequential enough to have survived many rounds of adaptations and may or may not have once been useful (Beatty & Lucero-Wagoner, 2000). Others suggest that the adaptation of the pupil serves a purpose in communication with conspecifics. In humans, social communication often involves direct eye contact (Farroni et al., 2002, but see Kleinke, 1986 for cultural divergences). The size of a pupil can therefore be an additional cue for the other's arousal during a conversation, including interest (in the content or the social partner) and honesty of the speaker. Another explanation is that the pupil dilates in order to heighten visual sensitivity which is especially beneficial during high-arousal situations, because it facilitates the organism's fight-or-flight response (Mathôt & van der Stigchel, 2015).

1.3.1.3 Activity in the Locus Coeruleus and the Pupil Response

The locus coeruleus (LC) in particular has been named as the brain structure whose activity is mirrored in the oscillation of the pupil. The relationship between activity in the LC and pupil diameter has been well documented in humans and other mammals (Aston-Jones, Rajkowski, Kubiak, & Alexinsky, 1994; Beatty, 1982; Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010; Richer & Beatty, 1987). The neural pathways connecting the LC with the iris are currently not fully understood. Some authors suggest that the absence of evidence of a specific pathway may reflect parallel downstream influences of a common source mechanism instead of a direct connection (Gilzenrat et al., 2010). The paragigantocellularis (PGi) nucleus of the ventral medulla is a likely contender for a common source, as it is both a critical relay for the sympathetic pathways projecting from the hypothalamus to the pupil (Hilton & Smith, 1984; Loewy, Wallach, & McKellar, 1981) and a major afferent projecting to the LC (Aston-Jones, Ennis, Pieribone, Nickell, & Shipley, 1986). Because of the close correlation, some authors even use pupil diameter as a reporter variable for LC activity (Gilzenrat et al., 2010).

Therefore, a look at studies concerning the role of the LC can be helpful in understanding what we are measuring when we talk about pupil dilation. The LC is a structure in the brainstem that is characterized by the presence of neurons synthesizing norepinephrine (also called noradrenaline). It has widely distributed, ascending projections to the neocortex and plays an important role in the autonomic nervous system (Aston-Jones & Cohen, 2005). Traditional views placed the LC as a regulator for arousal, as the processing of salient, arousing stimuli is associated with the release of norepinephrine. The more norepinephrine is produced, the higher is the arousal (Aston-Jones & Bloom, 1981; Berridge & Waterhouse, 2003). The *adaptive gain theory* however challenges this simplistic view (Aston-Jones & Cohen, 2005). Aston-Jones and Cohen (2005) ascribe two functions to the LC based on two modes: the phasic mode and the tonic mode (Aston-Jones & Cohen, 2005; Sara & Bouret, 2012). In the phasic mode, event-locked release of norepinephrine acts as a temporal filter, which facilitates task-related attention. This mode corresponds to the commonly described heightened arousal to stimulating events. The LC stays in the phasic mode as long as utility of the task and performance remain above a critical threshold. When utility declines, the LC transitions to the tonic mode, in which norepinephrine release increases to a stable level, activating a broader number of target neurons, while event-locked release to the previous task declines. Here, arousal towards the stimulus wanes even though norepinephrine is continually released. According to the adaptive gain theory, the two modes serve to regulate the trade-off between exploitation of stable sources of reward and exploration of new, possibly more rewarding opportunities (Aston-Jones & Cohen, 2005; Corbetta, Patel, & Shulman, 2008). In terms of cognitive processes, the LC plays a role in focusing attention towards the task at hand, at least as long as gains of the reward outweigh the costs of re-orienting attention. Gilzenrat et al. (2010) put the adaptive gain theory to test by examining adaptive cognitive control in an auditory oddball paradigm. They found that a large pupil diameter at baseline predicted task disengagement and exploration of different reward opportunities, which indicates an elevation of tonic LC activity. A smaller pupil diameter at baseline however corresponded to task engagement and exploitation of the current source of the reward. These findings were corroborated by several other studies, mostly investigating attentional tasks in adults, correlating the pupil diameter with the P3 component as a psychophysical marker of LC activity (P. R. Murphy, Robertson, Balsters, & O'Connell, 2011), with BOLD activity localized in the LC (P. R. Murphy, O'Connell, O'Sullivan, Robertson, & Balsters, 2014) and with fMRI during a multiple object tracking task (Alnæs et al., 2014). In an effort to provide a direct link, as opposed to correlational data, Joshi, Li, Kalwani, and Gold (2016) implanted recording cylinders in rhesus monkey brains at several sites in and surrounding the LC. Pupil diameter and activity at the LC were measured simultaneously during passive phases, presentation of arousing auditory stimuli, and during electrical microstimulation. They found that both spontaneous and evoked activity at the sites of stimulation is reflected in the pupil diameter. Interestingly, not just the LC, but other interconnected structures (the inferior and superior colliculus and the anterior and posterior cingulate cortex) also predicted change in pupil diameter. Joshi et al. (2016) therefore argue that not just the LC itself, but a network mediated by the LC coordinates norepinephrine-related activity related to attentional processes.

In further establishing the link between pupil oscillation and the LC-norepinephrine system, one recent study revealed two components underlying the pupil dilation response (PDR) signal, suggesting that the phasic and tonic activation of the LC are mirrored in the oscillatory dilation and restriction of the pupil (Wetzel et al., 2016).

How does the adaptive gain theory explain arousal context of the violation of infants' expectations (Jackson & Sirois, 2009; Sebastián-Gallés, 2013) or an index of emotional or social processing (Fitzgerald, 1968; Geangu, Hauf, Bhardwaj, & Bentz, 2011; Gredebäck, Johnson, & von Hofsten, 2010; Hepach et al., 2012)? In the adult literature, diverse phenomena such as explore-exploit trade-off, surprise, salience, decision biases and other effects have been interpreted in terms of activation of LC-norepinephrine system (Joshi et al., 2016). All psychological paradigms, whether they use straight-forward manipulation of perceptual properties of the stimuli (e.g. occlusion) or higher-level social cues, need the infant's attention toward the task in order to be successful. This is why adult's RT studies often exclude participants who make too many errors, suspecting that they did not take the task seriously enough (or failed to understand it), and why infants' eye tracking studies often have an minimal time watched criterion during the critical manipulation phase. Therefore, the processes underlying violation of expectation and processing social cues may well be correlated with attentional control. For instance, a video sequence that conforms to the infant's expectations is of low difficulty, and it is easy for the infant to stay engaged. Arousal is low, and correspondingly, the pupil is small. Contrarily, a VOE sequence is more demanding: there is a mismatch between how the scene should be playing out, and what the infant sees on the screen. Arousal heightens, and the pupil becomes bigger.

1.3.1.4 Role of the Superior Colliculus in the Pupil Response

Because of the strong correlation between LC activity and the pupil on the one hand, and the discrepancy between pupillometry as a "tool" in psychology and the research focusing on the specific functions of the LC on the other hand, other pathways have not been strongly investigated (Larsen & Waters, 2018; Mathôt, 2018). However, recently, another neural substrate has received attention as being an important moderator of the PDR (C.-A. Wang & Munoz, 2015): the superior colliculus (SC).

The SC is mostly known for activity relating to saccadic eye movements and spatial attention (Gandhi & Katnani, 2011; Krauzlis, Lovejoy, & Zénon, 2013), but may also be responsible for coordinating the PDR and therefore may be an important link between cognitive processes and the pupil response. Recent studies have shown that direct stimulation of the SC in monkeys led to pupil dilation (C.-A. Wang, Boehnke, White, & Munoz, 2012), and human's pupils dilate more for faster and for more difficult anti-saccadic eye movements (look in the opposite direction of a stimulus) than for pro-saccadic eye movements (look at the stimulus, C.-A. Wang & Munoz, 2015). C.-A. Wang and Munoz (2015) therefore propose that projections from the SC are a major contribution to the pupil control circuit. According to Einhäuser (2017), the circuit theory involving the SC and the adaptive gain theory involving the LC are not mutually exclusive. He suggests that the SC and the LC may fulfill complementary roles in modulating the PDR.

1.3.1.5 Pupillometry in Infancy Research

Pupillometric measures in infant studies have been utilized for a variety of questions. Both static changes and time-locked responses have been measured. Sometimes, the pupil diameter average over a period of time is taken before and after a manipulation in order to determine the relative increase in pupil size (Gredebäck & Melinder, 2010). Similarly, another approach is correcting the pupil size after the event with the pupil size at baseline to attain a measure of relative change (Hepach et al., 2012). The advantage of this approach is that it corrects for the naturally large variance of baseline pupil size between participants (Hepach & Westermann, 2016). Taking the dynamic aspect of the pupil into account, other authors have used b-splines to map the data over time, enabling them to directly compare curves and accurately determine the exact time two conditions become significantly different from each other (Jackson & Sirois, 2009; Sirois & Jackson, 2012). Similarly, the continuous signal has been decomposed using principal component analysis, revealing two components underlying the PDR (Wetzel et al., 2016). So far, there are no standards as to the way pupil data should be analyzed, and different approaches to data reduction, baseline-correction and interpolation exist. With the recent rise in use of pupillometric measures in infancy research however, the field is starting to develop best-practice approaches and methodological recommendations (Eckstein, Guerra-Carrillo, Miller Singley, & Bunge, 2017; Hepach & Westermann, 2016).

Similar to the adult literature, the application of pupillometry in infancy research covers a wide range of subjects. Pupil dilation is employed as a measure of sympathetic arousal in response to emotional processing (Geangu et al., 2011; Hepach & Westermann, 2013), motivational aspects of behavior (Hepach et al., 2012; Hepach, Vaish, Grossmann, & Tomasello, 2016) and social cognition (Gredebäck, Eriksson, Schmitow, Laeng, & Stenberg, 2012; Hochmann & Papeo, 2014; Jessen, Altvater-Mackensen, & Grossmann, 2016; Fawcett, Arslan, Falck-Ytter, Roeyers, & Gredebäck, 2017). Aside from more general questions of cognitive developmental, pupil dilation lends itself particularly well to the study of language acquisition, partly because stimulus presentation can easily be constricted to the auditory modality, which has no perceptual influence on the pupil. Recent studies on mispronunciation (Tamási, McKean, Gafos, Fritzsche, & Höhle, 2017), information mismatch (Renner & Włodarczak, 2017), language control in bilingual children (Byers-Heinlein, Morin-Lessard, & Lew-Williams, 2017) and sentence comprehension (Lum, Youssef, & Clark, 2017) prove the wide adaptation of pupillometry in infant cognitive research.

Because pupillometry lends itself especially well for VOE paradigms, it is suitable to supplement, or possibly even replace traditional methods of infancy research such as the more coarse global looking-time and preferential-looking measures (Gredebäck & Melinder, 2010; Jackson & Sirois, 2009; Sirois & Jackson, 2012).

1.3.1.6 How to Measure the Pupil

Today, in both infant and adult research, pupillometry is almost exclusively measured in combination with eye tracking using the pupil and corneal reflection system (Holmqvist, 2015). The most common output is horizontal diameter in mm, as the vertical diameter is more susceptible to error due to eyelid closure. Depending on the eye tracker type, high sampling rates (˜500 Hz) may reveal a fine-grained picture of the pupil's oscillation, whereas lower sampling rates (˜60 Hz) are sufficient for questions of static changes in pupil size from a baseline state to an end state (Hepach & Westermann, 2016). Some authors have challenged the use of commercially available eye trackers for pupillometric studies, stating that gaze position may be a source of error in pupil size estimation (Brisson et al., 2013). In particular, it appears that the system used in the current studies, the Tobii x120 eye tracker (Tobii Technology, Stockholm, Sweden), overestimates pupil size in top-left areas of the tracked screen, and underestimates it in bottom-right areas. In consequence, the pupillometry studies presented here (Chapters 2 and 3) take gaze position into account when reporting pupil data.

For the experiments presented here, the eye tracker was set up in a windowless room with controlled artificial ceiling light, which is crucial when recording pupil size (Holmqvist, 2015). The testing booth, and all visible equipment, were black, which is ideal to discriminate looks to the stimulus material (bright screen) from looks away in the pupil as well as the eye gaze. Only equipment that was necessary for visual access – the presentation screen, eye tracking system and scene camera – were visible to the participant, whereas all other equipment was hidden out of view behind a mollitan wall. Both the screen and the eye tracker were fixed to a hydraulic arm which was attached to a table behind the wall. The presentation screen could therefore be adapted in height to each participant.

Importantly, the pupil does not react immediately to changes in visual stimuli. In adults, the PLR sets in with a latency of about 200 ms, which decreases with the brightness of the stimulus (Mathôt, 2018). In infancy research, where bright stimuli could be startling and therefore a low variance in stimulus luminosity is preferred, a latency of the PLR of about 500 ms can be expected and has to be taken into account (Verschoor, Paulus, Spapé, Bíró, & Hommel, 2015). The latency can easily be determined in the data when the time of luminosity change is known, because the pupil restricts fast at first before fully adjusting to the new ambiance light.

1.3.2 EEG Band Power Analysis

Electroencephalography (EEG) is one of the oldest and yet one of the most versatile neurophysiological methods in use today (Seifert, 2008). Electrodes directly applied to the scalp pick up changes in voltage stemming from neurons firing on the other side of the cranium. The method is praised for its high temporal precision, with most systems being able to sample at 500 Hz or higher. Compared to other imaging techniques, EEG does not offer a high spatial resolution. This is because a signal measured on a surface is attributed to a three-dimensional structure beneath, and neurons, with the exception of pyramidal cells, do not necessarily fire orthogonally to the surface. Therefore, EEG is sometimes compared to putting microphones on the walls of a room with a party of a hundred people in hopes of understanding what the conversations are about (Eugenio Parise, personal communication).

Consequently, deciphering the signal is the main goal when working with EEG. One way to do this is to look for patterns of recurring activity that correlate with certain behaviors or psychological functions. The most obvious patterns of the EEG are its wave-like changes in frequency, also called oscillations. Traditionally, oscillatory frequency bands of the adult EEG are defined as delta (0 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 12 Hz), beta (13 – 30 Hz), and gamma (30 – 100 Hz), having been distinguished that way as each frequency band is associated with specific functional characteristics (Herrmann, Grigutsch, & Busch, 2005). Where do oscillations come from? The mean frequencies of the empirically observed bands form a linear progression on a natural logarithmic scale (Penttonen & Buzsáki, 2003), which is evidence that the boundaries are not arbitrary, but mirror the activity of specific *oscillators* associated with different functions. Therefore, the definition of frequency bands is not only based on apparent functional correlations, but may actually be evidence of underlying synchronized networks that “broadcast” in certain frequencies (Buzsáki & Draguhn, 2004), loosely following the Hebbian principle (Hebb, 1949) that neurons "wire together" by "firing together" using the same frequency.

Typically, the EEG shows synchronized rhythms in quiet states or sleep, and desynchronized rhythms in alert brain states. However, this pertains mostly to *global* synchronization. In alert states, local spots of synchronization may actually be a sign of communication between different regions of the brain, and therefore evidence of cognitive processing (Ahmed & Cash, 2013).

While some frequency bands, like the alpha rhythm, may be visible to the human eye, the EEG signal is layered with many different frequencies that only become apparent in a frequency domain analysis and thus have only been in the focus of research since computers have advanced to automatically decompose the EEG signal into component frequency bands in the 1970s (Seifert, 2008).

1.3.2.1 Wavelet Analysis of EEG Oscillations

The most common way to analyze the frequency components of a signal is by performing a Fast Fourier transform (FFT), which displays all frequencies of a signal in sine and cosine waves. However, the downside of the FFT is that while frequencies are precisely localized, the time domain is lost in the process (Samar, Bopardikar, Rao, & Swartz, 1999). This poses a problem in many applications of power change over time because frequency band power can be both phase-locked to a stimulus onset (similar to an ERP, evoked response) or vary in latency (induced response, Tallon-Baudry & Bertrand, 1999), which means the exact timing may not be known from trial to trial.

When the time domain is needed for a proper estimation of the data, wavelet transforms (WTs) may be used on the EEG signal (Samar et al., 1999). In a WT, the original signal time series is convoluted with a scaled and translated version of a mother wavelet function. The convolution leads to a new signal of wavelet coefficients which quantify the similarity between the original signal and the wavelet function at a specific scale and latency. For EEG analysis specifically, Complex Morlet wavelets are most often used as the mother wavelet function (Herrmann et al., 2005). In contrast to the simple Morlet wavelet, Complex Morlet wavelets are complex functions consisting of a multiplication of a harmonic sinusoidal function and a Gaussian envelope function for both their real and imaginary parts (see Figure 1.3).

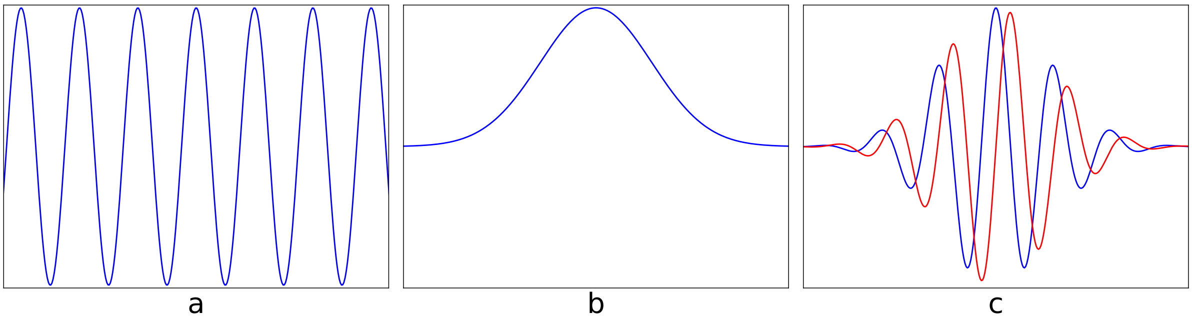


Figure 1.3: Illustration of the construction of a Complex Morlet Wavelet: A sinusoidal function (a) is multiplied with a Gaussian envelope function (b) to result in a wavelet (c). Imaginary parts in red.

Complex Morlet wavelets are especially suited for detecting oscillatory EEG activity because their shape aligns with that of the natural sine waves (Herrmann et al., 2005). Therefore, a wavelet transform works almost like a detector of whether the signal contains the frequency range defined in the wavelet. Because of its function, the Complex Morlet wavelet transform has a different time and frequency resolution at each scale, which leads to a general trade-off between temporal and frequency accuracy. At higher frequencies, the temporal resolution is higher because the number of cycles spread over a shorter interval than for lower frequencies. At the same time, frequency resolution decreases for higher frequencies because of the smaller temporal width. At lower frequencies however, temporal resolution decreases, but frequency resolution increases because of the wide temporal extension. Therefore, depending on the question asked, the parameters of the Morlet function should be adjusted. To obtain a higher resolution in the frequency domain, a relatively larger value for the Morlet Parameter c should be selected. In order to obtain a higher resolution in the time domain, a relatively smaller value for the Morlet Parameter c should be selected.

Because the distortion of accuracy increases exponentially at the ends of the wavelet in the temporal dimension, edge effects have to be taken into account. A conservative method to compute the length of the edge effect is half the wavelet length *WL* (Herrmann et al., 2005; Roach & Mathalon, 2008), which can be calculated from the Morlet Parameter c and the central frequency f with the Equation 1.1:

$$WL = c/f \tag{1.1}$$

For example, if c is 5, and f is 20 Hz, the wavelet length is .250 s and consequently, an edge of .125 s should be discarded from the beginning and the end of the time series.

While a time-locked response is stable after averaging, an induced response is canceled out by averaging over time. Therefore, it is crucial to consider which type of response is likely to be present in the data before averaging. For the evoked response, data are first averaged before the WT is applied. For the induced response, the WT is performed first, then the data are averaged and the evoked response is subtracted from the averaged data, because the evoked response is already included in the averaged signal (Roach & Mathalon, 2008). In studies using novel paradigms where the timeline of expected effects is not completely predicable (Tallon-Baudry & Bertrand, 1999), such as the ones described here, it is also possible to use the total activity without distinguishing between evoked and induced responses.

1.3.2.2 Recording EEG in Infants

There are several factors to consider when recording the EEG of an infant. In contrast to adult participants, infants cannot be instructed to stay still. That means artifacts from movements, including full body jerks, playing with hands or feet, head turns and eye blinks will add noise to the signal in almost all sessions.

Infants are also special in the regard that their EEG signal looks significantly different from that of adults. Contours, ERPs and frequency band activity take on shapes distinct from that of adults (Saby & Marshall, 2012; St. Louis, Frey, & Britton, 2016). In general, effects tend to change in latency across age occurring earlier or later than in the signal of adults, and are less pronounced (de Haan, 2007). Infants' attention spans are shorter than those of adults (Colombo, 2002). At the same time, because of movement artifacts, the noise to signal ratio is so high that a large number of trials are required for averaging to get significant differences between conditions (Stets, Stahl, & Reid, 2012). Therefore, it is crucial to design an experiment in a way that holds the child's attention as long as possible, while keeping the duration of each trial short enough to allow the presentation of at least 30 trials per condition so that 10-15 artifact free trials may be collected (Stets et al., 2012). Attention getters in the form of short animated video clips accompanied by lively music help to refocus the child's attention to the screen in between trials. If several conditions are tested within-subject, it makes sense to balance the order in a way that allows data analysis even if only the first few minutes of the experiment are attended to.

In favor of transparency, attrition rates, as well as reasons for attrition, should therefore always be reported alongside the data. Stets et al. (2012) examined the attrition rates of infant ERP studies in relation to the stimuli characteristics. Analyzing 149 published studies, they found an average of 47.3% attrition (range 0% to 83.8%). They also found that purely auditory stimuli led to a 8.7% decrease in attrition compared to purely visual auditory stimuli, with combined auditory and visual stimuli in between. While previous authors have described attrition rates between 50% and 75% as normal in the field (DeBoer, Scott, & Nelson, 2007), Stets et al. (2012) warn researchers of accepting high attrition rates too easily, and instead suggest that means be taken to create infant-appropriate stimuli to avoid terminations of EEG sessions in the first place, and find ways to use meaningful data from all participants in order to reduce selection biases.

1.3.2.3 Gamma Oscillations as a Marker of Object Representation

In infants and children, the boundaries of corresponding bands generally appear to be lower than in adults (Saby & Marshall, 2012). Alpha, theta and gamma band activity have been mostly in the focus of infancy research. In the following, I am going to focus on gamma band activity only, as it is the frequency band most associated with object cognition.

Gamma oscillations have gained particular interest among the traditionally defined EEG rhythms for their role in cognitive and perceptual processes and because they have divided the scientific community in believers and non-believers (Buzsáki & Wang, 2012, see also Section "Gamma Activity as a Manifestation of Microsaccades"). In infancy, high-frequency oscillations (20 — 60 Hz) map onto the adult gamma rhythm (Saby & Marshall, 2012), with effects related to object cognition in the lower half (20 - 40 Hz, Kaufman, Csibra, & Johnson, 2003; Kaufman et al., 2005).

Gamma rhythms were first associated with perceptual binding and object maintenance in adults in the 1990s (Herrmann & Mecklinger, 2000; Lutzenberger, Pulvermüller, Elbert, & Birbaumer, 1995; Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). One of the first findings was published by Tallon-Baudry et al. (1996) who demonstrated that induced (not phase-locked) gamma-band responses between 200-300 post stimulus distinguish between illusionary coherent stimuli (Kanisza triangle) and non-coherent stimuli. They therefore argued that synchronization in the gamma-band is an indicator for feature-binding in the visual modality. Taking the implications of gamma synchronization even further, Tallon-

Baudry, Bertrand, Peronnet, and Pernier (1998) conceived an experiment that manipulated whether participants who saw novel blob shapes were motivated to keep them in their active working memory or not. Participants who were asked to be prepared to compare the blob to another one after a delay showed increased bursts of activity in the 20-30 Hz range, at the presentation of the shape, at its disappearance, and crucially, after it was no longer visible. Participants who had a task unrelated to the presented blob did not have this burst of activity after occlusion, possibly because they did not have to maintain a representation. The authors took these results to signify a rehearsal of the object representation in short-term memory.

It did not take long for infancy researchers to pick up this thread. Csibra, Davis, Spratling, and Johnson (2000) reported data of 6– and 8-month-old infants presented with the Kanizsa square or a control stimulus. They found increased activity around 40 Hz in the 8-month-olds, who are known to be capable of perceptual binding, but not in the 6-month-olds. Building upon these findings, Kaufman et al. (2003) tested infants in a visual VOE paradigm. Six-month-old infants watched videos of a toy train disappearing inside of a tunnel. The tunnel was then lifted by a human hand. Kaufman et al. (2003) found that gamma activity in right temporal regions increased compared to baseline when an object was unexpectedly missing from the tunnel compared to instances where the train had visibly left the tunnel prior to it being lifted. The authors argued that infants at 6 months, who are sensitive to object permanence in looking-time experiments, have a mental representation of the occluded object that becomes re-activated as the visual reality mismatches their expectation.

The same authors further dissected this object maintenance effect by testing separate sets of stimuli that were either consistent or inconsistent with the continued existence of an object that was being occluded (Kaufman et al., 2005). They used the picture of a ball to be occluded by a moving square. In the occlusion condition, the ball did not change its appearance while being occluded. However, in the disintegration condition, the ball dissolved completely as the occluder moved to the spot where it used to be. They found that right temporal gamma-band oscillatory activity was higher when the ball was occluded than when it disintegrated. Crucially, this difference was significant right after the ball was completely out of view, a phase during which the two conditions were perceptually identical. The authors therefore suggest that the heightened activity corresponds to the representation of the object as opposed to its current state. Hence, they take gamma oscillation to be a marker for object maintenance in the infant brain.

Reid, Csibra, Belsky, and Johnson (2007) presented 8-month-olds with videos of complete and incomplete actions. They found heightened gamma-band activity (defined here as 32 – 48 Hz) starting around 200 ms after the conditions first deviated in the incomplete action videos compared to the complete action videos in left frontal channels. In contrast, they found an increase in the complete condition compared to the incomplete condition in occipital regions. The authors argued that heightened gamma activity in frontal regions in the incomplete condition mirrored greater attention to the events as they diverge from the preconceived expectations. They also suggested that what is happening in the infants' brains may correspond to "forward mapping" as they continue to process an action that is no longer happening. They explained the occipital effects as results of the visual processing of the scene, with both conditions having higher activity during the action than during baseline. The reason the complete action had a heightened response may be that it had a larger motion trajectory than the incomplete action, which stopped earlier.

Continuing the line of research that linked gamma-band responses to object permanence (Kaufman et al., 2003, 2005), Southgate, Csibra, Kaufman, and Johnson (2008) investigated whether graspable objects like toys are encoded differently from featural objects like faces. They presented 6-month-old infants with a VOE paradigm in which a screen moved before the object, briefly occluded it, and opened again to reveal either the same object or another one from the same category. In an attempt to replicate earlier findings, they first looked at changes in gamma activity as the object was being occluded. Interestingly, only the occlusion of a toy led to a burst in the gamma-band, but not the occlusion of a face. Secondly, they compared events in which the same object re-appeared to events in which the object was switched. They found that infants showed a heightened gamma activity in response to the exchange only when the object was a featural object (a face) but not when it was a graspable object (a toy). The author explained this apparent incongruity by discerning between the gamma-band activation resulting from looking at faces in general (see above) and the object specificity lacking in infants in the first year.

Gliga, Volein, and Csibra (2010) looked at gamma-band activity in response to the presentation of familiar and unfamiliar objects. In order to further distinguish between the representation of a spatiotemporal visual object and the representation of a concrete thing, the authors manipulated whether the 1-year-old children had a verbal label for the object. In the first experiment, they found that labeled familiar objects elicited higher gamma-band activity in posterior regions around $500 - 800$ ms after stimuli onset than unlabeled familiar and unfamiliar. The authors took this finding to mean that the gamma response is related to semantic information rather than familiarity per se. In the second experiment, they included a teaching phase in which a novel object was paired with a novel label. They again presented infants with pictures of objects, those whose labels they had just learned and novel, but familiar ones for which they had no label as well as unfamiliar objects. Again, they found that only the objects for which the infants had a label elicited a strong gamma response in occipito-temporal regions, but not the other ones. The authors argue that visual processing of labeled objects was elevated by feedback from other cortical areas that are correlated to extracting semantic information. Therefore, object processing in infants seems to be heavily influenced by language as they are beginning to understand labels.

More recently, Kampis et al. (2015) took the scope of what gamma oscillations may stand for even further by addressing the question of representing another agent's mental state. Widely discussed experiments in adults (Samson et al., 2010) and infants (Kovács et al., 2010) have proposed that humans automatically compute what kind of knowledge another has access to (but see Phillips et al., 2015 for a recent alternative explanation of Kovács et al.'s results). Kampis et al. presented 8-month-old infants with videos of an agent facing an open box with a toy inside. In their first experiment, they manipulated whether the object remained visible or disintegrated before the box turns to occlude the view of the inside to the agent, but not to the infant. They found that both in the full occlusion condition (the object was occluded to both the infant and the agent) and the agent-only condition (the object was still visible to the infant), activity in posterior channels increased in the 25 —35 Hz gamma band. Kampis et al. argued that this activity proves the infant's attribution of a representation to the agent. In their second experiment, the object always disintegrated, but the authors manipulated whether this was visible to the agent, or occluded from her view. The infant always had visual access to the object in this experiment. They found that infants had a continued higher gamma activity in the so-called "false belief" condition, in which the agent has no knowledge that the object has disintegrated, compared to the "true belief" condition, in which both the agent and the infant have seen the object disintegrate. The authors argue that the agent's perspective influences the way the infant represents the scene, therefore ascribing the infant multiple concurrent mental representations.

S. Leung et al. (2016) used gamma oscillations to assess small number representations in 6 to 8-month-olds and presented infants with one or two objects, which were then occluded. They found a significantly greater gamma-activation during the occlusion of two objects compared to one object in the right occipital region. These results may be interpreted as an expansion of the object maintenance effect (Kaufman et al., 2003, 2005).

1.3.2.4 Gamma Activity as a Manifestation of Microsaccades

In recent years, the common interpretations surrounding increased activation in the gamma band – object binding, object recognition, object representation (Gruber, Müller, & Keil, 2002; Herrmann & Mecklinger, 2000; Herrmann, Munk, & Engel, 2004; Tallon-Baudry et al., 1996, 1998) have come under scrutiny due to findings that these activations co-occur with microsaccades following the stimulus onset (Yuval-Greenberg, Tomer, Keren, Nelken, & Deouell, 2008) and may in fact be a muscular instead of a neural response.

Microsaccades are involuntary eye movements shorter than 100 minutes of arc (Yuval-Greenberg et al., 2008). These movements are normal functional processes of the human eye, associated with focus and prevention of perceptual fading (Rolfs, 2009). Yuval-Greenberg and colleagues argue that microsaccades may be suppressed at stimulus onset due to attentional processes and then resume about 200 – 300 ms after the onset, the exact time frame that is often reported for induced gamma activations. Indeed, they found a high correlation between the time of occurrence of microsaccades (as measured by an eye tracker) and the time of heightened activation in the gamma-band within a trial. The very fact that the onset of gamma activation is not time-locked to the stimulus, they claim, is a clue that it is not connected to neural activity and therefore a mere artifact of muscular activity associated with the stimulus presentation.

These findings brought many previous papers reporting gamma activation under fire and led to efforts trying to eliminate ocular artifacts from EEG using independent components analysis, eye tracking, or a combination of both. In removing activity related to microsaccades, some authors have been able to re-establish the link between object representation and gamma activation, while cautioning against using uncorrected data (M. X. Cohen, 2014; Friese et al., 2013; Keren, Yuval-Greenberg, & Deouell, 2010). How does the controversy surrounding gamma-oscillations relate to the data presented here?

Importantly, all of the literature considered in the critique reports adult EEG signal. As noted above, the adult EEG differs from infant EEG in several important dimensions, including amplitude, frequency range, artifact contamination, maximum length of recording etc. The same goes for ocular activity. We know that 6-month-old infants do not show saccade-related spike potentials that produce artificial gamma-band oscillations (Csibra, Tucker, & Johnson, 1998) and that even 12-month-olds produce a very low amplitude (Csibra, Tucker, Volein, & Johnson, 2000). Additionally, we know from eye tracking research that infants' lengths of fixations are shorter than that of adults (Harris, Hainline, Abramov, Lemerise, & Camenzuli, 1988) and exploration patterns gradually develop from salience-driven scanning to adult-like volitional controlled saccades over time (Amso & Johnson, 2008; Bronson, 1994). Therefore, it might as well be that microsaccades present themselves in different ways than in adults. Reliably collecting data on microsaccades would require an eye-tracking system of at least 200 Hz, though 500 Hz is most often used (Holmqvist, 2015). Microsaccades have not been measured in infants, nor would it be sensible to do so with the methods currently available for adult participants (Kampis, Parise, Csibra, & Kovács, 2016). However, independent components analysis relies on long periods of artifact-free EEG which, in infant sessions, are rare exceptions rather than regular occurrences. In terms of correcting for microsaccades, therefore, methods in place for adult data may not be useful for infant data.

Without knowing whether microsaccades are present, and without having a reliable way to exclude them from the data, how can we be certain they are not contaminating the signal? First of all, the problematic correlation to microsaccades was discovered because the time frame of the gamma-burst found in some studies of the adult literature (Tallon-Baudry & Bertrand, 1999; Tallon-Baudry et al., 1998) suspiciously matched the occurrence of microsaccades following the presentation of a visual stimulus. In infant data, however, heightened gamma activity is consistently found later after stimulus onset, around 500 – 800 ms (Kaufman et al., 2003, 2005; Gliga et al., 2010) and does not have the prototypical "candle shape" described in adult literature (Tallon-Baudry et al., 1998). Instead, activity appears like a "burst" that is usually sustained for several hundred milliseconds. The time frame, therefore, does not correspond to activity related to microsaccades.

Secondly, in the experimental designs reported here, care was taken that visual stimuli are similar enough to warrant no difference in the occurrence of microsaccades between different conditions. To be certain of this presumption, the visual stimuli of the EEG studies reported here were presented to infants in the same age range as the EEG participants in an eye tracking setting. No significant differences in exploration patterns were found. Therefore, if differences between conditions occur, they may not be explained by difference in ocular activity.

Lastly, there are ways to actively look for oscillations that should be related to microsaccades, and correlate them with the activity that shows an effect in the experimental conditions. One way of doing so is to examine the central parietal and occipital electrodes, which correlate with microsaccades (Yuval-Greenberg et al., 2008). Microsaccadic activity is not lateralized, so if an effect appears to be clearly restricted to one hemisphere, this adds evidence that is it not based on micro-ocular activity (Gliga et al., 2010). Another way of estimating the influence of ocular activity on the EEG in a broader sense is to subtract activity from the two electrodes closest to the outer corner of the eye. By making use of the dipolar property of the eye, this correction gives an approximation of the activity associated with horizontal eye movements (in humans, microsaccades most often describe horizontal movements, Rolfs, 2009). This activity can then be correlated with activity from channels of interest. If this correlation does not become significant, activation is unlikely to be result of ocular activity (Kampis et al., 2016).

1.3.2.5 Neurophysiological Paradigms of Social Cognition

Object cognition, with its clean-cut perceptual stimuli, can easily be studied using EEG methodology, because stimuli may be presented in picture form in fast succession. Social cognition, however, is more complex and does not easily allow a break-down into pictures with short presentation times. Still, joint attention in infancy has been increasingly moving into the focus of neurophysiological studies in recent years (e.g. Grossmann, Johnson, Farroni, & Csibra, 2007; Grossmann & Johnson, 2010; Grossmann, Parise, & Friederici, 2010; Hoehl, Michel, Reid, Parise, & Striano, 2014; Hoehl, Wahl, & Pauen, 2014; Michel, Wronski, Pauen, Daum, & Hoehl, 2017; Marno, Davelaar, & Csibra, 2014; Mundy, Card, & Fox, 2000; Striano, Reid, & Hoehl, 2006), using several approaches including ERPs, NIRS and EEG power band analyses to bridge the gap between a rich social context and an identifiable stimulus format. However, few studies so far have directly investigated specifically the role of referential-communicative cues, like pointing gestures.

In a pioneer work investigating the connection between infant referential communication and EEG power, Henderson, Yoder, Yale, and McDuffie (2002) performed a longitudinal study looking at declarative and imperative pointing between 14 and 18 months of age. Infants came into the lab at both ages and took part in a video-taped play session which was used to determine the occurrences of the two types of pointing gestures for each child. Infants took then part in a 12-minute recording of background EEG, during which an animated visual attention getter was a presented on a computer screen. The occurrence of infant-initiated pointing acts increased from 14 to 18 months. The authors found a significant negative correlation between higher activity in two frequency bands (4 – 6 and 6 – 9 Hz) in frontal regions at 14 months and occurrence of pointing at 18 months. A decrease in activity in the alpha band means heightened cognitive processing, whereas an increase in activity, or synchronization, in the alpha band is associated with inactivity. Importantly, there were not correlations between decrease in alpha activity to occurrences of imperative, or behavior-regulating pointing. The results suggest that the ability to engage in communicative joint attention can be predicted by decreased synchronization in the frontal regions even months earlier.

Brunetti et al. (2014) asked whether pointing gestures induced communicative intentions, but tested adults, not infants. They tested conditions with declarative pointing (interpersonal goal) against conditions with imperative pointing (instrumental goal). Analyzing gamma band activity from MEG recordings, they found higher activity in regions previously described as a "ToM circuit" for declarative (communicative) pointing than for imperative pointing (Corbetta et al., 2008). This suggests that networks specifically linked to social cognitive skills become active in communicative pointing situations, which could be an indication that a person is aware of the pointer's intention during the point.

Thus far, a similar study has not been undertaken with infants. Therefore, the EEG study presented in Chapter 4 provides an innovative perspective on infant's processing of referential-communicative cues by applying measures of object representation (see Section "Gamma Oscillations as a Marker of Object Representation") to the communicative context. This fills a gap in our current understanding of what infants represent when they see a referential-communicative act.

1.4 Aim and Outline of this Thesis

As detailed above, infants around 12 months of age are able to react appropriately to language and gestures used in referential ways. They are also capable of tracking an object during occlusion and respond appropriately to retrieve it. However, the literature currently available does not offer direct evidence of the representations and expectations infants form during these situations in which object expectation is induced through communication or occlusion. The aim of this thesis was to take what we know about referential expectation and object expectation at the end of the first year and use new physiological methods to provide a clearer picture of infants' mental representations.

1.4.1 Objectives

On the one hand, nativist theories on infant cognitive development propose that infants are born with systems or modules that allow them to understand referential communication. They must be innate because infants so readily pick up the most subtle social cues that enable them to infer meaning. These theories also posit a system to track and recognize objects. Following this view, cognitive mechanisms are already in place at birth to successfully decode social interaction and object occlusion. Therefore, we would expect referential understanding and object cognition to be in place at the same time or earlier than behavioral evidence tells us, since infants' brains should be ready to process information independent from the demands of a coordinated motor response. Nativist accounts tend to lean towards a rich interpretation of infant cognition, because the structures that allow adult mental representations should already be present in the infant.

Following constructivist accounts of infant cognitive development on the other hand, it may be assumed that infants do not thrive if they do not get the meaningful social interaction that healthy caregivers naturally supply. And even though human newborns are sensitive to social cues, the reciprocal nature of a conversation does not arise overnight but develops and refines over large parts of infancy and childhood. At the same time, infants may show appropriate responses to object manipulation, but lack the ability to represent what they see. Following this view, we would expect mental representations to arise later than in the evidence we find in the behavior, because the purpose of social interactions and the true nature of objects may not be immediately evident to the infant. Rather, infants may be motivated by rewards following successful interaction, and this success then gives rise to understanding and mental representations. Constructivist accounts tend to provide a rather lean view of infant cognition, because mental representations, like all other cognitive processes, have to be built up through development.

As so often in psychological research, these two views are not fully mutually exclusive. Innate readiness does not exclude learning through experiential input. Rather than deciding for one or the other, the goal here is to look at a point in development in which the study of referential understanding and object permanence has provided plenty of evidence to posit mental representations – innate or learned – but has remained susceptible to lean accounts due to the nature of the methods used. Here, I used physiological and neurophysiological methods that are robust against lean interpretations because they do not rely on behavioral responses. I used methods specifically devised to investigate infants' expectations and maintenance of objects. I also used carefully balanced violation-of-expectation (VOE) and occlusion paradigms that control for attention and other lower-level perceptual explanations of the results.

1.4.2 Summary and Work Plan

Above, I have summarized the current theoretical approaches and empirical evidence on referential understanding and object cognition. In the following, I will present new evidence I collected in three different empirical studies to investigate these topics with (neuro-)

physiological measures. In the first two studies (Chapters 2 and 3), I investigated expectations following referential cues and object occlusion using violation-of-expectation (VOE) paradigms with pupil dilation as the dependent measure. In the third study (Chapter 4), I focused on the representation of objects during occlusion, and used EEG power band analysis of the gamma band as a signature for object maintenance.

In Study 1 (Chapter 2), I provide new evidence for object expectation following pointing using pupillometry. Using a social VOE paradigm, I demonstrated that at 12 months of age, infants are surprised to find a container empty after an agent had pointed at it. This effect was not replicable with a nonsocial, purely attention-directing cue, and could not be reproduced in a younger age group of 8 months. This finding establishes the utility of pupillometry in the study of referential understanding of social-communicative cues.

The second study using pupillometry (Chapter 3) served to establish a baseline for object expectation in a pupillometric design. Here I complemented the findings from Chapter 2 with a VOE paradigm independent of referential cues, and therefore sensitive to the spontaneous expectation of objects. Object expectation was violated by revealing a space to be empty, which was inconsistent with the previously presented occlusion of an object in the same space. 18-month-old infants succeeded in this task, while 10-month-olds did not show sensitivity for the violation of object expectation in this nonsocial setting.

In the third study (Chapter 4), I investigated both social cognition and object cognition in the same paradigm, using a new and exploratory measure of infant cognition, EEG gamma band analysis. In the first experiment, I established the object maintenance effect in the gamma band and explored online processing of a pointing gesture with 12-month-olds. In the second experiment, I investigated the object maintenance effect in the context of a referential cue, or the absence of such a cue. Findings suggest that 10-month-old infants process occlusion events differently with communicative relevance than without. Study 3 establishes the utility of EEG power band analysis in regard to communicatively induced object representations.

In all studies, the super-ordinate theme is communicative reference and relevance: Reference, in the form of a communicative cue, to an object, and high or low relevance of the object induced by the cue, or the lack thereof. Together, these findings are used to support the notion that readiness to join in referential communication with a social partner is present even in young infants, predates language and facilitates object representation.

2 Pupillometry reveals communication-induced object expectations[1]

RESEARCH HIGHLIGHTS:

- *We used a novel pupillometric paradigm to test for violations of expectations.*
- *Pointing induced object expectations at 12 months of age.*
- *These induced object expectations were specific to pointing and absent for non-social endogenous cues.*
- *Pointing did not induce object expectations at 8 months of age.*

What do infants expect when they see someone point? Following this question, in this chapter I present evidence that a communicative pointing gesture can elicit object expectations in 12-month-olds. I hypothesized that infants would show violation of expectation through dilation of the pupil (Jackson & Sirois, 2009; Sirois & Jackson, 2012) when a communicative pointing gesture suggests the presence of an object, but no object appears. Conversely, there should be no such dilation when there is either no cue to elicit an expectation, or when the expectation is met by presenting an object after the cue. Twelve-month-old infants showed violation of expectation when they saw the agent point, but no object appeared at the location of reference. In a nonsocial control experiment, a directional light-and-sound cue did not elicit the same effect in 12-month-olds. At 8

---

[1] This study is under revision for publication entitled "Pupillometry reveals communication-induced object expectations in 12- but not 8-month-old infants" (Pätzold & Liszkowski, 2018).

months, infants did not differentiate between expected and unexpected outcomes in the experiment using the referential pointing gesture.

This finding extends the current literature on object expectation following a communicative cue in two ways: One, it is the first direct evidence of communicatively induced object expectation in the absence of a visible referent, in contrast to earlier studies in which a toy is often hidden in plain sight (Behne et al., 2012) or even not hidden at all (Woodward & Guajardo, 2002), which confounded the results through additional perceptual and locational cues to the infants. Two, pupillometry is established as a new implicit measure of violation of object expectation which is less vulnerable to low-level, attentional explanations than looking time and behavioral measures used in conceptually similar studies (Csibra & Volein, 2008; Gliga & Csibra, 2009). This study therefore sets the stage for the use of pupillometry in the investigation of object permanence, as reported in Chapter 3.

Acknowledgements: We thank all families who participated in these studies. We thank Jenny Yang and Katharina Orloff for their help in creating the stimuli and running pilot studies. We are grateful to Nicola Ballhausen, Janine Baumann, Jula Brüning-Wessels, Maria Häberlein, Saskia Kiehn, Jana Klose, Tine Maikowski and Lydia Rihm for their assistance with data collection. We thank Susanna Jeschonek-Seidel for valuable feedback in the early stages of the project.

2.1 Abstract

Several interaction-based and looking-time studies suggest that 1-year-old infants understand the referential nature of deictic gestures. However, these studies have not unequivocally established that referential gestures induce object-expectations in infants prior to encountering a referent object and they have remained amenable to simpler attentional highlighting interpretations. The current study tested whether pointing induces object expectations in infants by using a novel pupil dilation violation-of-expectation (VOE) paradigm. In Experiment 1, 12-month-olds watched videos of a protagonist who either pointed towards an occluder in front of her or remained still. At test, the occluder opened to reveal one of two outcomes: an empty surface or a toy. Results showed that infants' pupils were larger for the unexpected outcome of an empty surface following a point compared to the control condition (an empty surface following no point). These differences were not explainable by looking times or directions. In Experiment 2, a nonsocial, attention-directing control cue replaced the pointing gesture. The cue did direct 12-month-olds' attention, but it did not induce any object expectations. In Experiment 3 we tested 8-month-olds in the setting of Experiment 1. In contrast to 12-month-olds, 8-month-olds did not reveal any object-expectations following communication. Findings demonstrate that pointing gestures induce object expectations at 12 months of age, but not at 8 months of age, and that these expectations are specific to social-communicative as opposed to nonsocial cues.

Keywords: Pupillometry, Social cognition, Object expectation, Reference, Occlusion events, Point comprehension

2.2 Introduction

A ‘meeting of minds’ is a pivotal aspect of human communication. Infants begin to follow adults’ gestural attention-directing reference to entities in the environment in the first year of life (E. H. Leung & Rheingold, 1981; Mundy & Newell, 2007; M. Murphy & Messer, 1977). However, the underlying cognitive complexities of infants’ point-following, and its ontogenetic emergence, are still contested. On the one hand, predictive relations to language acquisition and ToM suggest that point-following is causal in the emergence of higher social-cognitive skills and involves a mental understanding of others’ attentional states (Brooks & Meltzoff, 2015). On the other hand, several other species can follow the gaze of others, and sometimes also the canonical human pointing gesture (e.g. Call & Tomasello, 1994; Itakura, 2004; Range & Virányi, 2011; Tomasello & Call, 2008), suggesting that point-following involves simpler cognitive processes which are not directly related to higher social-cognitive skills.

Classic behavioral studies established that around the age of 10 – 12 months, depending on the distance and position of target stimuli, infants will shift their head and gaze towards a lateral target to which an interactant points, more often than to the opposite side (Butterworth & Cochran, 1980; Butterworth & Grover, 1988; C. Moore & Corkum, 1998; Scaife & Bruner, 1975). One interpretation is that infants follow the vectorial direction of a point because the cue orients their visual attention, resulting in the encounter of objects pointed at (Butterworth, 2003; C. Moore & Corkum, 1998). More recent studies have used visual cueing paradigms and measures with a high temporal resolution, including EEG and eye tracking (Bertenthal et al., 2014; Gredebäck, Melinder, & Daum, 2010; Rohlfing et al., 2012).

These studies establish that younger infants around 4 – 8 months of age already orient covertly to a proximal stimulus that has been centrally cued. At 6 months, this cueing effect becomes specific to a hand with an extended index-finger and appears within 100ms, but not 500ms, after cue onset, suggesting that point following is initially an automatic, cue driven response (Bertenthal et al., 2014).

A cognitively richer interpretation is that infants also understand the meaning of the point by inferring the pointer's communicative intentions on several layers (Tomasello et al., 2007). On a referential level, infants would at least need to form expectations about the presence of a referent prior to encountering it (Csibra, 2003). That is, the pointing act should instigate an expectation about a referent which then leads infants to follow the point in search of the appropriate referent. Several interaction-based studies have tested the referential interpretation of point-following at 12 months of age using occlusion paradigms (for gaze-following, see Butler, Caron, & Brooks, 2009; Moll & Tomasello, 2004). Behne et al. (2012) and Liszkowski and Tomasello (2011) have argued that in a hiding game 12-month-olds will not only follow a point to an indicated site at which an object is hidden, but expect to find the object at that site, and so uncover and retrieve it. Further, in both these studies point comprehension to occluded referents correlated with infants' own production of pointing, suggesting a bidirectional understanding of gestural reference. In support, in a looking-time study (Csibra & Volein, 2008) infants watched videos in which an actor turned to look to the right or left side of a surface in front of her, occluded to the infant's view. At test, the actor disappeared and the occluders were removed to reveal an object either on the indicated or the opposite side of the surface. Twelve- and even 8-month-old infants looked longer at the empty side when it had been cued by the head turn than when it had not been cued, suggesting that this ability may emerge already before infants begin to point themselves.

Despite this solid body of research, several questions have remained. First, one problem is that the latter studies have remained amenable to an alternative, intermediate-level interpretation on which point-following does not involve a clear expectation about the referent object prior to finding it but is only a consequence of some form of attentional highlighting. For example, the point to an occluded site, as in the Behne et al. (2012) and Liszkowski and Tomasello (2011) studies, may have served as a directive to go, or just look, to that location, but the uncovering of the object then happened independently of any a priori object expectations. Similarly, in the looking-time study (Csibra & Volein, 2008), infants may have looked longer to the cued but empty location simply because the location had been cued, but not necessarily in expectation of an object (Krehm, Onishi, & Vouloumanos, 2014). Second, it has remained unclear whether non-communicative attention-directing cues would also instigate similar object expectations. If so, this would rather indicate general attentional processes than a specific referential understanding. Third, it is possible that infants' point-following starts out simple and that a more complex understanding still develops across the first year of life. However, direct developmental comparisons probing differences in the cognitive complexities of point-following are still sparse.

One way to distinguish empirically between the different interpretations of point-following would be to test for object expectations while controlling for the allocation of visual attention. In that way one could exclude the leaner interpretation that the amount of allocated attention leads to searching for objects in that location. A promising method in this respect is pupillometry. When infants look at the same scene for the same amount of time, differences in pupil size can still reveal different attentional-cognitive processes underlying the processing of the scene. Pupil dilation has been a measure of cognitive effort for more than 50 years (Bradshaw, 1968; Hess & Polt, 1960, 1964; van der Meer et al., 2010) and recent advances in eye tracking technology make it possible to effortlessly track infants' pupil sizes (Hepach & Westermann, 2016; Hochmann & Papeo, 2014; Sebastián-Gallés, 2013). For example, 8-month-old infants' pupils dilate to violations of object identity in impossible visual scenes (Jackson & Sirois, 2009; Sirois & Jackson, 2012); to violations of action goals at 6 months (Gredebäck & Melinder, 2010); and in real-life settings to failures to assist a person (Hepach et al., 2016, 2012).

In the current study, we designed a novel occlusion paradigm. Across three experiments, infants watched videos in which a door occluded their line of sight, and then opened to reveal either a toy or an empty surface. In Experiment 1, 12-month-olds watched as an actress sat behind the door and either pointed or did not point to the occluded site. If pointing indeed induces expectations about a referent object, the point should instigate a referential expectation about the presence of an object. On pointing trials, the toy outcome should thus be expected, and the empty outcome should result in a violation of that expectation. On non-pointing trials, there should be no specific expectation as to whether there is or is not an object behind the occluder.

Because the two different outcomes naturally differ in their luminance, we did not compare across outcome events but only across manipulation types. Our main comparison concerned the case of the empty outcome: Infants' pupils should be more dilated in the point manipulation, when the empty outcome was a violation of a referential expectation, than in the no point manipulation, when the empty outcome was not a violation, because no expectation had been formed. In the toy outcomes, one possibility was that the appearance of an expected object in the point manipulation would result in a relative decrease in pupil diameter compared to the no point manipulation, because in the former case the state of affairs was already expected and required no cognitive effort, while in the latter case infants were in a state of uncertainty and the new state of affairs still needed to be processed. Alternatively, the toy outcome in the point manipulation could yield no further decrease in pupil diameter compared to the no point manipulation, because infants in the no point manipulation were equally ready to expect a toy or no toy. If infants had no expectations whatsoever, then pupil sizes should not differ between test events of the same outcome, which would be especially noteworthy for our key comparison between the empty outcomes.

In Experiment 2, we addressed the question whether cues other than pointing would also induce object expectations. We used the same basic paradigm and measures as in Experiment 1 but used an endogenous light cue instead of a pointing cue. If simple attention-directing was enough to induce an object expectation, we expected that infants would dilate their pupils in the cued empty outcome condition, like for the pointing gesture.

In Experiment 3, we addressed the ontogenetic question about the origins of communication-induced object expectations. We re-ran Experiment 1 with 8-month-olds, who orient attention following a point, to test whether they also expect an object following others' points.

2.3 Experiment 1

2.3.1 Methods

2.3.1.1 Participants

Seventeen 12-month-old infants (7 males, 10 females) were included in the final sample. Mean age was 12 months, 24 days (range: 12 months, 17 days — 13 months, 1 day). Four additional infants participated (3 males, 1 female) but were excluded from the sample due to failure in reaching minimum looking times during manipulation (1), missing data in one or more conditions during baseline or test (1) or fussiness (2). Each infant provided data for a median of 10 (range 4 – 12) out of 12 trials.

All infants were recruited via birth records and had a middle to high socioeconomic, western cultural background. Infants were included in the final analysis when they provided data for at least one trial of each condition. Trials were included when looking times indicated that the infant had watched at least 50% of the manipulation.

2.3.1.2 Apparatus

We measured eye gaze and pupil dilation with a Tobii x120 eye tracker (Tobii Technology, Stockholm, Sweden) which was attached to a standard 51.50 x 32.00 cm computer screen. The presentation screen and the eye tracker were placed in a testing booth built for this purpose, with black canvases behind and on both sides of the screen. A video camera above the screen allowed the researchers to monitor the infant's behavior during the session. The size of the stimuli presentation was 1280 x 1024 px on a 1920 x 1200 px screen, with the rest of the screen appearing black throughout the experiment. The size of the display was 34.50 x 27.50 cm on the screen, which corresponds to a visual angle of 32.08° horizontally and 25.81° vertically.

2.3.1.3 Stimuli

The videos consisted of live action clips of a female protagonist sitting behind an empty table. During stimulus recording, props served to mark a location in the middle of the table to which the actress pointed. Additional objects were superimposed onto the recording using Adobe After Effects®. In all videos, the same clip of the pointing gesture was used. The edited clips showed the actress sitting behind the table. In front of her, part of the table was occluded by a green door frame that was animated to open and close. When the door opened, it revealed a view on the table surface that was either empty or held a toy. There were 3 toys (car, helicopter and truck) in total, one for each trial. They were constructed from Lego Duplo® blocks and were similar in size. It was apparent in all constellations that the protagonist could always see the target object, even when it was occluded to the infant. In order to hide the agent from view, a white curtain was superimposed onto the video and lowered between the agent and table after the manipulation.

All trials started with a 4 second sequence of animated soap bubbles on an evenly illuminated blue background in order to create a baseline of luminance for each trial (Hepach

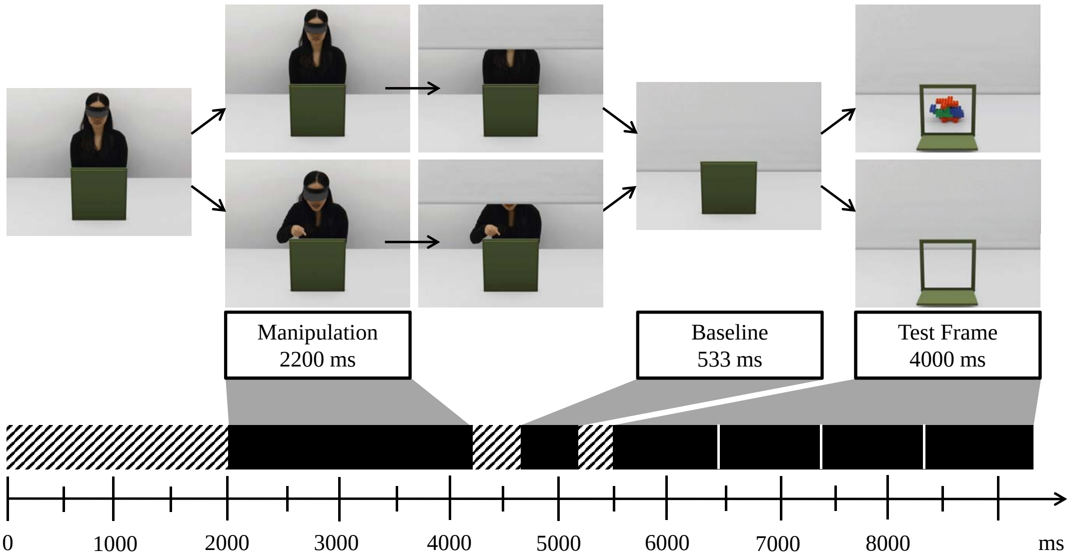


Figure 2.1: Scheme of the experimental design for Experiments 1 and 3.

et al., 2012). The bubbles then cut to the video showing the actress sitting at the table, with the door closed in front of her from the infant's view (2000 ms, see Fig. 2.1).

There were four conditions: two types of manipulation each paired with one of two outcomes. The protagonist either pointed to the area behind the occluder with her right hand, saying "ah" with an excited expression ("point" manipulation) or remained still ("no point" manipulation) for 2200 ms. Following this manipulation, the curtain hid the agent. The door of the occluder then opened and revealed one of two outcomes: a toy ("toy" outcome) or the empty table ("empty" outcome). After 4000 ms of showing a still frame of the outcome, the same animated bubbles were presented again for 4000 ms. Each infant was presented with 3 blocks each containing all 4 conditions. The sequence of presentation was counterbalanced across blocks and participants.

2.3.1.4 Procedure

The experimenter explained the study to the caregiver and obtained their informed consent. The caregiver was seated on a swivel chair in the testing booth 64 cm away from the presentation screen with the infant on his or her lap. A 9-point-calibration ensured that the infant's eyes were properly captured by the eye tracker. The experimental stimuli were then presented on the screen as long as the infant was willing to watch or until 12 trials were completed. The caregivers were instructed to neither talk to nor point for the infant and to keep their eyes closed for the duration of the experiment.

2.3.1.5 Data Processing

Pupil size and gaze location of both eyes were recorded at a 120 Hz rate. Pupil size from the left and right eye was averaged to mean pupil size. If data from one eye were missing, the data from the other eye were used. If data from both eyes were missing, no substitution was made.

Gaze data were processed in relation to a priori defined areas of interest (AOI) that included the face of the protagonist (200 x 330 px) during the manipulation phase and the door area (380 x 355 px). Gaze data at the entire display (1280 x 1024 px) were also calculated in order to allow analysis of looking time. Time looked away from the screen was defined as maximum looking time – looking time to display. For our time window of the main analyses we chose the first 1000 ms of the test sequence because we reasoned that infants would pay equal amount of attention to the outcomes during the first second.[2]

After careful visual inspection, pupil data were shifted to the right by 500ms. Infants' pupils react slower than adults (Verschoor, Spapé, Bíró, & Hommel, 2013; Verschoor et al., 2015) and therefore need more time to adapt to a change in stimulus. Also, by not including the phases of pupillary constriction (PC) following a stimulus onset in the average, a more accurate representation of the tonic pupil dilation is achieved (see Hepach & Westermann, 2016).

---

[2] Infants tended to look away from the screen over time, which is why we chose to report the pupil dilation of the first 1000 ms for our main analyses. However, to provide a fuller picture of the data set, we also report secondary analyses on the entire test period of 4000 ms.

Because pupil data collected by remote eye trackers such as the Tobii X120 may be susceptible to error introduced by gaze direction (Brisson et al., 2013) we checked that the gaze patterns did not differ systematically during the test phase between the two manipulations point and no point ($F_{(1,16)} = .164$, $p = .691$, $n_p^2 = .010$) and that there was no interaction with outcome ($F_{(1,16)} = .829$, $p = .376$, $n_p^2 = .049$). On average, 77.95% of all gaze points (range = 69.98% – 87.62%) fell into the central AOI of the door area during the test event. We report looking time patterns for looking at the video and at AOIs during the test phase. Concerning our main comparisons, infants attended perceptually identical stimuli the same amount of time.

2.3.2 Results

Figure 2.2 displays the mean pupil diameters for the four conditions across a trial. Visual inspection revealed an unexpected difference between point and no point trials in the time window right after the manipulation, before the test phase starts.

During this 533 ms long period — here called still phase – the curtain has come down and the door frame is still closed, rendering videos of all conditions perceptually identical. On no point trials, infants had larger pupils than on point trials (M$_{\text{“point”}}$ = 3.29 mm, SD$_{\text{“point”}}$ = .38 mm, M$_{\text{“no point”}}$ = 3.49 mm, SD$_{\text{“no point”}}$ = .37 mm, $t_{(16)} = 6.54$, $p < .001$). However, analyses of infants' looking pattern revealed that in point trials, infants looked longer at the screen in the time between the curtain starting to close and right before the door starts to open (1000 ms) than in the no point trials, see Fig. 2.4a (M$_{\text{“point”}}$ = 965.07 ms, SD$_{\text{“point”}}$ = 71.12 ms, M$_{\text{“no point”}}$ = 900.15 ms, SD$_{\text{“no point”}}$ = 135.61 ms, $t_{(16)} = 2.10$, $p = .052$). When infants looked away from the screen, they most likely focused on darker areas, as the visible equipment and the lining of the testing booth were black. The apparent dilation of the pupil on no point trials could therefore be explained by the adaptation of the pupil to the dark surrounding area.

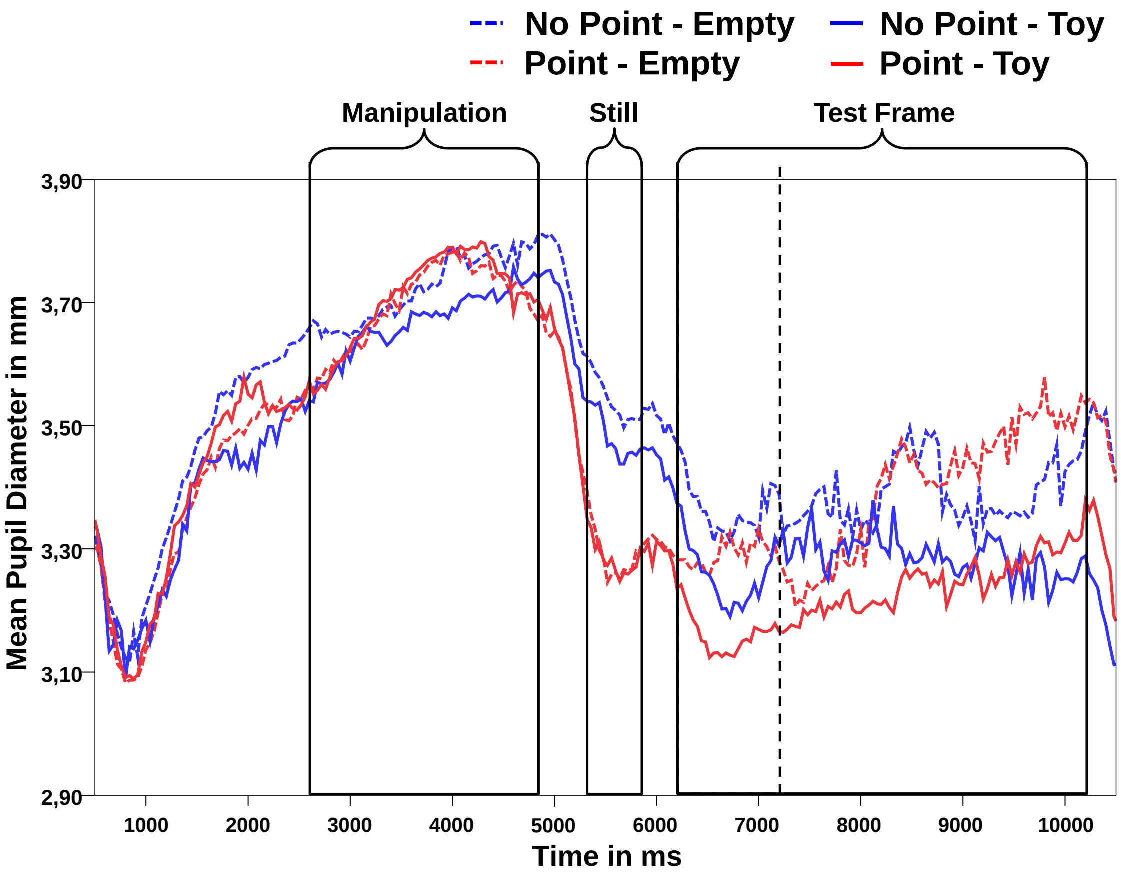


Figure 2.2: Mean pupil diameter of 12-month-olds in Experiment 1. Windows are shifted by 500 ms compared to the timeline in Fig. 2.1.

For our main analyses, in order to adjust for the different levels of pupil dilation before the outcomes, we therefore calculated a relative change score (Hepach et al., 2012), see Equation 2.1:

$$Relative\,change = \Big(pupil\,size(test) - pupil\,size(baseline)\Big)\Big/ pupil\,size(baseline) \quad (2.1)$$

Pupil size at test was defined as the averaged pupil size across the selected time bin (first 1000 ms of test phase). Pupil size at baseline was defined as the averaged pupil size during the still phase between the curtain having come to a stop and the door opening for the reveal (533 ms).

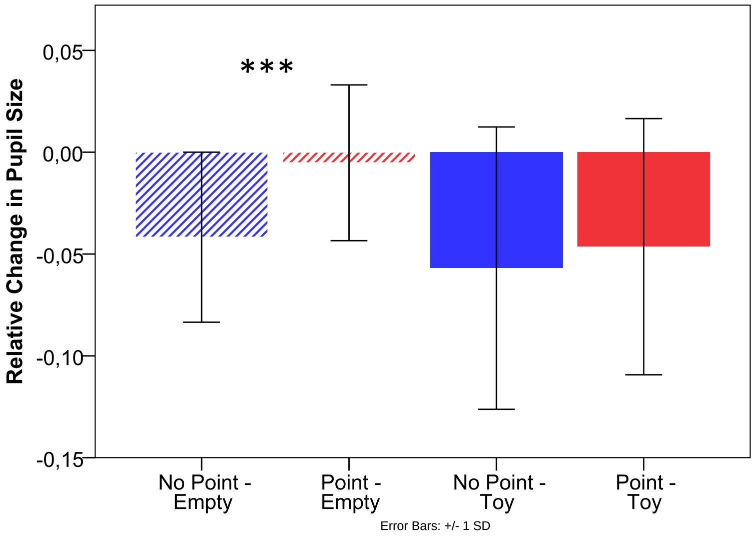


Figure 2.3: Change in pupil size of 12-month-olds in Experiment 1 relative to baseline. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 2.3 shows the relative change scores across conditions, averaged over the duration of the first 1000 ms.

Even though we our main hypothesis focuses on the comparison between the violation of expectation ('point - empty') and the perceptual control ('no point - empty'), we performed repeated measures ANOVAs for the relative change score to get the full picture of the interaction of all four conditions. Table 2.1 shows the results of both the first 1000 ms and the complete 4000 ms time windows. Note that the interaction between gesture (point and no point) and outcome (toy and empty) reaches one-sided significance in the 4000 ms time window, and there are significant main effects for both gesture and outcome.

Direct comparisons following our hypotheses confirmed that in the surprising 'point – empty' condition, the pupil was relatively larger compared to the no 'point – empty' condition ($M_{\text{“point – empty“}} = .005$, $SD_{\text{“point – empty”}} = .06933$, $M_{\text{“no point – empty“}} = -.042$, $SD_{\text{“no point – empty“}} = .043$, $t_{(16)} = 5.18$, $p < .001$). In contrast, pupils in the 'point – toy' condition did not differ in their relative change compared to the 'no point – toy' condition ($M_{\text{“point - toy“}} = .046$, $SD_{\text{“point – toy”}} = .063$, $M_{\text{“no point – toy“}} = -.057$, $SD_{\text{“no point - toy“}} = .069$, $t_{(16)} = .56$, $p =$

.582).

Paired comparisons between all conditions, Bonferroni-corrected with $n = 6$, show that only the condition ‘point – empty’ is significantly different from each ‘point – toy’, ‘no point – toy’ and ‘no point – empty’ ($t_{(16)} > 5.18$, $p < .007$), whereas all other conditions are not different from each other.

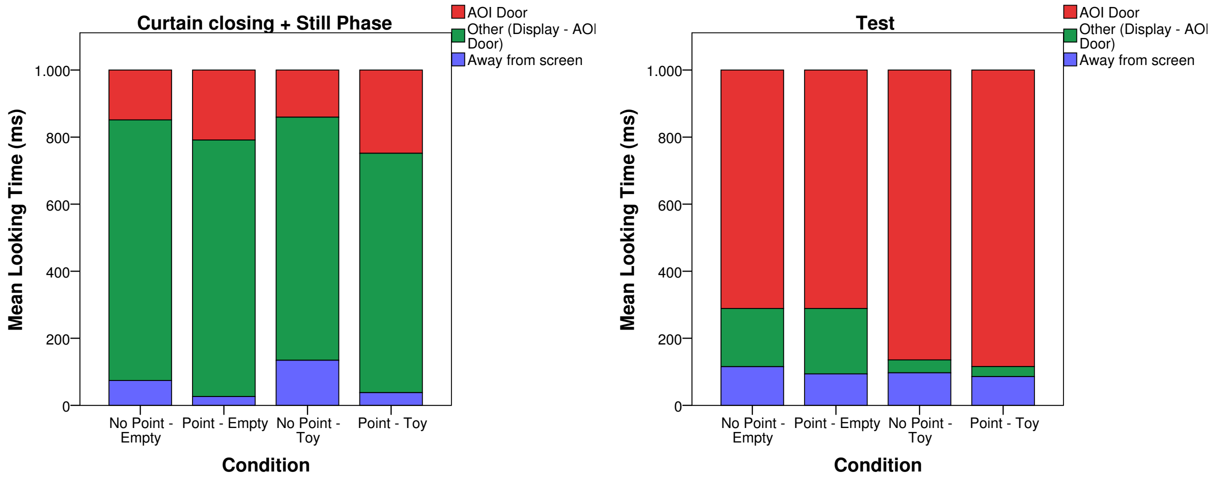
Table 2.1: ANOVA of time segments within test phase for 12-month-olds in Experiment 1.

Time Interval	Within-subject factors	*df*	F	p	η_p^2
0 – 1000 ms	Gesture	16	4.64	.047*	.225
	Outcome	16	7.55	.014*	.321
	Gesture*Outcome	16	2.05	.171	.114
0 – 4000 ms	Gesture	16	17.66	.001**	.525
	Outcome	16	13.62	.002**	.460
	Gesture*Outcome	16	3.11	.097	.163

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

The difference of the ‘point – empty’ condition relative to the other conditions is also apparent when inspecting the mean data across the time course displayed in Fig. 2.2: pupils decreased in all conditions except for the ‘point – empty’ condition (difference from baseline: no point – empty: $t_{(16)}$ = -4.122, p = .001, no point – toy: $t_{(16)}$ = -3.39, p = .004, point – toy: $t_{(16)}$ = -3.04, p = .008 point – empty: $t_{(16)}$ = -.56, p = .584).

To control for the possibility that differences in looking patterns led to differences in pupil sizes, we analyzed looking times to the three different AOIs during the first 1000 ms of test time: 1) Looks at the opened door, 2) looks at video in general and 3) looks away from the screen. Fig. 2.4b shows that infants looked longer to the door area when there was a toy than when it was empty (main effect for outcome: $F_{(1,\ 16)} = 5.89$, $p = .027$, n_p^2 = .269) however, regarding our crucial comparisons, the analysis revealed no significant differences between any conditions in terms of looking time to the video and looking time away from the screen. This means that influences from adaptation to the dark testing booth cannot explain our pupillary finding, and that the differences in pupil size reflect



(a) Baseline phase. *(b)* 1000 ms of test phase.

Figure 2.4: Distribution of looking time of 12-month-olds in Experiment 1 during baseline and test.

differences in the cognitive processing induced by the manipulation.

Since the experiment was designed with a test phase of 4000 ms, we also analyzed average relative change in pupil size for the entire duration of the test phase. Direct comparisons between conditions confirmed that in the surprising ‘point – empty’ condition, the pupil was relatively larger compared to the ‘no point – empty’ condition even for the average across 4000 ms ($\mathrm{M}_{\text{“point – empty“}} = .037$, $\mathrm{SD}_{\text{“point – empty“}} = .048$, $\mathrm{M}_{\text{“no point – empty“}} = -.014$, $\mathrm{SD}_{\text{“no point – empty“}} = .040$, $t_{(16)} = 5.02$, $p < .001$). Pupils in the ‘point – toy’ condition were relatively larger compared to the ‘no point – toy’ condition ($\mathrm{M}_{\text{“point – toy“}} = .013$, $\mathrm{SD}_{\text{“point – toy“}} = .070$ $\mathrm{M}_{\text{“no point – toy“}} = -.040$, $\mathrm{SD}_{\text{“no point – toy“}} = .059$, $t_{(16)} = 2.18$, $p = .045$).

The looking time pattern remained similar: infants looked longer to the door area when there was a toy than when it was empty (main effect for outcome: $F_{(1,\,16)} = 49.18$, $p < .001$, $n_p^2 = .754$), however, regarding our crucial comparisons, the analysis revealed no other significant differences between any conditions of the same outcome in terms of looking time in total or looking away from the screen.

2.3.3 Discussion

Infants' pupils were relatively larger in the 'point – empty' condition, when the actress pointed behind the door but the opening of the door revealed no object, compared to the control condition, when the opening of the door also revealed no object but there had been no pointing cue. This finding is consistent with the hypothesis that infants expect a referent object following a referential pointing gesture already before they encounter the referent object.

Although infants' pupils were generally larger on pointing than non-pointing trials before any outcome was presented, our additional analyses excluded the possibility that our main effect of relatively larger pupils in the 'point – empty' condition was solely driven by this initial difference between pointing and non-pointing trials. Firstly, infants' pupils were not significantly different between pointing and non-pointing trials for the toy outcome, thus rejecting the possibility that pupil size at test was solely driven by the differences after the manipulation. Secondly, by calculating the relative change from baseline to test and comparing it across conditions, we accounted for the differences at baseline. Further, our looking time analyses confirmed that infants focused equally long to the same areas in the stimulus material across conditions, thus excluding that differences in pupil size were spurious and just driven by differences in luminance and pupil orientation during test.

Although not target of our investigation, we note that non-pointing trials lead to larger pupils compared to pointing trials right after the manipulation. One interpretation in line with our main finding is that pointing reduces uncertainty about the next sequence, given a general expectation of referent objects following pointing. However, our looking time analyses revealed that infants looked on non-pointing trials significantly longer away from the screen into the dark surrounding of the testing booth, thus allowing for a leaner, perceptual interpretation. The looking-time difference was fairly small, leaving it somewhat unclear to what extent the different looking pattern contributed to the difference in pupil diameter. The difference during the manipulation between pointing and non-pointing trials should thus be investigated further. In the current context, we prefer to interpret it based on perceptual differences.

Pupil size in the ‘point – empty’ condition did not increase relative to the preceding still phase sequence after the manipulation. However, this lack of relative increase must be interpreted in the perceptual context of the stimulus material and its change in luminance which induced a pupil constriction in all conditions due to the pupillary light reflex. The crucial analysis pertained to the relative change compared to the control manipulation, in which the pupil size decreased to a significantly larger extent than in the ‘point – empty’ condition.

The current findings exclude leaner alternative interpretations in which infant point-following is just an orienting response (Bertenthal et al., 2014), or infants search for referent objects at indicated locations due to differences in attentional allocation (Behne et al., 2012). Instead, by 12 months of age, infants have developed referential expectations which make them follow a pointing gesture. It remains unknown from the current experiment whether this induced object expectation effect is specific to the communicative-referential pointing gesture, or whether other attention-directing, non-social cues would equally induce object expectations. We investigated this question in Experiment 2. A further question is whether the induced object expectations at 12 months of age appear already earlier in development, perhaps even before infants begin to point (Gredebäck & Melinder, 2010). We addressed that second question in Experiment 3.

2.4 Experiment 2

2.4.1 Methods

2.4.1.1 Participants

Twenty-three 12-month-old infants (13 males, 10 females) were included in the final sample. Mean age was 12 months, 12 days (range: 12 months, 3 days — 12 months, 28 day). Twenty-seven additional infants participated (16 males, 11 females) but were excluded from the sample due to failure in reaching minimum looking times during manipulation (12), missing data in one or more conditions during baseline or test (4), technical error (1) or fussiness (10). On average, each infant provided data for a median of 9 (range 4 – 12) out of 12 trials. Drop-out rates were higher in this experiment than in the other two experiments, due to the nonsocial, noncommunicative nature of the experimental material.

2.4.1.2 Stimuli

The videos consisted of a tabletop scene identical to the one in Experiment 1, but without a human agent. Instead of the pointing gesture, in the experimental manipulation a red circle appeared and disappeared on the closed door, accompanied by a bell chiming, directing attention to the door. The timing of this nonsocial audiovisual cue was identical to the pointing gesture and vocalization in Experiment 1. When the door opened, it revealed a view on the table surface that was either empty or held a toy, identical to the ones in Experiment 1. In the control manipulation, no cue appeared. The videos in Experiment 2 were presented in exactly the same way as in Experiment 1, including the presentation of bubbles in the beginning and the end, the order, and the number of trials. As in Experiment 1, we checked that the gaze patterns during the test phase did not differ systematically between the two manipulations cue and no cue ($F_{(1,22)} = .74$, $p = .398$, $n_p^2 = .033$) and that there was no interaction with outcome ($F_{(1,22)} = .25$, $p = .620$, $n_p^2 = .011$). On average, 67.93% of all gaze points (range = 61.14% – 76.53%) fell into the central AOI of the door area during the test event. Apparatus, Procedure and Data Processing were identical to Experiment 1.

2.4.2 Results

Figure 2.5 displays the mean pupil diameters for the four conditions across a trial. As in Experiment 1, visual inspection revealed a difference in pupil size right after the manipulation. The nonsocial cue manipulation lead to smaller pupils right before the test phase compared to the no cue control manipulation ($M_{\text{«cue«}}$ = 3.4323 mm, $SD_{\text{«cue«}}$ = .36533 mm, $M_{\text{«no cue«}}$ = 3.5757 mm, $SD_{\text{«no cue"}}$ = .36498 mm, $t_{(22)} = 2.994$, $p = .007$).

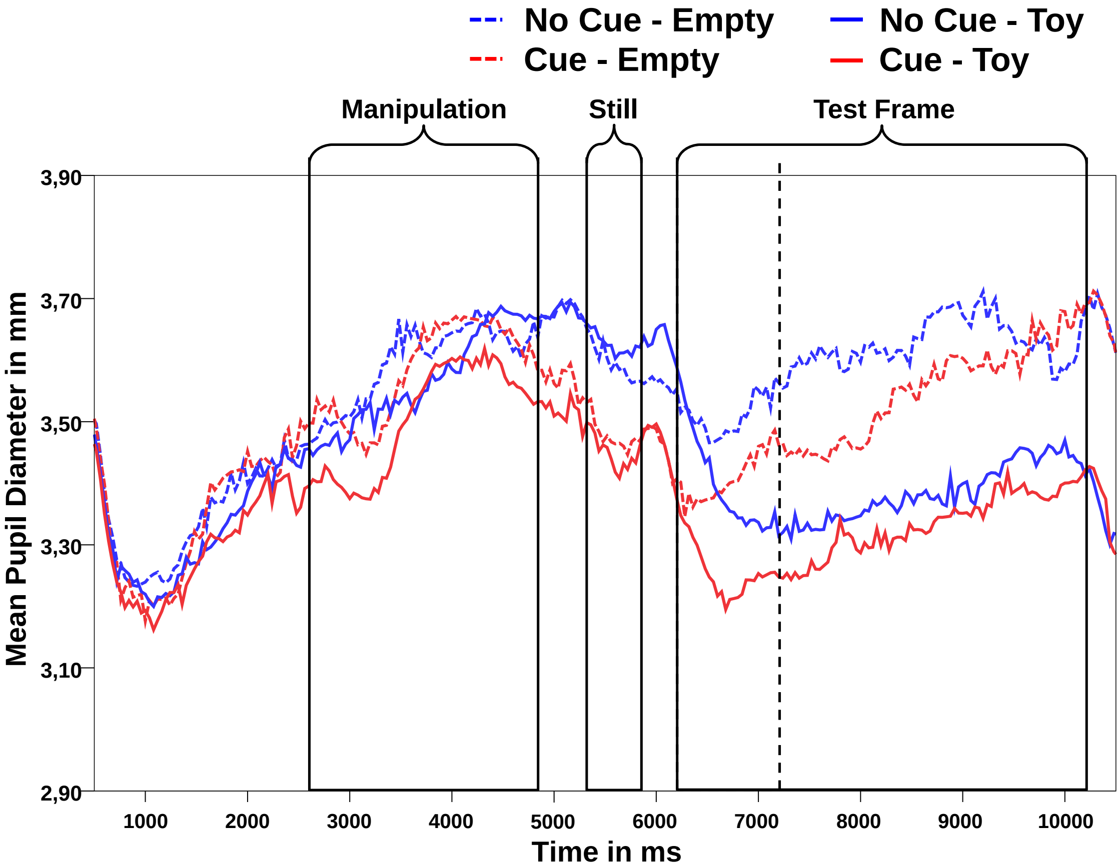


Figure 2.5: Mean pupil diameter of 12-month-olds in Experiment 2. Windows are shifted by 500 ms compared to the timeline in Fig. 2.1.

As in Experiment 1, our analysis of looking time after the manipulation, before the test, revealed that infants spent significantly more time looking at the screen when there was a cue than when there was no cue ($M_{\text{“cue“}}$ = 890.95 ms, $SD_{\text{“cue“}}$ = 215.84 ms, $M_{\text{“no cue“}}$ = 657.90 ms, $SD_{\text{”no cue”}}$ = 182.42 ms, $t_{(22)} = 4.044$, $p = .001$), see Figure 2.7a). This suggests that infants in the no cue manipulation looked away from the screen more, which may have resulted in a larger pupil as their eyes adapted to the darker surrounding of the testing booth.

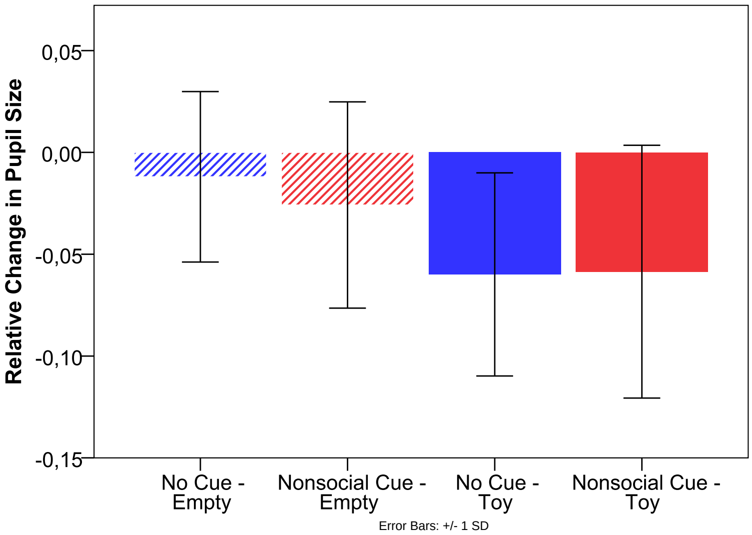


Figure 2.6: Change in pupil size of 12-month-olds in Experiment 2 relative to baseline. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

We applied the same baseline correction as in Experiment 1 (see Fig. 2.6). Again, we performed 2 (cue) x 2 (outcome) repeated measures ANOVAs to get a closer look at how the conditions interacted. Table 2.2 shows the results of both the 1000 ms and the 4000 ms time windows. Unsurprisingly, we find main effects for outcome in both time windows, but no interaction and no main effects for cue.

Direct comparisons following our hypotheses showed no differences in relative changes between the two manipulation conditions for the empty outcome (M“cue - empty“ = -.026, SD“cue - empty“ = .051, $M_{\text{“no cue - empty“}}$ = -.012, $SD_{\text{”no cue - empty”}}$ = .042, $t_{(22)} = 1.43$, p

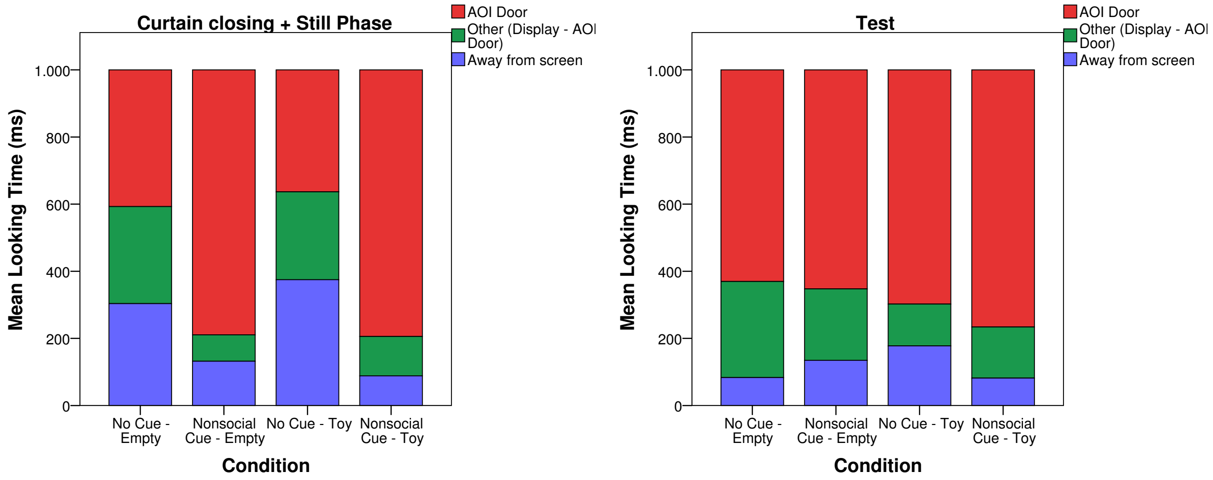
Table 2.2: ANOVA of time segments within test phase for 12-month-olds in Experiment 2.

Time Interval	Within-subject factors	*df*	*F*	*p*	η_p^2
0 – 1000 ms	Cue	22	.366	.551	.016
	Outcome	22	19.923	.000***	.475
	Cue*Outcome	22	1.211	.283	.052
0 – 4000 ms	Cue	22	.515	.481	.023
	Outcome	22	35.577	.000***	.618
	Cue*Outcome	22	1.354	.257	.058

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

= .166) and for the toy outcome ($\text{M}_{\text{“cue - toy“}} = -.059$, $\text{SD}_{\text{“cue - toy“}} = .062$, $\text{M}_{\text{“no cue - toy“}} = -.060$, $\text{SD}_{\text{“no cue - toy“}} = .050$, $t_{(22)} = .09$, $p = .927$).

Paired comparisons between all conditions, Bonferroni-corrected with $n = 6$, show that two other comparisons become significant: ‘no cue - toy’ and ‘cue - toy’ are each significantly smaller than ‘no cue – empty’ (t(22) = 4.18, $p < .001$ and$t_{(22)} = 3.46$, $p = .002$; respectively).



(a) Baseline phase.

(b) 1000 ms of test phase.

Figure 2.7: Distribution of looking time of 12-month-olds in Experiment 2 during baseline and test.

Again, we checked looking times to different AOIs during the test time in order to control for the influence of gaze direction on the pupil (Figure 2.7b). An 2 (cue) x 2 (outcome)

ANOVA of looking patterns revealed no significant differences between any conditions in terms of looking time to or away from the screen. As in Experiment 1, infants looked longer at the door AOI when there was a toy than when it was empty ($F_{(1,\ 22)} = 10.76$, $p = .003$, $n_p^2 = .328$). However, there was no interaction between the two factors, ruling out that the toy effect varied systematically with manipulation, and no significant differences concerning our crucial main comparisons between the two similar outcomes.

As in Experiment 1, we repeated the analyses with data averaged across the entire test phase of 4000 ms, to test for any prolonged effects. Repeated measures t-Test following our hypotheses revealed no differences between conditions in the empty outcomes ($M_{\text{„cue – empty“}} = .005$, $SD_{\text{„cue – empty“}} = .061$, $M_{\text{„no cue – empty“}} = .008$, $SD_{\text{„no cue – empty“}} = .039$, $t_{22)} = .17$, $p = .870$) and the toy outcomes ($M_{\text{„cue – toy“}} = -.039$, $SD_{\text{„cue – toy“}} = .055$, $M_{\text{„no cue – toy“}} = -.057$, $SD_{\text{„no cue – toy“}} = .047$, $t_{22)} = -1.24$, $p = .228$).

Infants looked longer to the door area over the duration of 4000 ms when there was a toy than when it was empty (main effect for outcome: $F_{(1,\ 22)} = 29.69$, $p < .001$, $n_p^2 = .574$), however, regarding our crucial comparisons, the analysis revealed no significant differences between any conditions of the same outcome in terms of looking time in total or looking away from the screen.

Because Experiment 2 differed from Experiment 1 only in one aspect (socialness of cue), we additionally performed a mixed ANOVA with cue (cue or no cue) and outcome (empty or toy) as within-subject factors and socialness (point or nonsocial cue) as a between-subject factor. We found a significant three-way interaction between cue, outcome and socialness ($F_{(1,38)} = 3.39$, $p = .037$, $n_p^2 = .082$[3]), a significant interaction between cue and socialness ($F_{(1,38)} = 3.81$, $p = .029$, $n_p^2 = .091$[3]), a significant main effect for outcome ($F_{(1,38)} = 24.94$, $p < .001$, $n_p^2 = .396$) and no other interactions or main effects.

---

[3] Note: Since the F-distribution is already one-tailed, p-values cannot be halved to achieve one-tailedness. Rather, if the hypothesis concerning groups is directional, contrasts should be tested.

2.4.3 Discussion

Experiment 2 showed that while effective in directing attention to an area, the nonsocial cue did not induce an expectation about a referent object. Infants' reaction to the outcome scene did not vary as a function of the previously seen manipulation. Therefore, we conclude that it is not the amount of allocated attention following a cue that causes infants to expect an object, but rather the referential aspect of the cue: Infants at 12 months know that a point is about something, so when nothing appears, they are surprised. However, simply highlighting an area does not excite this same expectation. In order to better understand the development of communication-induced object expectations, we tested a younger age group in Experiment 3.

2.5 Experiment 3

Experiment 3 was the same as Experiment 1, except that we tested 8-month-old infants instead of 12-month-olds. If results for the 8-month-olds were to mirror those of the 12-month-olds, it would provide strong evidence that an understanding of the referential nature of pointing gestures appears already before infants can point themselves; and before they reliably show point comprehension in behavioral settings. Such evidence would reject constructivist accounts which assume a protracted development of referential understanding (Carpendale & Carpendale, 2010) and instead support early cognition views which rather invoke maturational processes (Csibra, 2003).

2.5.1 Methods

2.5.1.1 Participants

Seventeen 8-month-old infants (8 males, 9 females) were included in the final sample. Mean age was 8 months, 18 days (range: 8 months, 4 days — 9 months, 5 days). Nine additional infants participated (6 males, 3 females) but were excluded from the sample due to failure in reaching minimum looking times during manipulation (8) or missing data in one or more conditions during baseline or test (1). On average, each infant provided data for a median of 8 (range 5 – 11) out of 12 trials.

Apparatus, Stimuli, Procedure and Data Processing were identical to Experiment 1. As in Experiment 1, we checked that the gaze patterns in the test phase did not differ systematically between the two manipulations point and no point ($F_{(1,16)}$ = .031, p = .862, n_p^2 = .002) and that there was no interaction with outcome ($F_{(1,16)}$ = .840, p = .373, n_p^2 = .050). On average, 72.31% of all gaze points (range = 62.53% – 81.11%) fell into the central AOI of the door area during the test event.

2.5.2 Results

Figure 2.8 displays the mean pupil diameters for the four conditions across a trial. As in Experiment 1, visual inspection revealed a difference in pupil size right after the manipulation, already before the outcome. The pointing manipulation lead to smaller pupils during the still phase right before the test phase than the control manipulation ($M_{\text{“point”}}$ = 3.52 mm, $SD_{\text{“point”}}$ = .51 mm, $M_{\text{“no point”}}$ = 3.69 mm, $SD_{\text{“no point”}}$ = .59 mm, $t_{(16)}$ = 6.19, $p <$ .001).

Similar to our results in Experiment 1, our looking time analysis revealed that in point trials, infants looked significantly longer at the screen than in the no point trials (Fig. 2.10a, $M_{\text{“point”}}$ = 913.93 ms, $SD_{\text{“point”}}$ = 103.78 ms, $M_{\text{“no point”}}$ = 725.20 ms, $SD_{\text{“no point”}}$ =

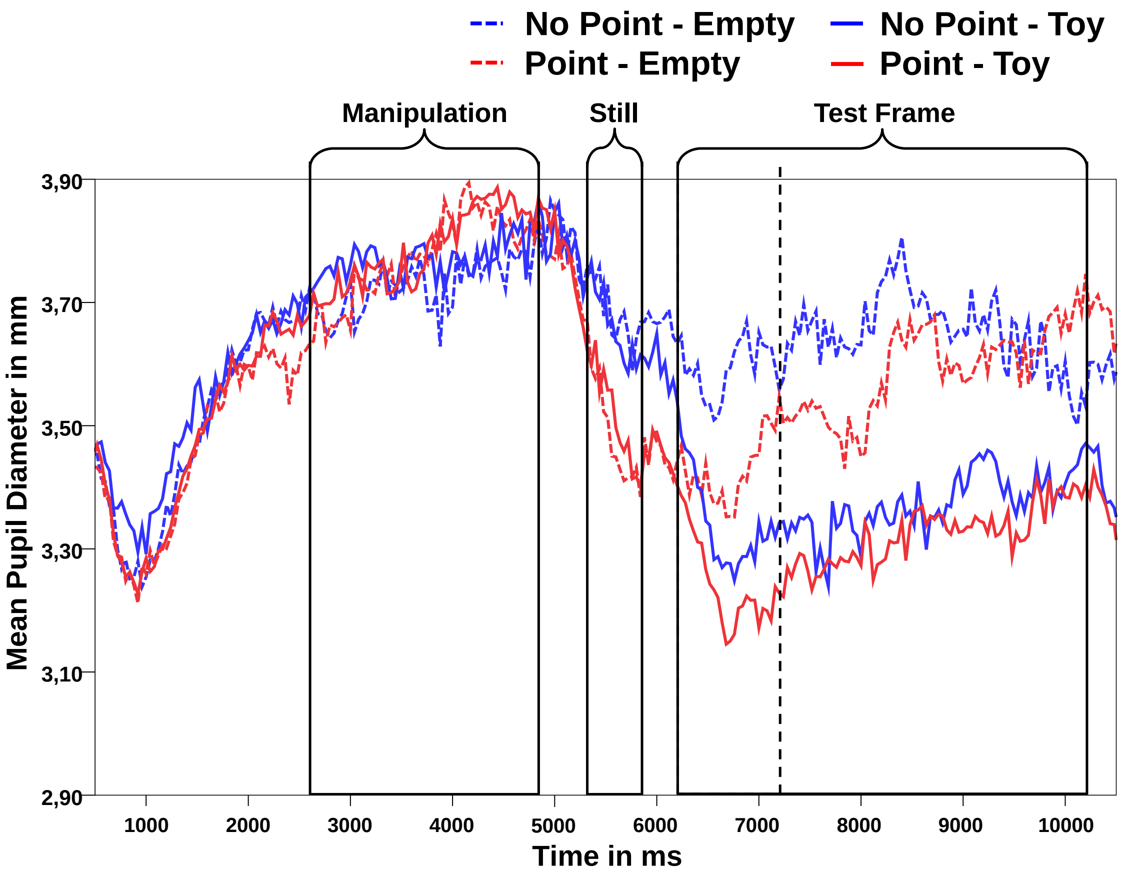


Figure 2.8: Mean pupil diameter of 8-month-olds in Experiment 3. Windows are shifted by 500 ms compared to the timeline in Fig. 2.1.

214.41 ms, $t_{(16)} = 3.92$, $p = .001$). Again, we suspect that when infants looked away they focused on the darker test booth, which may explain why the pupils in the no point trials were larger than in the point trials.

To test our main question, we applied the same baseline correction as in Experiment 1. We performed 2 (gesture) x 2 (outcome) repeated measures ANOVAs to get a closer look at how the conditions interacted. Table 2.3 shows the results of both the 1000 ms and the 4000 ms time windows. There are main effects for both outcome and gesture in both time windows, but no interaction between the two conditions.

Paired-samples *t*-Tests following our hypotheses revealed that the relative changes in

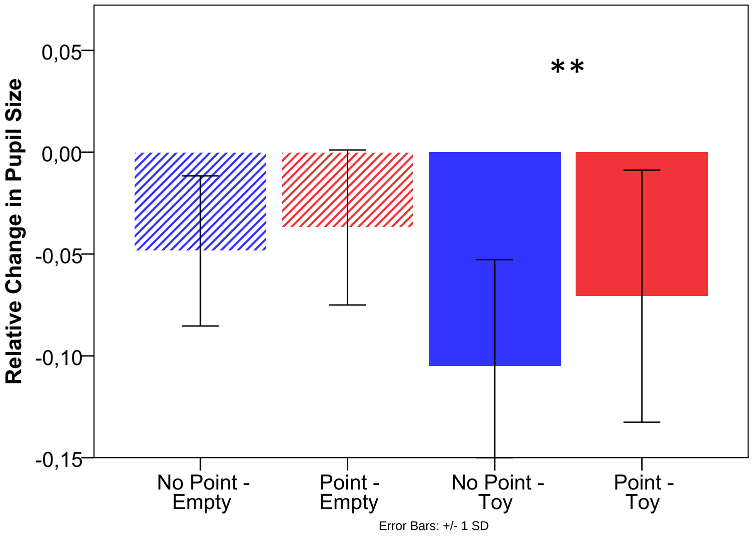


Figure 2.9: Change in pupil size of 8-month-olds in Experiment 3 relative to baseline. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

pupil size in the empty conditions were not significantly different after a point compared to after no point (M$_{\text{“point - empty”}}$ = -.037, SD$_{\text{“point - empty”}}$ = .038, M$_{\text{“no point - empty”}}$ = -.049, SD$_{\text{“no point - empty”}}$ = .037, $t_{(16)} = 1.31$, $p = .207$, see Fig. 2.9). Pupil size in the toy conditions was even significantly less decreased after a point compared to after no point (M$_{\text{“point - toy”}}$ = -.071, SD$_{\text{“point - toy”}}$ = .062, M$_{\text{“no point - toy”}}$ = -.105, SD$_{\text{“no point - toy”}}$ = .052, $t_{(16)} = 2.84$, $p = .012$).

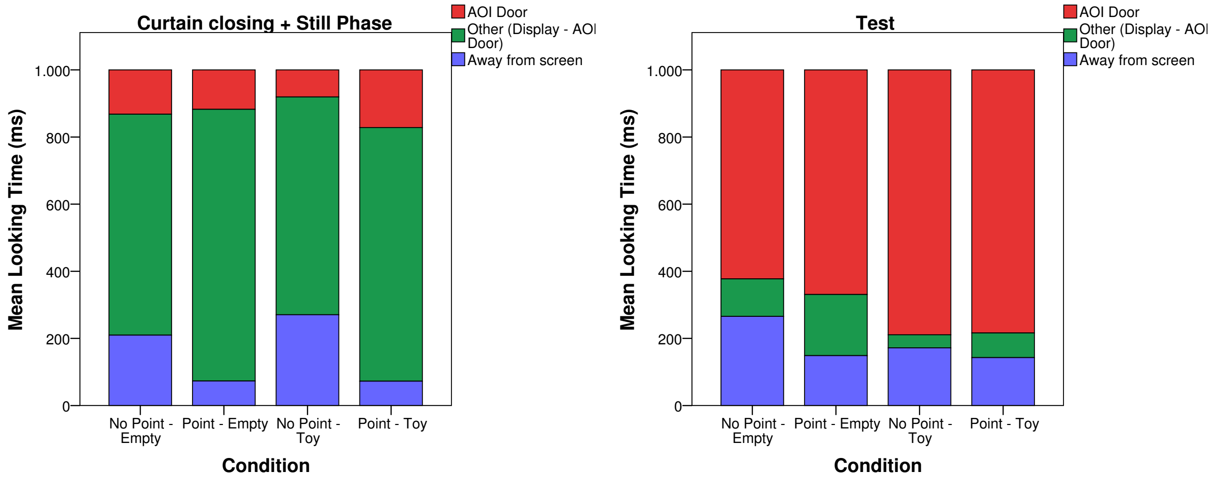
Table 2.3: ANOVA of time segments within test phase for 8-month-olds in Experiment 3.

Time Interval	Within-subject factors	*df*	*F*	*p*	η_p^2
0 – 1000 ms	Gesture	16	8.34	.011*	.343
	Outcome	16	23.72	.000***	.597
	Gesture*Outcome	16	2.68	.121	.143
0 – 4000 ms	Gesture	16	8.06	.012*	.335
	Outcome	16	32.87	.000***	.673
	Gesture*Outcome	16	2.14	.163	.118

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Paired comparisons between all conditions, Bonferroni-corrected with $n = 6$, show that two other comparisons become significant: ‘no point – empty’ and ‘point – empty’ are each significantly bigger than ‘no point – toy’ ($t_{(16)} = 4.89$, $p < .001$ and $t_{(16)} = 5.67$, $p < .001$; respectively).

To confirm the results from the pupil data, we again analyzed looking times to different AOIs during the test phase. A 2 (gesture) x 2 (outcome) repeated measures ANOVA revealed that infants looked more towards the door when there was a toy than when it was empty (main effect for outcome, $F_{(1,\ 16)} = 5.98$, $p = .026$, $n_p^2 = .272$) but no main effect for gesture and no interaction between gesture and outcome, thus excluding that the pupil data were an artifact of difference in looking patterns (see Fig. 2.10b). As in Experiments 1 and 2, paired comparisons between conditions revealed no significant differences between any conditions in terms of looking time to the video as a whole or looks away from the screen.



(a) Baseline phase.

(b) 1000 ms of test phase.

Figure 2.10: Distribution of looking time of 8-month-olds in Experiment 3 during baseline and test.

We repeated the analysis for the test phase of 4000 ms to check for a later emergence of the effect. Paired-samples *t*-Test following our hypotheses revealed that relative change in pupil size in the empty conditions were not significantly different after a point compared to after no point ($\mathrm{M}_{\text{“point – empty“}} = \text{-.009}$, $\mathrm{SD}_{\text{“point – empty“}} = .044$, $\mathrm{M}_{\text{“no point – empty“}} = \text{-.023}$, SD“no point – empty“ = .042, $t_{(16)} = 1.06$, $p = .304$). Relative change size in the toy conditions was significantly less decreased after a point compared to after no point (M“point – toy“ = -.059, SD“point – toy“ = .045, M“no point – toy“ = -.096, SD“no point – toy“ = .052, $t_{(16)} = 3.44$, $p = .003$).

Infants looked longer to the door area over the duration of 4000 ms when there was a toy than when it was empty (main effect for outcome: $F_{(1,\ 16)} = 23.81$, $p < .001$, $n_p^2 = .598$), however, regarding our crucial comparisons, the analysis revealed no significant differences between any conditions of the same outcome in terms of looking time in total or looking away from the screen.

Because the paradigm in Experiment 3 was identical to the one in Experiment 1, we additionally performed a mixed ANOVA with gesture and outcome as within-subject factors and age as a between-subject factor. We found a three-way interaction between gesture, outcome and age ($F_{(1,32)} = 4.55$, $p = .041$, $n_p^2 = .124$), main effects for gesture ($F_{(1,32)} = 11.83$, $p = .002$, $n_p^2 = .270$) and outcome ($F_{(1,32)} = 28.12$, $p < .001$, $n_p^2 = .468$), a tentative main effect for age ($F_{(1,32)} = 4.09$, $p = .052$, $n_p^2 = .113$) and no other interactions.

2.5.3 Discussion

8-month-olds showed no difference in relative change of pupil size from baseline between the expectation-violating conditions and the control conditions. Most tellingly, and in contrast to the 12-month-olds of Experiment 1, 8-month-olds did not differentiate between the ‘point – empty’ condition and the ‘no point – empty’ condition. Importantly, across our focal comparison conditions, the 8-month-olds watched the videos for the same amount of time, and paid attention to the same areas of interest, thus excluding perceptual differences as one explanation. By the same standards of Experiment 1, the finding thus suggests that pointing does not induce object expectations in 8-month-old infants.

Visual inspections suggest further differences in infants’ processing of the point during the manipulation phase. While 12-month-olds' pupils began contracting after about 1800 ms, as the pointing finger was outstretched at its farthest point, 8-month-olds' pupils did not contract during the point in the manipulation phase. Twelve-month-olds' shape may be reminiscent of that seen in adult problem solving, when the problem has been solved (Bradshaw, 1968; Hess & Polt, 1964). Eight-month-olds' shape, in contrast, appears reminiscent of that seen in adults when more difficult problems have not yet been solved. By analogy, 12-month-olds recognized the pointing gesture, and it was immediately telling to them, such that cognitive load decreased because they could 'solve' the problem of what would happen next (i.e. they developed an expectation). Eight-month-olds, instead, saw the pointing gesture, but processed it with increasing effort and did not 'solve' for a prediction. Instead, the manipulation phase ended, leaving the infant puzzled as to what to expect next.

The current finding of a developmental difference in the emergence of communication-induced object expectations contradicts previous findings which have suggested similar performance at 8 and 12 months of age (Csibra & Volein, 2008). One should note, however, that the latter study did not test each age group separately, and that it remained amenable to a leaner attentional highlighting interpretation. Our current finding is thus less supportive of views on which reference comprehension emerges early as a unitary complex system (Csibra, 2003) and instead provides room for developmental accounts on which several factors may interact to yield an increasing refinement in the development of understanding gestural reference.

2.6 General Discussion

The current study investigated the complexity and development of infant point comprehension across three experiments. The main question was whether the observation of a point led infants to expect a referent object. The alternative was that infants simply follow the vectorial direction of a point and only then encounter an object. Findings from Experiment 1 rejected the lean alternative for 12-month-old infants and revealed that referential communication with a pointing gesture induces the expectation of an object, prior to encountering it. Findings from the control Experiment 2 revealed that the effect of induced object expectation was specific to the pointing gesture and absent for a non-social endogenous cue, thus excluding general attention-directing or attention-highlighting processes as one alternative explanation of the effect. Results from the younger age group of Experiment 3 revealed that pointing did not induce object expectations at 8 months, an age at which infants themselves do not yet produce the canonical index-finger pointing gesture, suggesting a protracted development across the first year of life. The current study thus provides new evidence for a cognitively 'rich' interpretation of infant pointing at 12 months of age, while at the same time suggesting leaner alternatives to apply in the first year of life.

Our study established pupillometry as a measure of cognitive expectations in a referential comprehension task. Pupil dilation may be especially apt as a measure of cognitive expectations in infants because it allows controlling for the direction and amount of attentional allocation, and does not require additional motor skills or motivational components to elicit behavior. Although infants looked at the same perceptual stimuli for the same amount of time, the pupillary response still revealed differences in their processing of the scenes. Differently to looking time measures typically employed in VOE paradigms, the pupillary measure was more closely linked to the online processing of the situation and did not require a lengthy test phase, thus allowing for multiple trials and reducing variance. Nevertheless, looking time analyses seem necessary as a control measure to pupillary responses, especially because infants may not focus equally on the same stimuli across conditions. In that respect, our looking-time control analyses were crucial in establishing the same amount of attentional allocation and direction across our focal comparison groups. Our findings based on the pupillary measure thus provide less equivocal evidence for the referential interpretation of infant point comprehension than previous interaction-based and looking-time studies (Behne et al., 2012; Csibra & Volein, 2008).

The present study did not differentiate between mentalistic and non-mentalistic versions of the referential interpretation of pointing: Does the infant attribute a referential intention – and thus a mental representation of the target object – to the pointer (Tomasello et al., 2007), or does the pointing gesture elicit an object expectation in the infant, irrespective of the pointer's mental representation (Csibra, 2003)? It has been notoriously difficult to distinguish between these two possible interpretations, because the pointer's referential intention and reality typically conflate. What would be needed is to somehow manipulate the pointer's visual access to the object and decouple her referential intention from reality. The test would then be to see whether infants form expectations according to their view, or according to the pointer's referential intention. One interaction-based study to date has revealed positive evidence for a distinction at 17 months of age (Southgate, Chevallier, & Csibra, 2010, but see Dörrenberg, Rakoczy, & Liszkowski, 2018, for negative evidence at 24 months of age). Given the absence of decisive studies on this question for 12-month-olds, as well as the conceptually more parsimonious appearing non-mentalistic interpretation, it is less clear to what extent 12-month-olds comprehend referential intentions. A recent EEG study on object representations suggests that a representation of another person's mental object representation may be within 8-month-olds' capacity (Kampis et al., 2015).

The importance of the finding of a developmental difference between 8 and 12 months of age cannot be overstated. Current findings clarify that point comprehension does not emerge as one parcel early in ontogeny but instead likely undergoes developmental change: Very early point-following in the first half of the first year appears rather reflexive (Bertenthal et al., 2014). In the early part of the second half of the first year behavioral interaction-based point-following responses are governed only by the directional nature of the gesture, but not yet by referential expectations. Only toward the end of the first year of life, as the current findings attest, does gestural reference induce object expectations. Given this developmental gradient, it is plausible to assume that a cognitively even more advanced distinction between intended and non-intended reference (i.e. based on false belief) still awaits further development within the second or third year of life (Liszkowski, 2017).

Here, we can only speculate about the developmental mechanisms of change between 8 and 12 months of age. While infants certainly gain more experiences with processing pointing during this period (see Salomo & Liszkowski, 2013), another major developmental change pertains to the emergence of point production in that period (see Liszkowski & Tomasello, 2011; Behne et al., 2012). And so, while the current study demonstrates that by 12 months of age infants partake in a ‘meeting of minds’ which forms a basis for further cultural learning and socialization, this ‘meeting of minds’ is perhaps as much a product of socialization and earlier forms of social interacting in the first year of life.

3 Pupillometry reveals spontaneous expectation of objects following occlusion[1]

RESEARCH HIGHLIGHTS:

- *We used pupillometry to measure violations of expectations in infants.*
- *18-month-olds' pupils dilated when a door frame was closed containing an object, and opened up to be empty.*
- *10-month-olds did not show pupil dilation in response to the violation of expectation.*
- *Neither age group responded with pupil dilation to the violation of an emptiness expectation.*

In Chapter 2, I demonstrated that 12-month-olds expect an object to appear at a space that had been referred to by a pointing gesture. But what does it mean for infants to have an object expectation? The study on object permanence has provided a lot of evidence using many different methods (see Section "Object Permanence" in Chapter 1). In order to reinforce the use of pupil dilation as a valid measure of object expectation, here I present evidence on object permanence in a VOE paradigm using pupillometry. Whereas in Chapter 2, infants' pupils dilated in response to the unexpected absence of an object whose presence was previously communicated, here, expectation of presence or absence was manipulated through an action sequence. An object was either shown to be occluded within the scene (consistent with a continued existence behind the occluder) or removed

---

[1] This study is in preparation for submission entitled "Pupillometry reveals spontaneous expectation of objects (but not empty sets) following occlusion at 18 months but not at 10 months" (Pätzold & Liszkowski).

from the scene with a visibly empty occluder (inconsistent with a continued existence behind the occluder). Expectation of an object therefore was not directly induced by a cue, but spontaneously arose while the infant watched the scene unfold. Previous studies with similar VOE paradigms using implicit measures, including anticipatory looking (Ruffman et al., 2005), pupillometry (Sirois & Jackson, 2012) and EEG measures (Kaufman et al., 2003) have provided mixed evidence of the understanding of the continued existence of objects in the first year, whereas studies using global looking time measures show stable evidence of object permanence from as young as 3.5 months with various paradigms (Baillargeon, 1987; Baillargeon & DeVos, 1991; Baillargeon, 1998).

In Experiment 1, I established this novel paradigm of measuring object permanence with 18-month-olds, whose pupils dilated in response to the violation of object expectation. However, in Experiment 2, I failed to find the same effect in 10-month-olds. These results were unexpected, especially in the light of successful cued object expectation in 12-month-olds in Chapter 2. I discussed the results with regard to previously used object permanence paradigms and the reasoning that studies reporting early competence also rely on additional social cues that facilitate a build-up of expectation (Kaufman et al., 2003; Ruffman et al., 2005), whereas the current paradigm relies on a spontaneously occurring expectation. Therefore, the current findings offer an interesting clue of the interplay between social cognition and object expectation, which will be further explored in Chapters 4 and 5.

Acknowledgements: We thank all families who participated in these studies. We thank Celine Staack for her help in preparing the stimuli. We thank Nicola Ballhausen, Jula Brüning-Wessels, Sriranjani Karthik, Jana Klose and Lydia Rihm for their assistance with data collection. We thank Susanna Jeschonek-Seidel for valuable feedback in the early stages of the project.

3.1 Abstract

Depending on the measures used, there remains ambiguity concerning the emergence of object permanence in infants. Looking time studies have established early competence in representing occluded objects as young as 3 months of age, but are susceptible to attentional confounds. Recently, a variety of new cognitive measures, including EEG power band analysis, anticipatory looking and pupillometry have presented mixed evidence of object permanence in the first year.

In the current study, we continued to explore pupil dilation as a measure of object permanence. Infants watched videos of animated toys either stopping in an open frame or moving across the screen out of sight. The frame then closed for 2 s and opened up again to reveal either the toy, or an empty space, which was either consistent or inconsistent with the continued existence of the object depending on the previous motion sequence.

In Experiment 1, 18-month-olds's pupils dilated in response to the unexpected empty outcome, establishing the paradigm as a suitable measure of violation of object expectation. There was no difference in pupil dilation in response to the violation of an empty place, suggesting that infants do not encode absence of an object in the same way as its presence.

In Experiment 2, we did not find pupil dilation in the VOE condition in 10-month-olds, which is surprising considering both the success of the paradigm in the older group and the previous evidence of object permanence at 10 months using other measures.

The results are discussed in regard to the social-communicative context of some VOE paradigms, which may inadvertently have been an important cognitive facilitator for young infants, but was not present in the paradigm reported here.

Keywords: Pupillometry, Object permanence, Occlusion events, Violation of Expectation

3.2 Introduction

Object permanence, the ability to represent an object's continued existence when it is not visible (Piaget, 1954) has been the focus of cognitive development research for many decades. While classic behavioral studies reliably found that infants grasp for occluded objects at the earliest around 8 — 10 months but still commit search errors when an object is visibly hidden (Diamond, 1990; Wellman & Miller, 1986), these findings have been criticized to underestimate infants' competence because the dependent measure relies on additional complexities of coordinating action and knowledge (Miller et al., 1970; Willatts, 1984). Subsequent research has used violation-of-expectation (VOE) paradigms with looking time as dependent measure and suggests that object permanence emerges as early as 3 months (Baillargeon, 1987; Baillargeon et al., 1985). However, these VoE studies have been criticized to overestimate infants' competence because the looking-time paradigm is susceptible to lower level perceptual accounts, e.g. perceptual salience and novelty of the test event (Bogartz et al., 1997; Haith, 1998; Meltzoff & Moore, 1998).

For example, Dunn and Bremner (2017) recently showed that six-month-old infants looked longer at test trials with a violation of object expectation, but also at test trials with a novel object, thus making it hard to distinguish whether a longer looking effect is driven by object expectations or simply by longer processing of a novel object. In the classic 'drawbridge' paradigm (Baillargeon et al., 1985) it has been argued that infants may prefer the impossible test event simply because of its inherent properties (Rivera et al., 1999; Sirois & Jackson, 2012). Sirois and Jackson (2012) used a full factorial design and habituated infants not only to the 180° arc of the moving occluder, as in the original study (Baillargeon et al., 1985), but also to the 120° moving occluder displayed during test events. They found the original looking time effect, i.e. longer looking to the impossible event when the occluder moved 180° backwards despite a box being behind it; however, they also found that overall, infants looked longer at events involving the 180° angle (irrespective of habituation) and at events involving the box (irrespective of the impossibility of the event). The authors concluded that the original looking time effect does not warrant a conclusive interpretation in terms of object permanence. Others have further argued that the looking time effect depends on a fragile balance between familiarity vs. novelty preference, for example when varying the number of familiarization trials (Hunter & Ames, 1988; Schilling, 2000, but see S.-H. Wang, Baillargeon, & Brueckner, 2004).

One way to address the validity of object permanence findings is to employ diverse approaches and measures that have the potential to concur on the same interpretation. A promising measure that circumvents the interpretative problems of behavioral reaching and looking time measures is pupil dilation. Because pupils dilate not only to changes in luminance but also with regard to cognitive effort, one can show participants the exact same perceptual event under different cognitive expectations and measure involuntary changes in pupil diameter as a function of cognitive processing, while excluding lower level perceptual effects (Bradshaw, 1968; Hess & Polt, 1960, 1964; van der Meer et al., 2010). Pupillometric studies of object permanence, however, are still sparse and have yielded equivocal findings. One of the two to date published studies is the drawbridge replication study by Sirois and Jackson (2012), in which the authors also measured pupil dilation. In accordance with the perceptual interpretation of their looking time results, the authors did not find a selective effect of pupil dilation for the conditions which would violate cognitive expectations. These findings shed serious doubts on the robustness of object permanence at 10 months of age. However, Sirois and Jackson (2012) did not provide positive evidence that their paradigm works with older children, making it difficult to interpret the absence of evidence. However, in another pupillometric study by Jackson and Sirois (2009), the authors found indications that 8-month-olds do react to violations of the continued identity of objects, suggesting that pupil dilation is a sensitive measure to detect aspects of object expectations in infants.

In the current study, we followed up on the pupillometry method to test whether infants show object permanence in a simplified visual VOE task. Instead of the moving drawbridge violation we used static outcomes to avoid additional demands and potential confounds of processing and preferring distinct movements of the drawbridge. In a fully balanced 2 x 2 factorial design we showed infants an occluder that revealed either an object or an empty location. The appearances and disappearances of objects were either expected or unexpected because either an object or an empty location were shown before occlusion. This enabled us to compare perceptually identical outcomes that were thus also identical in luminance and differed only in their cognitive expectedness. Our main question concerned the comparison between the expected and unexpected disappearance of an object, that is, when the occluder opened to reveal an empty location. In the unexpected (violation) condition, an object that was previously present failed to appear after occlusion. In the expected condition, an object that had previously disappeared remained absent after occlusion. Success would require representing the object during its occlusion and expecting it to reappear, and the violation should yield larger pupils compared to the expected event.

We prevented infants from always expecting an empty location outcome by alternating empty location and toy outcomes across trials. Therefore, a secondary possibility for a comparison arose regarding the outcomes when an object was present. Thus, in the unexpected (violation) condition, an object that had already disappeared before occlusion re-appeared after occlusion; and in the expected condition, an object that had been present before occlusion remained there after occlusion. Success would require representing and expecting the ‘nothingness’, or empty set. Pupils should dilate when an object was displayed where an empty location was supposed to be. There is less literature on the latter topic, but it appears that infants are not surprised to see objects suddenly appear compared to when objects disappear. In a study by Wynn and Chiang (1998), eight-month-old infants did not look longer at an unexpected appearance event compared to an expected appearance, both in a scene with two toys (one that was expected to be there, and one that should not have been there) and a scene with a single toy (which should not have been there). The concept of "emptiness" appears far from trivial and may not be found until much later in childhood (Bialystok & Codd, 2000; Wellman & Miller, 1986). We therefore included the unexpected appearance comparison in our analyses only as a secondary question.

To test our main question whether pupillometry would reveal object permanence in infants, we tested in Study 1 an older age group of 18-month-olds for which we expected firm object permanence given their skills at searching for and communicating about absent objects. 18-month-olds' pupil should be larger in the unexpected condition when an object that should be there had suddenly disappeared, compared to the identical outcome scene when it was expected that the object should not be there. In Study 2 we then moved on to our target age group of 10-month-olds for which pupillometric findings of object permanence are more contested.

3.3 Experiment 1

3.3.1 Methods

3.3.1.1 Participants

Nineteen 18-month-old infants (10 males, 9 females) were included in the final sample. Mean age was 18 months, 13 days (range: 18 months, 3 days -– 18 months, 26 days). Six additional infants participated (3 males, 3 females) but were excluded from the sample due to failure in reaching minimum looking times during manipulation (4), fussiness (1) or technical failure (1).

All infants were recruited via birth records and had a middle to high socioeconomic, western cultural background. Infants were included in the final analysis when they provided data for at least one trial of each condition. Trials were included when looking times indicated that the infant had watched at least 50% of the manipulation. On average, infants provided data for 13.63 trials of out 16 trials (range 6 – 16).

3.3.1.2 Apparatus

Pupil dilation was measured with a Tobii x120 eye tracker (Tobii Technology, Stockholm, Sweden) which was attached to a standard computer screen. The presentation screen and the eye tracker were placed in a testing booth built for this purpose, with black canvases behind and on both sides of the screen. A video camera above the screen allowed the researchers to monitor the infant's behavior during the session. The size of the stimuli presentation was 1280 x 1024 px on a 1920 x 1200 px screen, with the rest of the screen appearing black throughout the experiment. The size of the display was 34.50 x 27.50 cm on the screen, which corresponds to a visual angle of 24.25° horizontally and 30.16° vertically.

3.3.1.3 Stimuli

The videos consisted of computer-animated clips of a stage-like scene (see 3.1). Centered in the video was a closed light-brown door frame. All videos start with the door opening from top to bottom to reveal the empty inside, which is lined with a bright turquoise checking pattern (500 ms). The pattern implicates depth inside the frame and also serves to balance out the luminance difference between an empty outcome and an outcome containing one of four toys: a boat, a cow, a train and a horse. Immediately following the complete opening of the door, one of the toys started moving into the screen from one side (left/right was counterbalanced so that each infant saw each condition from each side the same amount of time). The movement was accompanied by a sound typically associated with that toy in an animation (boat gliding on water for the boat, cow bells for the cow, railroad for the train, clip-clopping of hooves for the horse). After 3750 ms, the toy had moved into the middle of the screen. In half of the videos, the toy remained in the middle of the screen until it was occluded by the closing door (3250 ms). In the other half of the videos, the toy did not stop but continued its trajectory until it was completely out of sight. As it was leaving the screen, the door closed. The timing of the door closing was matched so that in both versions, infants saw the toy the exact same amount of time. In all videos, the second part of the toy animation – staying still or moving out – was accompanied by another toy-specific sound (fog horn for the boat, mooing for the cow, choo-choo for the train, neighing for the horse). The door then remained closed for 2000 ms before opening again, accompanied by the sound of a drumroll, to reveal the outcome: either the same toy as before or no toy at all (500 ms). The outcome was presented for 2000 ms before the door closed again (500 ms) and stayed closed for the remainder of the video (4000 ms). Sound was included to keep infants engaged with the videos. The timing of sounds was identical in all conditions.

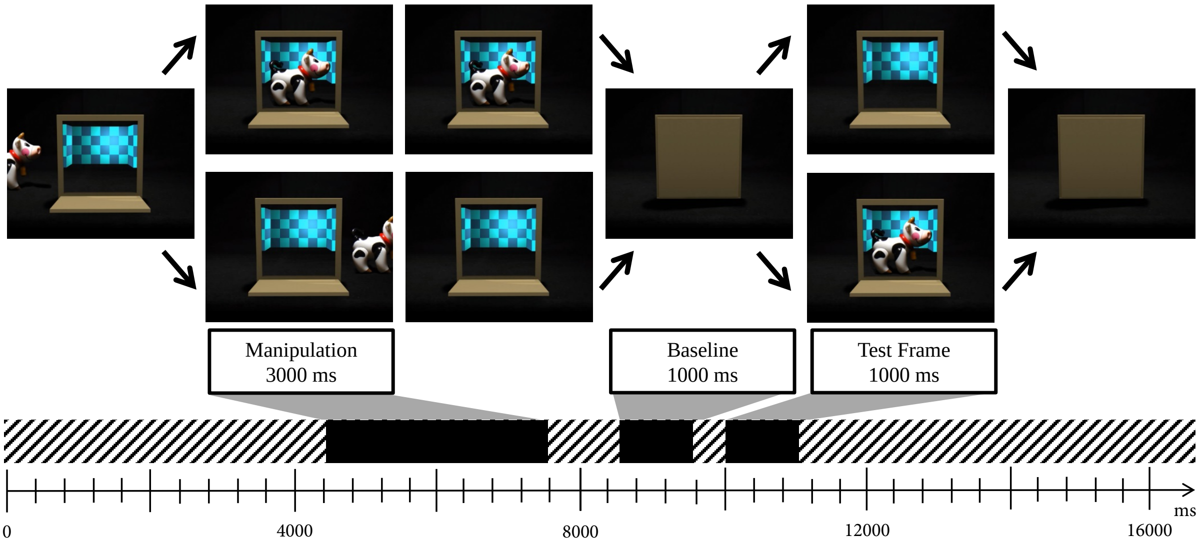


Figure 3.1: Scheme of the experimental design for Experiments 1 and 2.

Infants were presented with 16 videos, four of each condition. Appearance of toys was blocked in the same order for all infants, whereas order of conditions was semi-random: no same manipulation or outcome could be presented more than twice in a row. Each of the infants saw each toy in all four conditions, at different times in the study. Before each video, a short attention getter was played to redirect the infant's attention towards the middle of the screen (3 s). After a block of four videos, a longer attention getter was played to break up the routine of the videos (8 s).

Because we averaged data from several trials with different objects, we took care to match luminance levels of the four objects as closely as possible. Following the formula used in Jackson and Sirois (2009), we calculated photometric luminance from the red, green and blue levels of the images as given by a photo analysis software. On a scale from 0 to 1 with 1 being the brightest, luminance levels read .252 for the empty box, .272 for the boat, .261 for the cow, .236 for the train and .241 for the horse. Thus, the biggest difference in luminance was between the train and the boat with the boat being 3.7% brighter than the train. The empty box was exactly at the mean of luminance of the full boxes.

3.3.1.4 Procedure

The experimenter explained the study to the caregiver and obtained their informed consent. The caregiver was seated on a swivel chair in the testing booth 64 cm away from the presentation screen with the infant on his or her lap. A 9-point-calibration ensured that the infant's eyes were properly captured by the eye tracker. The experimental stimuli were then presented on the screen as long as the infant was willing to watch or until 16 trials were completed. The caregivers were instructed to neither talk to nor point for the infant and to keep their eyes closed for the duration of the experiment.

3.3.1.5 Data Processing

Pupil size and gaze location of both eyes were recorded at a 120 Hz rate. Pupil size from the left and right eye was averaged to mean pupil size. If data from one eye were missing, the data from the other eye were used. If data from both eyes were missing, no substitution was made.

Gaze data were processed in relation to an a priori defined area of interest (AOI) of the door frame (500 x 500 px). Gaze data at the entire display (1280 x 1024 px) were also calculated in order to allow analysis of looking time. Time looked away from the screen was defined as Maximum Looking Time – Looking Time to Display.

After careful visual inspection, pupil data were shifted to the right by 500ms. Infants' pupils react slower than adults (Verschoor et al., 2013, 2015) and therefore need more time to adapt to a change in stimulus. Also, by not including the phases of pupillary constriction (PC) following a stimulus onset in the average, a more accurate representation of the tonic pupil dilation is achieved (see Hepach & Westermann, 2016).

Because pupil data collected by remote eye trackers such as the Tobii X120 may be susceptible to error introduced by gaze direction (Brisson et al., 2013), we checked that the gaze patterns did not differ systematically from one condition to the other during the test phase. 68.89% of all gaze points (range 64.40% – 73.55%) fell into the central AOI of the door area during the test event, with no significant differences between conditions ($F_{(1,18)} = 2.649$, $p = .084$).

3.3.2 Results and Discussion

Figure 3.2) displays the time line of pupil dilation during baseline and test. A priori analyses of the baseline revealed that there was no difference in pupil dilation between the two types of manipulation at baseline ($\text{M}_{\text{“occlusion“}} = 4.83$ mm, $\text{SD}_{\text{“occlusion“}} = .55$ mm,

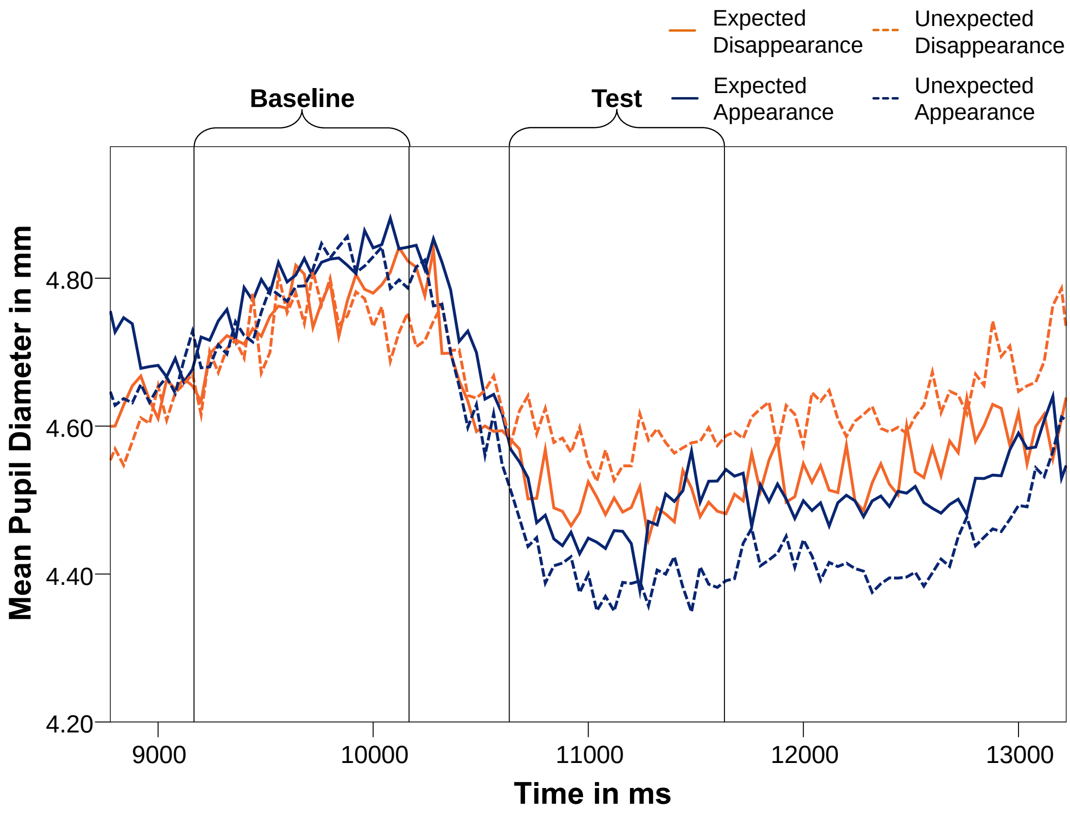
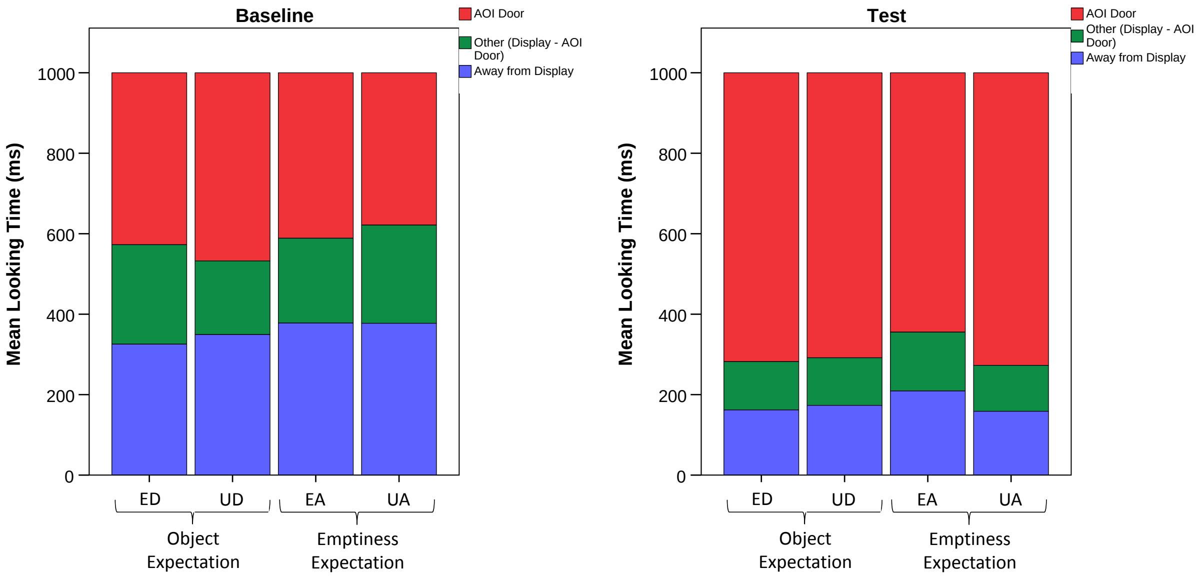


Figure 3.2: Mean pupil diameter in 18-month-olds in Experiment 1. Windows are shifted by 500 ms compared to the timeline in Fig. 3.1.

M$_{\text{„removal“}}$ = 4.87 mm, SD$_{\text{„removal“}}$ = .53 mm, $t_{(27)}$ = -.83, p = .416). Figure 3.3a further shows that infants looked during baseline for the same amount of time at the video, with no significant differences between the two types of manipulation (M$_{\text{„occlusion“}}$ = 622.49 ms, SD$_{\text{„occlusion“}}$ = 206.30 ms, M$_{\text{„removal“}}$ = 640.37 ms, SD$_{\text{„removal“}}$ = 177.99 ms, $t_{(18)}$ = -.57, p = .576). These results warrant the use of the baseline to calculate the relative change of pupil size for the different outcomes.

Figure 3.4 shows the results of the relative change score analysis at test, averaged over the duration of the first 1000 ms. A 2 (outcome) x 2 (expectation) repeated-measures ANOVA revealed a main effects for outcome ($F_{(1,\ 18)} = 19{,}39$, $p < .001$, $n_p^2 = .519$) no main effect for expectation and a significant interaction between outcome and expectation ($F_{(1,\ 18)} = 12.97$, p = .002, n_p^2 = .419). Regarding our hypothesis about the disappearance of the object,



(a) Baseline phase.

(b) Test phase.

Figure 3.3: Distribution of looking time during baseline and test of 18-month-olds in Experiment 1.

direct comparisons showed that infants' pupils dilated to the empty outcome significantly more when it was unexpected than when it was expected (expected disappearance = ED, unexpected disappearance = UD, $\text{M}_{\text{“ED“}}$ = -.054, $\text{SD}_{\text{“ED“}}$ = .046, $\text{M}_{\text{“UD“}}$ = -.034, $\text{SD}_{\text{“UD“}}$ = .039, $t_{(18)} = 3.24$, $p = .005$). Regarding the secondary question about the appearance of the object, there was no difference in infants' pupil sizes in response to the toy outcomes (expected appearance = EA, unexpected appearance = UA, $\text{M}_{\text{“EA“}}$ = -.070, $\text{SD}_{\text{“EA“}}$ = .056, $\text{M}_{\text{“UA“}}$ = -.082, $\text{SD}_{\text{“UA“}}$ = .049, $t_{(18)} = 1.54$, $p = .139$).

To control for attentional differences, we checked the gaze pattern during the test phase (Fig. 3.3b). Infants looked at the videos for the same amount of time during the first 1000 ms of the expected and unexpected empty outcomes ($\text{M}_{\text{“ED“}}$ = 812.06 ms, $\text{SD}_{\text{“ED“}}$ = 237.94 ms, $\text{M}_{\text{“UD“}}$ = 790.20 ms, $\text{SD}_{\text{“UD“}}$ = 224.44 ms, $t_{(18)} = -2.059$, $p = .304$) and during the expected and unexpected toy outcomes ($\text{M}_{\text{“EA“}}$ = 785.16 ms, $\text{SD}_{\text{“EA“}}$ = 238.91 ms, $\text{M}_{\text{“UA“}}$ = 835.31 ms, $\text{SD}_{\text{“UA“}}$ = 212.65 ms, $t_{(18)} = -.95$, $p = .353$). There were also no differences

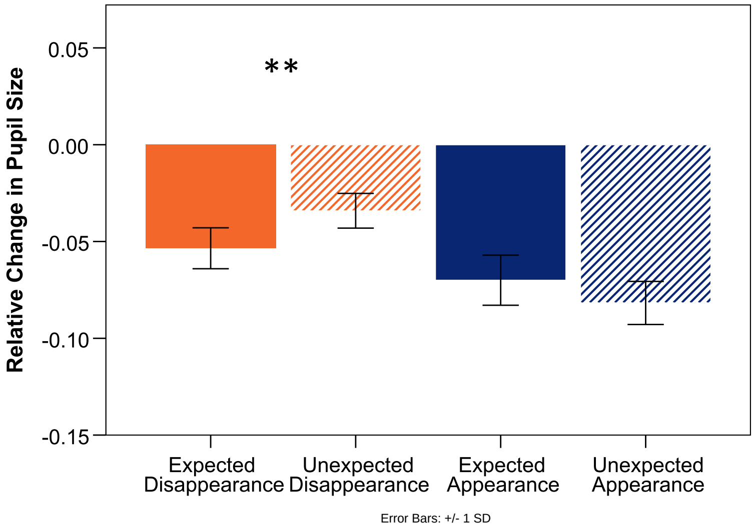


Figure 3.4: Change in pupil size of 18-month-olds in Experiment 1 relative to baseline. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

in looking towards the door or looks away from the screen when comparing the expected and unexpected outcomes ($p > .127$).

The current results thus reveal that infants' pupil dilation increases in response to an unexpected disappearance of an object, i.e. when an object does not re-appear after a temporary occlusion. Importantly, infants' pupils dilated significantly less when the object expectedly did not appear after occlusion (because it had already disappeared before occlusion). This excludes the possibility that purely perceptual differences drove the effect. Our further control analyses confirmed that the effect was neither due to differences in pupil sizes at baseline nor explainable by differences in visual attention during baseline or test. The pattern of findings supports an interpretation in terms of object permanence and provides confirmatory evidence to the established view of object representations in the second year of life. The new paradigm thus proves to be a useful test in detecting object permanence while it circumvents the interpretative challenges of habituation paradigms and looking time measures.

The current experiment did not find a differential effect of pupil dilation to the unexpected appearance of an object, i.e. when an object appeared after temporary occlusion of an empty location. Given that our control measures did not reveal any differences that could have accounted for the absence of the effect, we assume that 18-month-olds do not represent empty sets, at least in the current paradigm. While the absence of evidence certainly does not reveal evidence for absence, the lack of an effect for the unexpected appearance has to be interpreted in the context of our positive effect for the unexpected disappearance. In line with the current literature (Nieder, 2016), we therefore think that the representation of empty sets is beyond 18-month-olds' capacities, at least when compared to their ability of representing objects, as our interaction result must suggest.

3.4 Experiment 2

Given that our paradigm revealed object expectations at 18 months, in Experiment 2 we moved on to test the more contested target age group of 10-month-olds to probe whether they, as a group, would show evidence of object expectations in our new paradigm, and how it would compare to the current results with the 18-month-old infants. If 10-month-olds indeed represent occluded objects in our paradigm, they should dilate their pupils significantly more in the unexpected disappearance event compared to the expected disappearance event. Because the effect could be smaller in younger infants, we increased our *N*. In order to prevent infants from always expecting an empty outcome we included again the toy outcome in a full-factorial 2 x 2 design. However, given that the 18-month-olds did not show evidence of representing the empty set in Experiment 1, we did not expect that 10-month-olds would represent the empty set in the current experiment.

3.4.1 Methods

3.4.1.1 Participants

Twenty-eight 10-month-old infants (14 males, 14 females) were included in the final sample. Mean age was 10 months, 13 days (range: 10 months, 2 days -– 10 months, 30 days). Four additional infants participated (2 males, 2 females) but were excluded from the sample due to failure in reaching minimum looking times during manipulation (1), pupil size more than two standard deviations larger than mean (1) or technical failure (2). Recruitment and inclusion criteria were identical to Experiment 1. On average, infants provided data for 13.46 trials of out 16 trials (range 5 – 16).

Apparatus, Stimuli, Procedure and Data Processing were identical to Experiment 1. On average, 73.69% of all gaze points (range 70.47% – 76.22%) fell into the central AOI of the door area during the test event, with no significant differences between conditions ($F_{(1,27)}$ = 1.337, p = .285, n_p^2 = .138).

3.4.2 Results and Discussion

Figure 3.5 displays the mean pupil diameters for the four conditions during baseline and test. Visual inspection of the time line of pupil dilation revealed no apparent differences between conditions during the baseline phase. Paired t-tests confirmed that pupil dilation was not different during the two types of manipulation ($M_{\text{“occlusion“}}$ = 4.55 mm, $SD_{\text{“occlusion“}}$ = .54 mm, $M_{\text{“removal“}}$ = 4.56 mm, $SD_{\text{“removal“}}$ = .51 mm, $t_{(27)}$ = -.56, p = .584).

Figure 3.6a further reveals that infants looked at the screen the same amount of time, irrespective of which manipulation they had previously seen ($M_{\text{“occlusion“}}$ = 527.10 ms, $SD_{\text{“occlusion“}}$ = 173.32 ms, $M_{\text{“removal“}}$ = 543.98 ms, $SD_{\text{“removal“}}$ = 195.20 ms, $t_{(27)}$ = -.65, p = .522). Thus, infants were paying the same amount of attention towards the screen during baseline, with no significant difference in pupil size, assuring that the calculation

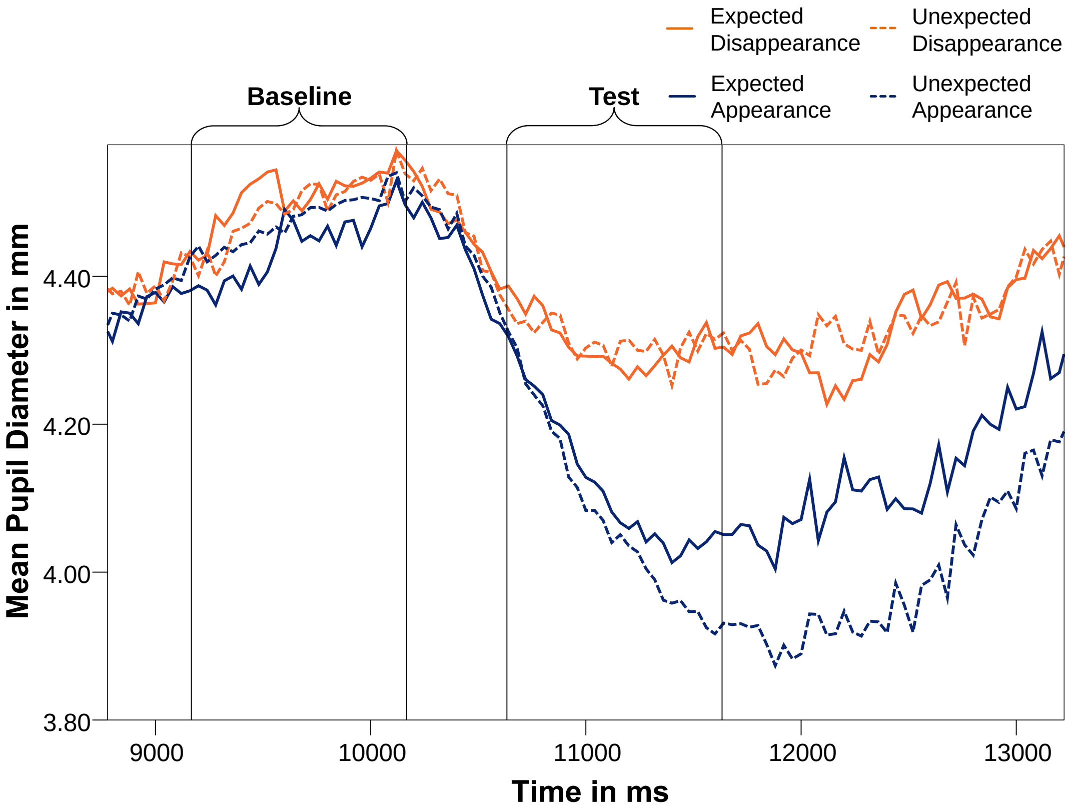
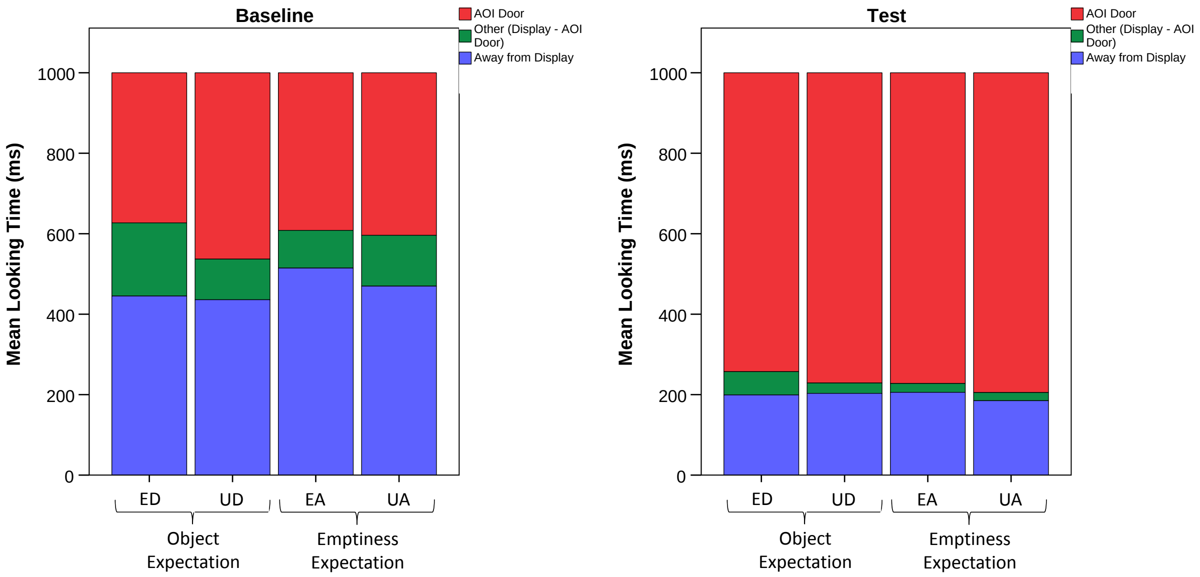


Figure 3.5: Mean pupil diameter of 10-month-olds in Experiment 2. Windows are shifted by 500 ms compared to the timeline in Fig. 3.1.

of the relative change of pupil size to the test outcomes was unbiased by differences in the baseline.

Figure 3.7 shows the results of the relative change score analysis at test, averaged over the duration of the first 1000 ms. A 2 (outcome) x 2 (expectation) repeated-measures ANOVA revealed a main effect for outcome ($F_{(1,\,27)} = 59.31$, $p < .001$, $n_p^2 = .687$) such that infants' pupils dilated more to the empty outcome compared to the toy outcome and no main effect for expectation. Unlike for the 18-month-olds, the interaction term did not reach significance ($F_{(1,\,27)} = 1.92$, $p = .177$, $n_p^2 = .067$).

The direct comparison following our hypothesis further confirmed that infants' pupils indeed did not dilate differently between the expected and unexpected empty outcomes ($\text{M}_{\text{“ED“}} = \text{-}.0385$, $\text{SD}_{\text{“ED“}} = .033$, $\text{M}_{\text{“UD“}} = \text{-}.039$, $\text{SD}_{\text{“UD“}} = .036$, $t_{(27)} = .06$, $p = .949$,



(a) Baseline phase. *(b)* Test phase.

Figure 3.6: Distribution of looking time during baseline and test of 10-month-olds in Experiment 2.

two-tailed). Although we did not expect differences between the toy outcomes, and the interaction term did not reach significance, we also report the comparison between the expected and unexpected object appearance for completeness. Counter to the empty set hypothesis, and to our expectation of no differences, the expected appearance event yielded a relatively larger increase in pupil size than the unexpected appearance event ($M_{\text{“EA“}}$ = -.083, $SD_{\text{“EA“}}$ = .043, $M_{\text{“UA“}}$ = -.099, $SD_{\text{“UA“}}$ = .062, $t_{(27)} = 2.18$, $p = .038$). While it is unclear how to interpret the effect, we must note that in the absence of a significant interaction term and any hypothesis or directed prediction, the finding might be spurious and would not survive an alpha-level correction appropriate for the unplanned post hoc comparison.

Figure 3.6b displays results of our control analyses of the looking times during the test phase. Infants looked at the videos for the same amount of time during expected and unexpected empty outcomes ($M_{\text{“ED“}}$ = 704.66 ms, $SD_{\text{“ED“}}$ = 310.20 ms, $M_{\text{“UD“}}$ = 736.66

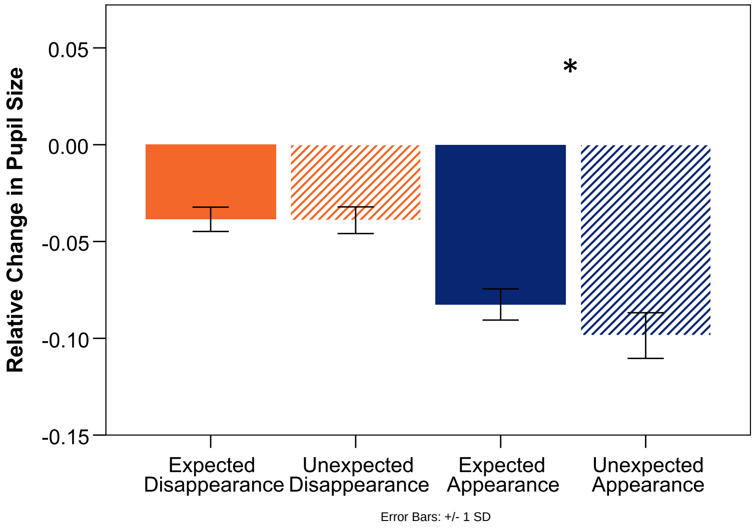


Figure 3.7: Change in pupil size of 10-month-olds in Experiment 2 relative to baseline. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

ms, SD$_{\text{«UD«}}$ = 266.01 ms, $t_{(27)} = .68$, $p = .503$), and during expected and unexpected toy outcomes (M$_{\text{«EA«}}$ = 743.95 ms, SD$_{\text{«EA«}}$ = 279.53 ms, M$_{\text{«UA«}}$ = 762.15 ms, SD$_{\text{«UA«}}$ = 259.06 ms, $t_{(27)}$ = -.758, $p = .455$). Similarly, there were no differences with regard to looks towards the door and looks away from the screen when comparing the expected and unexpected outcomes ($p > .455$). These results thus reveal no additional factor in our data that could account for the pupillometric findings. Notably, our measurement was sensitive enough to detect significant differences in infants' pupil sizes, as revealed by the main effect for the outcome, excluding the possibility that the pupillometric measure did not work at all.

Current results thus suggest that 10-month-olds do not form object expectations following occlusion, which is in line with the pupillometric and looking time findings by Sirois and Jackson (2012). The absence of the effect in our current paradigm is especially meaningful given the presence of the effect at 18 months of age. Comparing object expectations between experiments directly revealed no main effect of age ($F_{(1,\ 46)} = .25$, $p = .617$, $n_p^2 =$ .006), a relatively reduced effect of object expectation compared to at 18 months ($F_{(1,\ 46)} =$

3.10, $p = .043$, one-tailed, $n_p^2 = .064$) and an interaction with age in the predicted direction ($F_{(1,\ 46)} = 3.44$, $p = .035$, one-tailed, $n_p^2 = .071$), supporting the conclusion of our planned experimental comparisons that the effect of object expectations was driven entirely by the 18-month-olds, and absent at 10 months of age.

3.5 General Discussion

The current study employed a new VOE paradigm using pupillometry as a measure of object representation during occlusion to add to the debate about the validity of previous findings of object permanence. The paradigm revealed object permanence in 18-month-olds, an expected competence at that age given that infants successfully search for and communicate about perceptually absent objects (Liszkowski, 2006; Moll & Tomasello, 2004). These findings thus reveal that the paradigm is sensitive to assessing violations of object expectations in infants. Regarding the more contested age of 10 months, however, the paradigm revealed no evidence for expectations of objects during occlusion. What can we make of 10-month-olds' failure in light of 18-month-olds' passing in the current object permanence paradigm?

On the one hand, our findings of 10-month-olds' failure of object expectations is in line with previous research. Most notably, our findings concur with the pupillary findings by Sirois and Jackson (2012) and provide confirmatory support for the absence of object expectations at 10 months of age. The current study substantiates this absence of evidence by presenting positive evidence with the same paradigm for an older age group. The absence of evidence is further supported by leaner interpretations of looking time studies which suggest that results can more parsimoniously be explained by perceptual preferences and biases (Schilling, 2000; see also Baillargeon, 1999). Further doubts on the robustness of object expectations derive from infants' search for visibly displaced objects. For example, (Topál et al., 2008) removed the social and communicative pull that might induce incorrect searches at an initial hiding place and found that 10-month-old infants still searched randomly for a visibly displaced object, indicating that infants' search is not guided by strong expectations about the whereabouts of visibly hidden objects.

On the other hand, the lack of evidence at 10 months appears to be at odds with other research. For example, two studies using oscillatory gamma activity as a signature for object maintenance during occlusion found positive evidence at 6 months of age (Kaufman et al., 2003, 2005). Gamma oscillations typically increase immediately in response to occlusion of objects and the increase is short lived (Tallon-Baudry et al., 1998), suggesting that infants maintain a trace of the object at the moment it becomes occluded. It is less clear whether this measure also reveals a longer lasting cognitive expectation that the object continues to exist. Reaching studies reveal that 6-month-olds can reach for objects in the dark (Clifton et al., 1993; Hood & Willatts, 1986) or in opaque liquids (Shinskey, 2012), suggesting that their immediate reaching is guided by representations of invisible objects. The reaching is less proficient when compared to their reaching for occluded entities (Munakata, 1997), and reaching in the dark appears still different from that of adults (Babinsky et al., 2012). Further, infants do not perform above chance in simplified visible displacement task (Topál et al., 2008). While these findings may suggest that occlusion is hard (Hespos et al., 2009), they must also mean that young infants' expectations about hidden objects are initially less robust. An anticipatory looking study (Ruffman et al., 2005), however, in which the placing of an object was ostensively communicated to 4-month-olds before the object was occluded in one of two locations, revealed that infants expected the object to appear where they last saw it (after a 2-second-delay, not after an 8-second-delay). The cue to anticipate was the experimenter's utterance ('Doors up, here comes the hand') and in familiarization trials infants had been shown a hand waving the object while the experimenter called the infant's name. This may suggest that infants were expecting not just an object but an action on the object where they had last seen it.

One difference between the aforementioned paradigms and the current paradigm is that these did not involve violations of expectations. In the paradigms employing reaching and anticipatory looking, infants are prepared to execute these behaviors in familiarization trials, to ensure that infants will process the scenario with regard to acting in it (and then act). This is not the case in looking time VoE paradigms, but as reviewed before looking time VoE paradigms may have their own weakness in interpretation. In the current VoE paradigm, there were no familiarization trials, and the neurophysiological measure required no acting. One possibility why 10-month-olds' did not show the effect is then that they simply did not process the relevance of the scene in terms of the objects' permanence and violations thereof. This proposal entails that task-inherent demands masked an existing competence in 10-month-olds. Any explanation for 10-month-olds' failure must of course be seen in light of 18-month-olds' competent performance in the task.

The interpretation we thus offer includes developmental change in infants' stable representation of occluded objects, a change from cued (or scaffolded) to spontaneous (or internalized) use of the ability. In line with our contention that our paradigm did not instigate the relevance of processing object permanence in the 10-month-olds, those studies providing positive evidence with younger infants – all looking time measures quarantined from this argument – all include rich socially embedded or mediated cues to object processing. Typically, infants are encouraged to reach for objects and the placing of the objects is done by an interactive experimenter. Also the anticipatory looking study (Ruffman et al., 2005) heavily relied on social cues to attending to the object, including rich cues of acting on the object. Recent studies show that social aspects of the task influence infants' performance in tasks of object expectations, from creating illusions of object locations and features (Topál et al., 2008) to resolving ambiguity through social referencing (Dunn &

Bremner, 2017). The lack of any social and action-relevant context as in our paradigm then likely hampered correct performance at a younger age. The finding that the older age group was not affected must then suggest that the development of spontaneous, seemingly automated mental representations of occluded objects is mediated by social relevance.

4 Gamma-oscillations as a signature of object representations following occlusion and pointing events[1]

RESEARCH HIGHLIGHTS:

- *12-month-old infants showed right-frontal gamma activity in response to object occlusion, but not object disappearance.*
- *10-month-olds showed right-temporal gamma activity in reponse to the occlusion of an object, but only when it occurred in a communicative context.*
- *During a social cue, both age groups exhibited bilateral temporal gamma activity.*
- *In the communicative, but not the noncommunicative social condition, we also found left frontal gamma activity associated with joint attention.*

In the previous two chapters, I used pupillometric measures to show that infants expect objects to appear after a communicative act (referential expectation, Chapter 2) and after occlusion (object permanence, Chapter 3). In the current chapter, I am using neurophysiological markers of object representation to investigate whether infants' representation of an object may be elicited by a communicative cue, without a subsequent reveal of an object. The paradigm therefore is based upon the online maintenance of an object representation, not the violation of expectation. Kaufman et al. (2003, 2005) first introduced increased activity in the gamma frequency band as a signature for object maintenance in infants (see also Kampis et al., 2015; S. Leung et al., 2016). Therefore, in Experiment 1, I first established this object maintenance effect in 12-month-olds and explored online processing

---

[1] This study is in preparation for submission entitled "Gamma-oscillations as a signature of object representations following occlusion and pointing events in 10- and 12-month-old infants" (Pätzold & Liszkowski).

of a pointing gesture or a nonsocial, attentional cue following the occlusion event. I report the occlusion maintenance effect in right-frontal regions, and find an increase in gamma activity during the social, but not the nonsocial attentional cue.

In Experiment 2, I investigated the object maintenance effect in the context of a communicative cue, or the absence of such a cue, in 10-month-olds. The results suggest that infants actively maintain object representations after occlusion events in a referential-communicative context, but not in a social, but non-communicative context. Furthermore, I report that processing of the communicative cue elicits widespread frontal and temporal gamma activation, whereas the matched non-referential cue only elicits temporal activation, but no frontal activation. While the results pertain to a small sample, the indications of these findings are discussed in the context of a readiness for referential-communicative signals that facilitates object processing in infancy, and the significance of this finding for current developmental theories.

Acknowledgements: We thank all families who participated in these studies. We thank Kirstin Burckhardt for her help in running a pilot study. We are grateful to Jula Brüning-Wessels, Maria Häberlein, Sriranjani Karthik, Jana Klose, Zoé Kröckel, Lydia Rihm, Mieke Röder, Eva Rüdinger and Betty Timmermann for their assistance with data collection. We especially thank Susanna Jeschonek-Seidel for valuable feedback in the early stages of the project and Sriranjani Karthik for conceptual and practical support throughout data collection and analysis.

4.1 Abstract

Increased activity in the lower gamma-band (20-40 Hz) has been correlated with object representation in infants. In the current study, we investigated how infants process the act of pointing towards an object that has been occluded as referential to that object.

In Experiment 1, 12-month-old infants watched short videos in which a toy was either occluded or disappeared. In half of the videos, an agent then pointed at the occluder. In the other half, the occluder was illuminated. The results showed increased gamma activity in right frontal regions in response to the occlusion event, but not to the disappearance, confirming gamma as a signature for object representation. During the pointing gesture, we noted bilateral temporal gamma activity.

In Experiment 2, 10-month-olds watched videos of an agent looking at a container, which then opened up to reveal a toy inside. After the container closed again, occluding the toy, the actor lifted her hand to point at the container, or, in the control condition, lifted her hand to cup her mouth while coughing. Interestingly, we found a significant increase in gamma activity after the toy was occluded in right-temporal sites only in the pointing condition, but not in the control condition. Gamma activity also increased in both frontal and temporal regions during the pointing gesture, whereas the control condition showed temporal, but no frontal effects.

The results shed a new light on the development of object cognition in relation to referential-communicative cues at the end of the first year.

Keywords: Gamma oscillation, Object maintenance, Infant EEG, Social cognition

4.2 Introduction

We know that infants comprehend the referential intention of a point gesture around their first birthday (Tomasello et al., 2007). Behavioral studies have shown that infants around 12 months of age readily search for a toy when adults direct their attention towards a covered container or occluded area (Behne et al., 2012; Moll & Tomasello, 2004). A recent study investigating object expectation using pupillometry reveals that infants expect an object following a point towards an occluded area at 12 months, but not at 8 months (Pätzold & Liszkowski, 2018). But what is it exactly that they represent when they understand the communicative intent of the point? At this time, we do not know whether it is actually the expectation, and therefore representation, of an object, that infants form, or if their expectation is a broader form of anticipation or excitement for something interesting soon to happen. Moreover, we do not know whether infants represent the referent they have already seen while they are searching for it (Behne et al., 2012). We therefore need a way to shed light on the nature of the representations that infants have during communicative acts.

We approached this question from what we know about object representation in infancy. Remedying the biases of the traditional habituation tasks (Baillargeon, 1987; Baillargeon et al., 1985), S.-H. Wang et al. (2004) showed that as early as 4 months of age, infants looked longer to events that were inconsistent with the concept of a full occlusion or containment. In this study, the authors removed the habituation phase in order to avoid transient novelty or familiarity preferences that may offer leaner explanations for earlier findings of object representation (Bogartz et al., 1997; Schilling, 2000). Alternatives to looking time measures, including pupillometry (Jackson & Sirois, 2009) and anticipatory looking (Ruffman et al., 2005) have recently been used to provide evidence of object permanence during occlusion in the first year, but do not offer an online measure of object representation.

Neurophysiological measures present a much more unambiguous way to investigate infants' minds, because they do not depend on behavioral output (Hoehl & Wahl, 2012). Recently, the decomposition of the overall power in the EEG signal into individual frequency bands has allowed researchers to investigate the increase in activity within different ranges of frequencies associated with specific functions (Saby & Marshall, 2012).

When investigating object representations, the gamma band (30 – 100 Hz in adults; 20 – 60 Hz in infants, Saby & Marshall, 2012) has proven particularly useful. In adults, gamma band activity is associated with perceptual binding (Müller et al., 1996) and active maintenance of objects in memory (Bertrand & Tallon-Baudry, 2000; Tallon-Baudry et al., 1998). In a series of EEG studies investigating object occlusion in infants, six-month-olds showed an increase in gamma activity in temporal regions after the object's occlusion (Kaufman et al., 2003, 2005) but not after its disintegration (Kaufman et al., 2005), signifying a continued maintenance of the object representation. The authors explained the findings in terms of consistency with continued existence: An occluded object is still there and may again become relevant, whereas a dissolved object is gone. We will refer to these findings as the "object maintenance effect". Recently, it has been demonstrated that gamma band activity may increase in infants when another person represents an occluded object, even when it is still visible to the infant (Kampis et al., 2015), and that gamma band activity increases with the number of objects represented (S. Leung et al., 2016), emphasizing the utility of the gamma band power analysis in the study of object representation. These findings led us our main question: If infants understand communicative cues to be referential, do they show increase in gamma band activity when an occluded object is referred to by a social partner?

For this, we took a two-step approach: First, we established the object maintenance effect as described above in 12-month-olds, the age when infants readily pick up referential cues in behavioral settings. Second, we investigated whether a referential communicative cue following the occlusion would generate an activation similar to the short term memory effect by Tallon-Baudry et al. (1998) in terms of timing, location and frequency band, which would indicate the repeated activation of a memory trace associated with the object.

In Experiment 1, we included the occlusion/disappearance comparison to establish the object maintenance effect in 12-month-olds. Our hypothesis was that infants would show an increase in gamma activity after the object was completely out of sight only in the occlusion condition, but not in the disappearance condition. Following this occlusion or disappearance event, we then highlighted the occluded area with an attentional cue: either a referential-communicative pointing gesture (social cue) or a comparable attentional cue of sound and light (nonsocial cue). Our hypothesis was that infants would again show an increase in gamma activity only in the communicative cue condition, but not in the attentional cue condition, akin to the second burst found in adults when they had to represent a previous stimulus in preparation for a response (Tallon-Baudry et al., 1998).

In Experiment 2, we regarded the object maintenance effect as established and abolished the control condition. After the occlusion of the toy, 10-month-olds saw a social cue. In this experiment, the communicative pointing gesture was contrasted with a gesture that was visually matched to the pointing gesture but did not entail a communicative cue (coughing). Therefore, in contrast to Experiment 1, both conditions contained a social aspect. We hypothesized that would infants show an increase in gamma activity only in the referential communicative cue condition, but not in the non-referential, non-communicative cue condition,

We defined regions of interest (ROI) a priori based on previous literature. For the object maintenance effect, we expected activity in temporal or parietal regions based on findings investigating sustained object representations in infancy (Kampis et al., 2015; Kaufman et al., 2003, 2005; S. Leung et al., 2016). While Kaufman et al. (2003, 2005) and S. Leung et al. (2016) find activation restricted to the right hemisphere, Kampis et al. (2015) find activation in both hemispheres, with a more pronounced effect on the left side. Therefore, we included both hemispheres in our analyses.

Additionally, previous infant studies using gamma frequency band analysis have looked at prefrontal regions (Csibra, Davis, et al., 2000; Grossmann et al., 2007). Investigating object processing, Csibra, Davis, et al. (2000) found an enhancement of induced gamma-band activity in response to the Kanizsa square over three left frontal electrodes (corresponding to an area between F3-F7-Fp1). Grossmann et al. (2007) looked at gamma activation in response to direct and averted gaze. Their results indicated that a picture for an upright face elicited a significant gamma burst over right prefrontal channels (corresponding to an area between Fp2 and F4) compared to a picture of an inverted face. There are also adult fMRI studies (Kampe, Frith, & Frith, 2003; Schilbach et al., 2006) indicating that prefrontal brain structures are activated by the detection of direct gaze, or possibly communicative intent. We therefore hypothesize that activity pertaining to the processing of the communicative intent of the scene may be found in prefrontal regions.

4.3 Experiment 1

4.3.1 Methods

4.3.1.1 Participants

Twenty-one 12-month-old infants (ten girls) were part of the final sample. Mean age was 12 months, 16 days, (range: 12 months, 4 days – 12 months, 29 days). Only trials in which infants watched the critical manipulation sequence were included in the analysis (frame by frame coding of participant video).

An additional 41 infants (20 girls) participated in this study but did not contribute data due to fussiness (21 infants), too much noise in the signal (16), or failure to wear the cap (4 infants). The excluded infants did not differ from the included infants in age (mean age: 12 months, 16 days, range: 12 months, 3 days – 12 months, 31 days), which is an indication that the results were not biased by a developmental advance of the included sample. Note that only a subset of the included sample provided data for the social/nonsocial comparison (16 infants, seven girls). Although the attrition rate appears high (66.13 %), an attrition rate between 50% and 75% is common in infant EEG studies, especially when using paradigms that rely on attention to visual stimuli (DeBoer et al., 2007; Stets et al., 2012).

4.3.1.2 Apparatus

The BrainVision actiCHamp system including actiCAP active electrodes and the BrainVision recording and analysis software (Brain Products, Munich, Germany) was used for this study. The stimuli were presented using Presentation© software (Neurobehavioral Systems Inc., Berkeley, CA). The recording room was shielded from electrical fields by means of a conducting mesh in order to ensure an EEG recording free of electromagnetic interferences. The presentation and recording computers, the presentation screen, presentation speakers and a camera monitoring the infants' behavior simultaneously to the EEG-recordings were installed in an adjoining room, accessing the recording room through a window covered by a clear shielding foil (marTECH Systems).The presentation screen was a 1920 px x 1200 px LCD computer screen (51.50 x 32.00 cm). From the inside of the experimental chamber, only screen itself was visible through the window, whereas the speakers and infant camera where concealed behind a cloth frame.

4.3.1.3 Procedure

Before beginning preparation, the experimenter explained the study to the caregiver and obtained their informed consent. The experimenter worked together with a research assistant to prepare the infant for the EEG recording. After the EEG cap had been fastened on the infant's head, one person applied electrically conductive gel (preheated to 37°C) and adjusted the impedances of the active electrodes while the other one played with and entertained the infant. Testing was generally started when impedances were below 20 kΩ.

Infants were seated on their caregiver's lap on a chair facing the presentation screen. The infants' eyes were approximately 80 cm from the screen. The lights were dimmed and the video presentation started with an attention getter to orient the infant towards the screen. The session lasted as long at the infant remained still and was stopped when the infant exhibited crying, fidgeting or excessive movement, or until a total of 64 experimental videos had been presented. Low-density EEG was recorded at 32 channels (20-10 system) at 250 Hz and synchronized video of the infant was recorded at 24 fps.

After the study, the experimenter debriefed the caregiver and the infant was given a small toy and a certificate for participating.

4.3.1.4 Stimuli

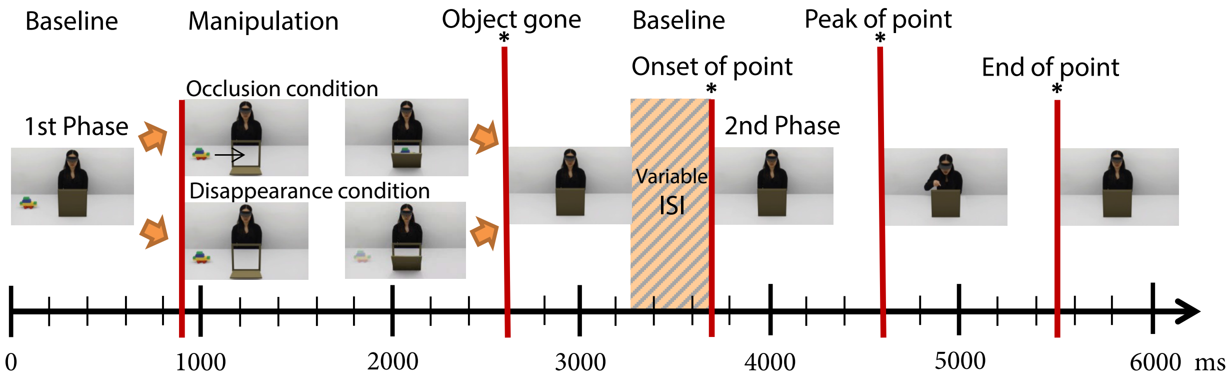


Figure 4.1: Scheme of the experimental design for Experiment 1.

Figure 4.1 shows the sequence of each trial schematically. In the social manipulation videos, each video started with an agent sitting at a table, with the face partially covered with a visor. In front of her in the middle of the table was a closed door frame and to the left or right side, there was a toy vehicle. Sides were counterbalanced during presentation. The video started with the door frame opening to fully reveal the empty table behind it (800 ms). 2000 ms later, the vehicle started moving. In the occlusion condition, the vehicle moved forward to the middle of the frame (1333 ms), which immediately started closing, completely occluding the vehicle 500 ms later. In the disappearance condition, the vehicle moved forward and then backward to its initial position. As the door closed (synchronized to the time line of the occlusion event), the vehicle disappeared from the scene. At this point, a variable time (400 ms, 600 ms or 800 ms) passed before the second phase of the scene. The agent raised her right hand, formed an index finger point towards the door and exclaimed "Ah!", with the peak of the point being reached after 800 ms. She then retracted her hand and returned to her original position (800 ms). The video went on for another 400 ms in order give time to record lingering effects on the oscillatory activity. A white fixation cross on a grey screen was presented for 400 ms between trials.

The nonsocial videos were identical to the social videos, except that there was no agent present. Instead of the social cue, the door frame was highlighted by an intensifying illumination, reaching its peak luminance at the same time as the peak of point, and then dimming back to the original light. A bell sound was played in place of the vocalization in the nonsocial version, time-locked to the voice of the agent of the social videos.

There were eight types of vehicles (car, bus, sailboat, plane, truck, helicopter, train engine and ship) in four color variants each, built from Lego Duplo® blocks to have the same size. The door frame was designed and animated in Blender® in eight different colors. Objects and door frames were super-imposed on the scene using Adobe After Effects®.

Videos were rendered to the size 1280 x 800 px. Presentation size of the video was 34.5 by

21.5 cm, which corresponds to a visual angle of 24.3° horizontally and 15.3° vertically. The order of presentation of the four conditions (occlusion/disappearance, social/nonsocial), the length of trial (three variable lengths of ISI between the first and the second part of the trial), the type of vehicle and the door color were pseudo-randomized with the constraints of not appearing more than twice in a row. Because of the constraints, four fixed stimuli orders were prepared and implemented in Presentation© scripts.

The four conditions were tested within-subjects. Infants were randomly assigned to a fixed stimuli order when they came in. All videos were uploaded for presentation before the start of the session, to minimize buffering times. Additionally, an at-will attention-getter (an animation of three red balloons, dancing in the middle of the screen), as well as a stop function, were implemented so that infants could be re-oriented towards the screen if they were fussy. When the attention-getter or the stop button was pressed, the current trial was aborted. On resuming the presentation, the next trial in line automatically started.

Triggers for syncing the stimulus presentation with the EEG data were written into the script to co-occur with key events in the videos. Each experimental trial contained two triggers: The first one at 2600 ms, the moment of complete occlusion or disappearance of the toy, and the second one at 4200 ms (4400; 4600 ms for the longer variants), the peak of the cue. In addition to these conceptual triggers, trigger codes were automatically sent at the start of each video, identifying the type of video (occlusion, disappearance, social, nonsocial) for later recognition in the EEG software.

4.3.1.5 Data Processing

Data were imported into the Analyzer2 analysis software (Brain Products, Munich, Germany). First, data were low-pass-filtered at .3 Hz. A rater then coded the subject video frame by frame into categories “watched” (when the infant’s eyes were clearly focused on the video) and “not watched” (when the infant’s eyes were averted, not on camera or focus was unclear).

Next, data were segmented for the occlusion or disappearance event of each trial from -800 to 800 ms, with 0 ms being the time of complete occlusion or disappearance of the object. For the second phase (cued events), data were segmented at the time window around each cue starting at -1200 to 1000 ms, with 0 ms being the peak of the cue. During both segmentations, all data that were not marked “watched” were discarded. This was done to make sure that only EEG data of events that were attended to were used in the analysis process.

Artifact rejection was done semi-automatically, with maximal difference of values in intervals set to 200 µV (interval length: 100 ms). Segments that were affected by artifacts were discarded, or channels marked for interpolation. Following the first artifact rejection, channels were individually interpolated by linear interpolation if needed. Not more than 3 channels were interpolated per data set. Channels TP9 and TP10 were discarded in all data sets because they tended to be noisy and were not relevant to the question of the study.

Next, data were re-referenced from Cz to Average. After re-referencing, Morlet wavelets were calculated for each segment (Morlet Complex 20-60 Hz, 41 Frequency Steps, Morlet Parameter c: 7, Logarithmic steps). A 200 ms time period before the door closed, after the toy had stopped moving, was defined as a baseline for occlusion and disappearance events (-400 ms to -200 ms in relation to time of occlusion/disappearance). For cue events, a 200 ms time period before the person started moving was defined as baseline (-1000 ms to -800 ms in relation to peak of cue). Total Gamma was calculated by averaging trials for each infant.

Following this step, data from individual channels were pooled into different combinations of two or more channels. This was done in order to a) produce results better comparable to studies were high-density EEG was used and b) better define the ROI for each event. We defined ROI based on previous findings in temporal (Kampis et al., 2015; Kaufman et al., 2003) and frontal regions (Csibra, Davis, et al., 2000; Grossmann et al., 2007). Temporal poolings were CP5 and T7 (left) and CP6 and T8 (right). Frontal poolings were F3, F7 and FC5 for the left hemisphere and F4, F8 and FC6 for the right hemisphere.

For the occlusion/disappearance event, time bins of 200 ms lengths were then averaged and exported from Analyzer2 for statistical analysis, comprising the time from complete occlusion or disappearance to 600 ms post event. Because changes in gamma activity could be short-lived, we chose overlapping bins, so that five time bins in total were exported. Similarly, data for the social/nonsocial cue event were exported in 200 ms time bins, comprising the time -200 to 600 ms. A total of seven time bins were exported.

4.3.2 Results

In segmenting the data for artifact rejection and further analysis, we deliberately separated the occlusion/disappearance event from the social/nonsocial cue in each trial. This was done because artifacts impairing one part of the trial jeopardized the analysis of the part of the trial with good data, leading to the exclusion of a large number of trials. Therefore, results for the occlusion/disappearance event and the social/nonsocial cue event are reported separately, and have to be treated as between-subjects results.

4.3.2.1 Results of the Occlusion/Disappearance Events

Because the occlusion/disappearance event preceded the cue event in each video, data from the social/nonsocial cue conditions were merged for analysis. Data sets were included in the final analysis if they provided artifact-free data for at least 8 experimental trials. All 21 infants (ten girls) were included for this comparison. On average, infants provided 14 trials (range: 8 – 22) in the occlusion condition and 15 trials (range: 9 – 24) in the disappearance condition.

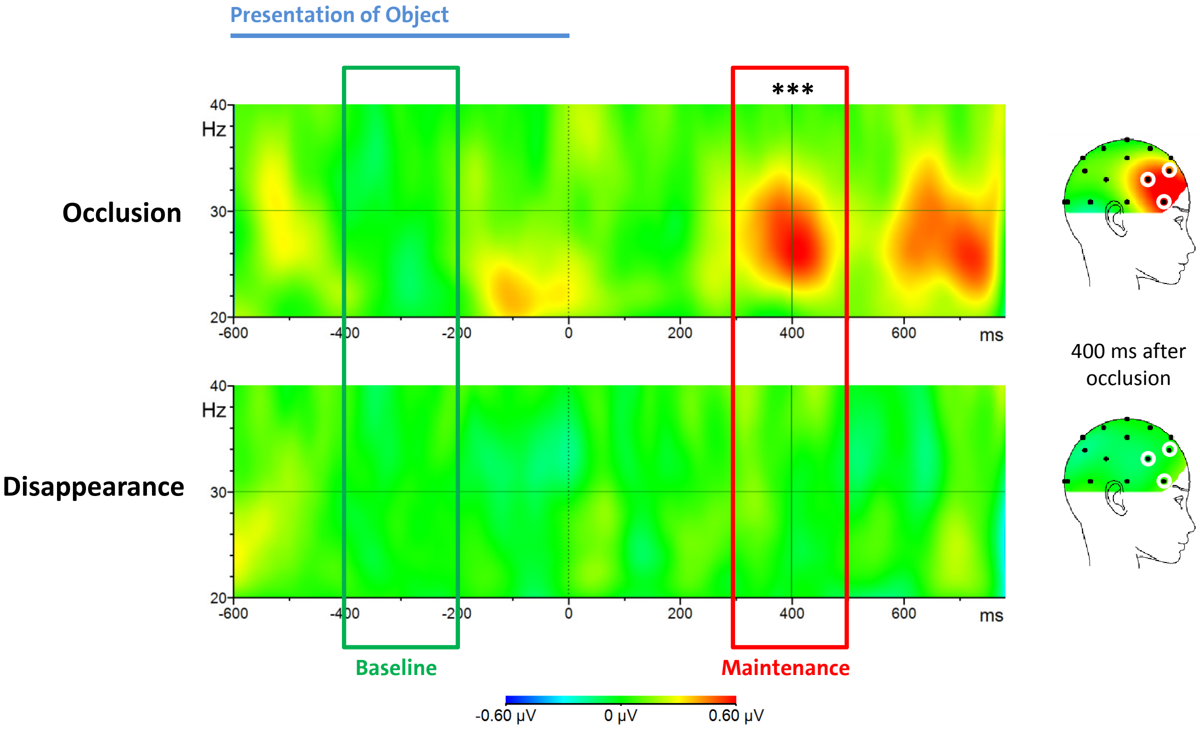


Figure 4.2: Time-frequency plots (left) and topographic maps (right) of activity in the gamma band at right frontal sites after the occlusion event (top) or the disappearance event (bottom) in 12-month-olds in Experiment 1. F4, F8 and FC6 pooled for analysis. Baseline from -400 to -200 ms. Object is completely occluded at 0 ms. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 4.2 shows an increase in gamma activity in right frontal areas (F4, F8 and FC6) at 300 ms – 500 ms post-occlusion in the frequency band 20 – 40 Hz in the occlusion condition (M“baseline”= .002 µV, SD“baseline”= .006 µV, M“test”= .306 µV, SD“test”= .364 µV) but not in the disappearance condition (M“baseline”= .0005 µV, SD“baseline”= .006 µV, M“test”= .037 µV, SD“test”= .311 µV). A 2 (time) x 2 (condition) repeated measures ANOVA showed a significant interaction between time and condition ($F_{(1,20)} = 4.92$, $p = .038$ $n_p^2 = .197$), a main effect for time ($F_{(1,20)} = 17.18$, $p = .001$, $n_p^2 = .462$) and a main effect for condition ($F_{(1,20)} = 4.82$, $p = .040$ $n_p^2 = .194$). Paired comparisons following our hypothesis confirmed that activity in the occlusion condition was significantly higher compared to baseline ($t_{(20)} = 3.87$, $p = .001$) and compared to the disappearance condition ($t_{(20)} = 2.21$, $p = .039$).

The disappearance condition was not different from baseline ($t_{(20)} = .54$, $p = .592$).

For the left frontal pooling (occlusion condition: M$_{\text{“baseline”}}$= -.003 µV, SD$_{\text{“baseline”}}$= .007 µV, M$_{\text{“test”}}$= .152 µV, SD$_{\text{“test”}}$= .369 µV; disappearance condition: M$_{\text{“baseline”}}$= .002 µV, SD$_{\text{“baseline”}}$= .009 µV, M$_{\text{“test”}}$= .088 µV, SD$_{\text{“test”}}$= .308 µV), a 2 x2 ANOVA showed a main effect for time ($F_{(1,20)} = 6.69$, $p = .018$, $n_p^2 = 251$), no main effect for condition ($F_{(1,20)} = .26$, $p = .615$, $n_p^2 = .013$) and no interaction ($F_{(1,20)} = .35$, $p = .562$, $n_p^2 = .017$). Paired comparisons confirmed that gamma activity was significantly increased from baseline in the occlusion condition ($t_{(20)} = 1.96$, $p = .035$, one-tailed), but not in the disappearance condition ($t_{(20)} = 1.31$, $p = .206$).

For the occlusion event, there were no changes in activity in temporal or parietal regions.

4.3.2.2 Results of the Social/Nonsocial Cue

As the social/nonsocial cue event took place later in the trial, fewer infants provided enough data for analysis for this part of the experiment. Therefore, data from occlusion/disappearance events had to be merged for this analysis. 16 infants (seven girls) provided enough data for both the social cue and the nonsocial cue conditions. Data sets were included in the final analysis if they provided artifact-free data for at least 5 experimental trials. On average, infants provided 12 trials (range: 5 – 20) in the social condition and 10 trials (range: 5 – 14) in the nonsocial condition.

Figure 4.3 shows the pattern of gamma activity 400 – 800 ms after the onset of the cue at the left temporal pooling CP5 and T7 (social cue condition: M$_{\text{“baseline”}}$= .0005 µV, SD$_{\text{“baseline”}}$= .006 µV, M $_{\text{“test”}}$= .133 µV, SD$_{\text{“test”}}$= 0.286 µV,), nonsocial cue condition: M$_{\text{“baseline”}}$= -.003 µV, SD$_{\text{“baseline”}}$= .007 µV, M $_{\text{“test”}}$= .086 µV, SD$_{\text{“test”}}$= 0.329 µV). A 2 (time) x 2 (condition) repeated measures ANOVA revealed a main effect for time ($F_{(1,15)} = 5.21$, $p = .037$, $n_p^2 = .258$), no main effect for condition ($F_{(1,15)} = .18$, $p = .675$, $n_p^2 = .012$) and no interaction ($F_{(1,15)} = .13$, $p = .721$, $n_p^2 = .009$). Based on the main effect of time,

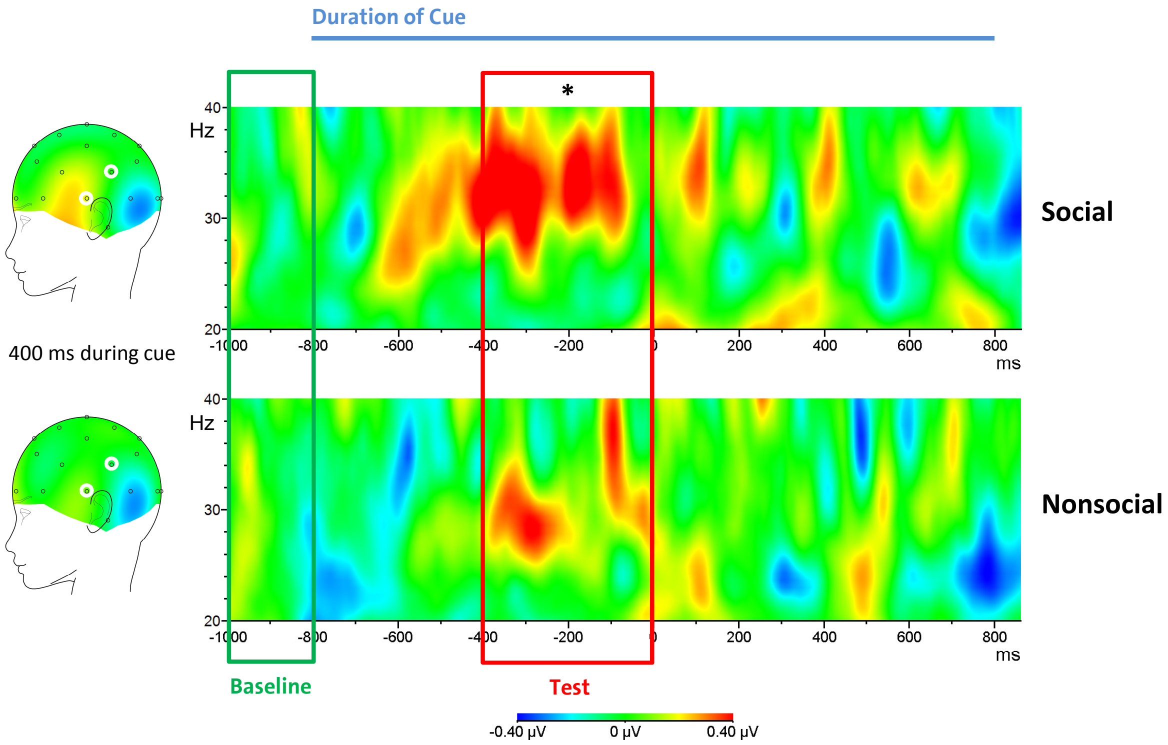


Figure 4.3: Topographic maps (left) and time-frequency plots (right) of activity in the gamma band during the social cue condition (top) and the nonsocial cue condition (bottom) at left-temporal electrode sites CP5 and T7 in 12-month-olds in Experiment 1. Baseline from -1000 to -800 ms. Cue starts at -800 ms and ends at 800 ms. Peak of point at 0 ms. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

we performed paired comparisons between baseline and test within conditions and found that the pointing gesture lead to an increase in gamma activity compared to baseline ($t_{(15)} = 1.87$, $p = .041$, one-tailed), whereas the nonsocial cue did not ($t_{(15)} = 1.09$, $p = .293$).

Figure 4.4 shows the pattern of gamma activity 400 – 800 ms after the onset of the cue at the right temporal pooling CP6 and T8 (social cue condition: M$_{\text{“baseline”}}$= .001 µV, SD$_{\text{“baseline”}}$= .004 µV, M $_{\text{“test”}}$= .131 µV, SD$_{\text{“test”}}$= 0.342 µV) compared to the nonsocial cue condition (M$_{\text{“baseline”}}$= -.00003 µV, SD$_{\text{“baseline”}}$= .009 µV, M $_{\text{“test”}}$= -.146 µV, SD$_{\text{“test”}}$= 0.329 µV). A 2 x 2 ANOVA revealed no main effect for time ($F_{(1,15)} = .02$, $p = .890$, $n_p^2 = .001$), a significant main effect for condition ($F_{(1,15)} = 4.58$, $p = .049$, $n_p^2 = .234$) and a significant interaction between time and condition ($F_{(1,15)} = 4.72$, $p = .046$, $n_p^2 = .239$).

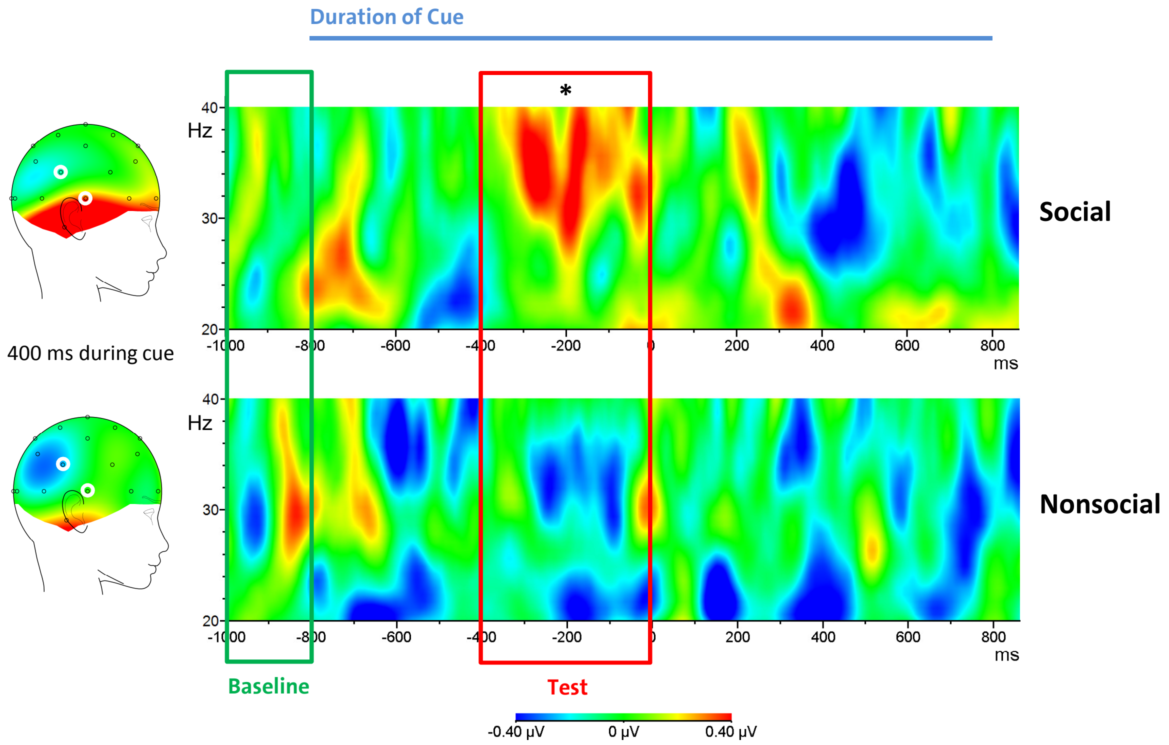


Figure 4.4: Topographic maps (left) and time-frequency plots (right) of activity in the gamma band during the social cue condition (top) and the nonsocial cue condition (bottom) at left-temporal electrode sites CP6 and T8 in 12-month-olds in Experiment 1. Baseline from -1000 to -800 ms. Cue starts at -800 ms and ends at 800 ms. Peak of point at 0 ms. Note. * p < .05, ** p < .01, *** p < .001.

Paired comparisons confirmed that the social cue showed heightened activity compared to the nonsocial cue condition ($t_{(15)}$ = 1.81, p = .046, one-tailed), but not compared to baseline ($t_{(15)}$ = 1.52, p = .149). The nonsocial cue condition showed a significant decrease in activity compared to baseline ($t_{(15)}$ = -2.16, p = .048).

4.3.3 Discussion

We found right frontal gamma activity around 300 – 500 ms after occlusion events, but not after disappearance events. This effect is identical in timing and frequency band to the previously reported object maintenance effect, but localized in right frontal, as opposed to right temporal regions in 6-month-olds (Kaufman et al., 2003, 2005) and in 8-month-olds (Kampis et al., 2015). The differences in location may be explained by advanced cognitive maturation of 12-month-olds compared to younger infants. We report this finding in data in which social and nonsocial conditions were merged, which means it remains to be investigated what role the social context plays for the object maintenance effect.

Experiment 1 therefore offers consolidating evidence for the object maintenance effect reported by Kaufman et al. (2003, 2005) and Kampis et al. (2015) with a control of disappearance. This shows that even „disappearing in thin air“ and not just gradual dissolution leads infants to believe an object to be gone, whereas occlusion, in this case by a solid other object, leads the infant to have a continued representation of the object. In the Experiment 2, we therefore abandoned a control for the occlusion condition and instead focus on the social context of the object maintenance effect further in younger infants.

When exploring the activity during the ongoing cue, there was an increase in gamma activity in temporal regions in both hemispheres co-occurring with the pointing gesture (ending shortly before peak of point), but no increase in activity during the nonsocial, attention-directing cue of light and sound. In this respect, Experiment 1 provides evidence for gamma activity resulting from a referential-communicative cue. This could be evidence that the pointing produces a “second burst” of object representation, as the infant is reminded of the occluded object by the referential gesture. This interpretation is supported by the fact that a mere attention-directing cue does not elicit the same gamma activation. However, in Experiment 1 we had to collapse data from both occlusion and disappearance conditions for the analysis of the social cue, somewhat clouding the interpretation we can draw on these results.

An alternative explanation for this "social effect" is that the richness of the social cue may have elicited gamma increase for a number of reasons, including salience of the facial expression of the agent, her movement, or her voice. Gamma activation right after the onset of the pointing gesture may be therefore be evidence of online processing of a relevant social cue as opposed to pointing comprehension or object representation.

For these reasons, in Experiment 2 we used a control that is perceptually much closer to pointing and is adjusted both for the social content and movement of the agent. Because we also dropped the disappearance condition, all cues were presented in the context of an occlusion event. This enabled us to test the hypothesis whether it is the social content of the stimulus that causes the effect, or the reference to the occluded object created by a communicative gesture. If the gamma activity is elicited by the reference to the object, infants should not show the same pattern in gamma activation with a cue that is neither referential nor communicative. However, if the mere presence of a social partner causes the effect, the change in activity seems to relate more to social processing and less to the referential content of the gesture. In order to avoid carry-over effects from the referential cue to the non-referential control condition, and to ensure we were able to present enough trials to each participant, we tested the two conditions between subjects.

4.4 Experiment 2

4.4.1 Methods

4.4.1.1 Participants

Forty-five 10-month-old infants (26 girls) watched 72 videos of toys being occluded while low-density EEG was recorded on 32 active electrodes at 250 Hz.

After the occlusion, half of the participants saw the actor point at the occluder (pointing condition, n = 23, 15 girls). Mean age was 10 months, 14 days (range: 10 months, 3 days – 10 months, 26 days). The other half saw the actor briefly cover her mouth with her hand while making a coughing sound (control condition, n = 22, 11 girls). Mean age was 10 months, 10 days (range: 10 months, 0 days – 10 months, 30 days). Only trials in which infants watched the critical occlusion event were included in the analysis (frame by frame coding of participant video).

An additional 77 infants (31 girls) participated in this study but were not part of the final sample. 35 of those infants (15 girls) took part in the pointing condition but did not contribute data due to fussiness (17 infants), too much noise in the signal (13), technical failure (4) or failure to wear the cap (1). 42 infants (16 girls) took part in the control condition but did not contribute data due to fussiness (20 infants), too much noise in the signal (19), failure to wear the cap (1), values more than two standard deviations from the mean (1), or experimenter error (1). The excluded infants did not differ from the included infants in age. In the pointing condition, the mean age was 10 months, 14 days (range: 9 months, 25 days – 10 months, 30 days). In the control condition, the mean age was 10 months, 11 days (range: 9 months, 25 days – 10 months, 26 days).

On average, participants included in the final sample of the pointing condition provided data for 12 experimental trials (range 8 – 25), whereas participants in the control condition provided data for 14 experimental trials (range 8 – 38). Apparatus and Procedure were identical to Experiment 1.

4.4.1.2 Stimuli

Each video started with a lateral view of an agent facing a large container on a table (Figure 4.5). Only the head down to the shoulders was visible in the video, making facial features especially salient. After a still period of 1000 ms, the container split open at about a third of its height (200 ms) and revealed an object. The object was visible for 1000 ms before the container closed again (200 ms), completely occluding it.

A variable time (either 1000 ms or 1400 ms) passed before the agent started moving and either stretched out her index finger to point at the closed container and exclaim

“Da!” (“There!”, pointing condition) or held up her hand to cover her mouth, making a coughing noise (control condition). Pointing and coughing movements and sounds were time-matched and lasted 1400 ms, with the peak of the movement (hand farthest away from the face) 600 ms into the movement. After the actor had withdrawn the hand completely, the end frame was shown for another 1200 ms before the next trial started. Each video lasted either 6000 ms or 6400 ms, depending on the variable time between the box movement and the social movement.

We chose a coughing gesture as a control because it is similar to the pointing gesture in some ways (the raising of a hand in front of the face, a sound that is co-occurring with the hand movement) but dissimilar in other ways: Coughing is not a gesture directed "at" something but directed inward – in comparison to other non-communicative gestures like grasping –, and it is non-communicative in that there is no referent for the cough. Because of the way we filmed both types of stimuli, we were able to present pointing and coughing in a visually very similar manner, and in post-processing we time-matched the two movements and sounds as precisely as possible, so that for both gestures, the time of oncoming, the peak, and the returning to the original position were all on the same timeline.

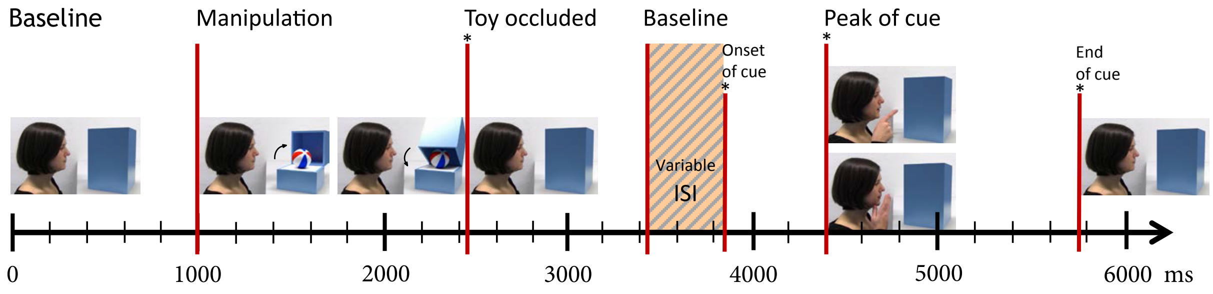


Figure 4.5: Scheme of the experimental design for Experiment 2.

There were 18 different shapes and 6 colors of containers, so that the content or events in the video could not be inferred from the container. 18 different objects comprising toys, food items or otherwise known objects (e.g. a rubber duck, a banana or a sippy cup) were used as content of the containers. All objects and containers were animated using Blender© and subsequently superimposed on the live action video using Adobe After Effects©.

In one third of the videos, the container was empty and there was no movement following it. These filler trials lasted for 4400 ms and were mixed with the test trials in order to minimize habituation effects for the test trials. Order of long, short and filler trials as well as object and type of container were pseudo-randomized so that the same type of trial never appeared more than twice in a row and the same object or container never appeared more than once within one block of nine trials.

In order to retain the infant's attention, videos were presented in eight blocks of nine trials each. After a block of experimental videos, a short, infant-appropriate video of 6 – 7 s was presented, showing colorful scenes (e.g. ducks on a pond or umbrellas dancing in the rain) accompanied by cheerful music. These attention-getters helped the infant refocus towards the screen and break up the repetitiveness of the stimulus material. Attention-getters were not repeated, with 11 different versions playing in a fixed order. After each attention-getter, the infant was eased back into the presentation of the test videos by a "hello" video (3 s). This video showed the same scene as the experimental videos, with the agent turning her head towards the infant and greeting her with a friendly "hello" before turning back towards the table. In the "hello" videos, no container was present on the table.

Since the lateral view on the agent drew attention towards one side of the screen, videos were mirrored and sides were flipped with every new block. Video size was rendered to 1200 x 690 px. Presentation size of the video was 32.0 by 18.5 cm, which corresponds to a visual angle of 22.6° horizontally and 13.3° vertically.

Four different versions of pseudo-randomized presentation orders were created, with the constraints that the same color or shape of containers and the same objects could not be repeated within the same block of presentation, and all objects appeared about the same amount of time overall. These fixed presentation orders were implemented in the Presentation© scripts.

Infants were randomly assigned to a fixed order when they came in. As in Experiment 1 - Communication-induced gamma oscillations at 12 months, all videos were queued to minimize buffering times, and the experiment included an at-will attention-getter and a stop function. Each experimental trial contained two triggers: The first one at 2200 ms, the moment of complete occlusion of the toy, and the second one at 4000 ms (4400 ms for the long version), the peak of the social cue.

4.4.1.3 Data Processing

Data were prepared for further processing in the same way as in Experiment 1 - Communication-induced gamma oscillations at 12 months. For the occlusion event of each trial, data were segmented from -1800 to 1400 ms, with 0 ms being the time of complete occlusion of the object. For the cue event of each trial, data were segmented around each cue starting at -2700 to 1600 ms, with 0 ms being the peak of the cue. Only segments that were entirely marked as “watched” were kept. The cue event segments were then further cut to -1000 to 1600 ms to reduce the chance of artifacts within a segment. Artifact rejection and interpolation was identical to Experiment 1 - Communication-induced gamma oscillations at 12 months.

The Morlet transform was identical to Experiment 1 - Communication-induced gamma oscillations at 12 months, with a 200 ms time period before the box starts opening defined as baseline for occlusion events (-1600 ms to -1400 ms). For cue events, a 200 ms time period before the person starts moving was defined as baseline (-800 ms to -600 ms). Total Gamma was calculated by averaging trials for each infant.

Following this step, data from individual channels were pooled into different combinations of two or more channels. For the occlusion event, time bins of 200 ms lengths were then averaged and exported from Analyzer2 for statistical analysis, comprising the time from -1200 ms before complete occlusion to 1200 ms post occlusion. This long time interval was defined so that it included the time frame in which the object was fully visible, for comparison with the time in which it was fully occluded. Because changes in gamma activity could be short-lived, we chose overlapping bins, so that 23 time bins in total were exported.

We defined ROI based on previous literature and on Experiment 1 - Communication-induced gamma oscillations at 12 months in temporal and frontal regions. For the object maintenance effect, we defined a larger, more central pooling on the right (C4, CP2, CP6, P4) in concordance with the original finding of Kaufman et al. (2003). For investigating the social effect, we enlarged the temporal ROI from Experiment 1 to also include a central electrode (left: C3, FC5 and T7, right: C4, FC6 and T8). The two frontal sites were F3, FC1 and Fp1 for the left hemisphere and F4, FC2 and Fp2 for the right hemisphere, mirroring Experiment 1 - Communication-induced gamma oscillations at 12 months, but moving more laterally to further distinguish between left and right hemisphere.

4.4.2 Results

In segmenting the data for artifact rejection and further analysis, we again separated the occlusion event from the communicative/noncommunicative cue in each trial. Therefore, results for the occlusion event and the cue event are reported separately. However, it has to be stressed that all communicative/noncommunicative cue events in the final analysis were preceded by a successfully watched occlusion event (see 4.4.1.3). Therefore, all cue events are considered in the context of the knowledge of a toy having appeared and now being hidden, which is important when the results are interpreted in terms of referential intent and meaning.

4.4.2.1 Results of the Occlusion Event

Because of the length of the segment (baseline, appearance of object, occlusion) increased the likelihood of artifact contamination, infants had to provide at least 5 artifact-free trials to be included in the analysis. In the pointing condition, 16 infants (11 girls) contributed an average of 10 trials each (range 5 – 18). In the control condition, 22 infants (11 girls) contributed an average number of 12 trials each (range 5 – 39).

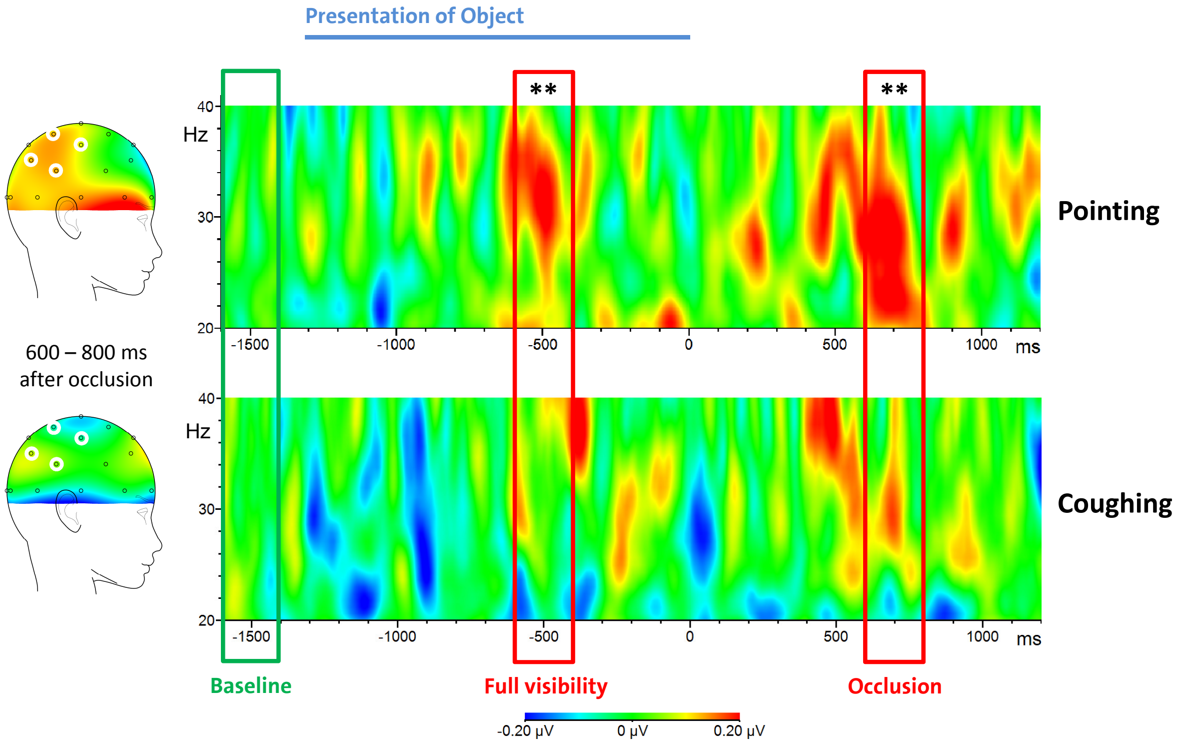


Figure 4.6: Topographic maps (left) and time-frequency plots (right) of of gamma activity at temporal channels (C4, CP2, CP6 and P4) after occlusion in the point condition (top) and the control condition (bottom) in 10-month-olds in Experiment 2. Baseline from -1600 to -1400 ms. Object is completely occluded at 0 ms. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

In response to the occlusion event, we found heightened activity in right temporal-parietal regions (C4, CP2, CP6, P4) 600 to 800 ms after the complete occlusion of the toy (pointing condition: $M_{\text{“baseline”}} = .00006$ µV, $SD_{\text{“baseline”}} = .003$ µV, $M_{\text{“test”}} = .117$ µV, SD«test» = .150 µV, see Fig. 4.6, top; control condition: M«baseline» = -.001 µV, SD«baseline» = .005 µV, M«test» = .031 µV, SD«test» = .246 µV, see Fig. 4.6, bottom). A mixed-design 2 (time) x 2 (condition) ANOVA showed a main effect for time ($F_{(1,36)} = 4.57$, $p = .039$, $n_p^2 = .113$), no main effect for condition ($F_{(1,36)} = 1.55$, $p = .221$, $n_p^2 = .041$), and no interaction between time and condition ($F_{(1,36)} = 1.48$, $p = .231$, $n_p^2 = .040$). Paired comparisons following the main effect of time found a significant increase in gamma activity compared to baseline in the pointing condition ($t_{(15)} = 3.11$, $p = .007$) but not in the control condition ($t_{(21)} = .610$, $p = .549$).

In the same pooling, we also found a burst of activity during the presentation of the toy (-600 to -400 ms before complete occlusion), pertaining to the pointing condition (M«baseline»= .00006 µV, SD«baseline»= .003 µV, M«test»= .098 µV, SD«test»= .132 µV, see Fig. 4.6, top; control condition: M«baseline» = -.001 µV, SD«baseline» = .005 µV, M«test» = .006 µV, SD«test» = .188 µV, see Fig. 4.6, bottom). A 2 mixed-design (time) x 2 (condition) ANOVA found a tentantive main effect for time ($F_{(1,36)} = 3.67$, $p = .063$, $n_p^2 = .092$), no main effect for condition ($F_{(1,36)} = 2.89$, $p = .098$, $n_p^2 = .069$), and no interaction between time and condition ($F_{(1,36)} = 2.73$, $p = .107$, $n_p^2 = .071$). A direct comparison confirmed the increased activity compared to baseline for the pointing condition ($t_{(15)} = 3.00$, $p = .009$), and no increase in the control condition ($t_{(21)} = .18$, $p = .860$).

In left temporal-parietal regions (C3, CP1, CP5, P3), there was no change in activity during occlusion (pointing condition: M«baseline» = .001 µV, SD«baseline» = .004 µV, M«test» = -.012 µV, SD«test» = .106 µV; control condition: M«baseline» = .001 µV, SD«baseline» = .005 µV, M«test» = .033 µV, SD«test» = .191 µV). An ANOVA showed no main effect for time ($F_{(1,36)} = .13$, $p = .718$, $n_p^2 = .004$), no main effect for condition ($F_{(1,36)} = .743$, $p = .394$, $n_p^2 = .020$), and no interaction between time and condition ($F_{(1,36)} = .73$, $p = .399$, $n_p^2 = .020$). Similarly, there were no differences between conditions during presentation (pointing condition: M«baseline» = .001 µV, SD«baseline» = .004 µV, M«test» = .060 µV, SD«test» = .142

µV; control condition: M“baseline” = .001 µV, SD“baseline” = .005 µV, M“test” = .022 µV, SD“test” = .268 µV). A mixed-design ANOVA showed no main effect for time ($F_{(1,36)}$ = 1.18, p = .285, n_p^2 = .032), no main effect for condition ($F_{(1,36)}$ = .25, p = .618, n_p^2 = .007), and no interaction between time and condition ($F_{(1,36)}$ = .27, p = .605, n_p^2 = .007).

4.4.2.2 Results of the Communicative/Noncommunicative Cue

Infants had to provide at least 8 artifact-free trials to be included in the final analysis. In the pointing condition, 21 infants (14 girls) contributed an average of 12 trials each (range: 8 – 25), and in the control condition, 17 infants (7 girls) contributed an average of 14 trials each (range: 8 – 38).

Based on the results of the social condition in Experiment 1 - Communication-induced gamma oscillations at 12 months, here we again looked at gamma activity during the presentation of the social cue. On right temporal channels (C4, FC6, T8), both the pointing and the control cue showed a significant increase in gamma activity during the cue (-200 pre to +600 ms post peak of cue, pointing condition: M“baseline”= .0009 µV, SD“baseline”= .004 µV, M“test”= .171 µV, SD“test”= .250 µV, see Fig. 4.7, top; control condition: M“baseline” = .001 µV, SD“baseline” = .006 µV, M“test” = .197 µV, SD“test” = .259 µV, see Fig. 4.7, bottom). A mixed-design 2 (time) x 2 (condition) ANOVA showed a main effect for time ($F_{(1,36)}$ = 19.87, p = .000, n_p^2 = .356), no main effect for condition ($F_{(1,36)}$ = .10, p = .759, n_p^2 = .003) and no interaction between time and condition) $F_{(1,36)}$ = .091, p = .765, n_p^2 = .003). Paired t-Tests confirmed that both the pointing condition ($t_{(20)}$ = 3.14, p = .005) and the control condition ($t_{(16)}$ = 3.16, p = .006) were significantly increased from baseline.

Increase in gamma activity in left temporal channels (C3, FC5, T7) did not reach significance compared to baseline (pointing condition: M“baseline”= .001 µV, SD“baseline”= .005 µV, M“test”= .077 µV, SD“test”= .209 µV; control condition: M“baseline” = .002 µV, SD“baseline” =

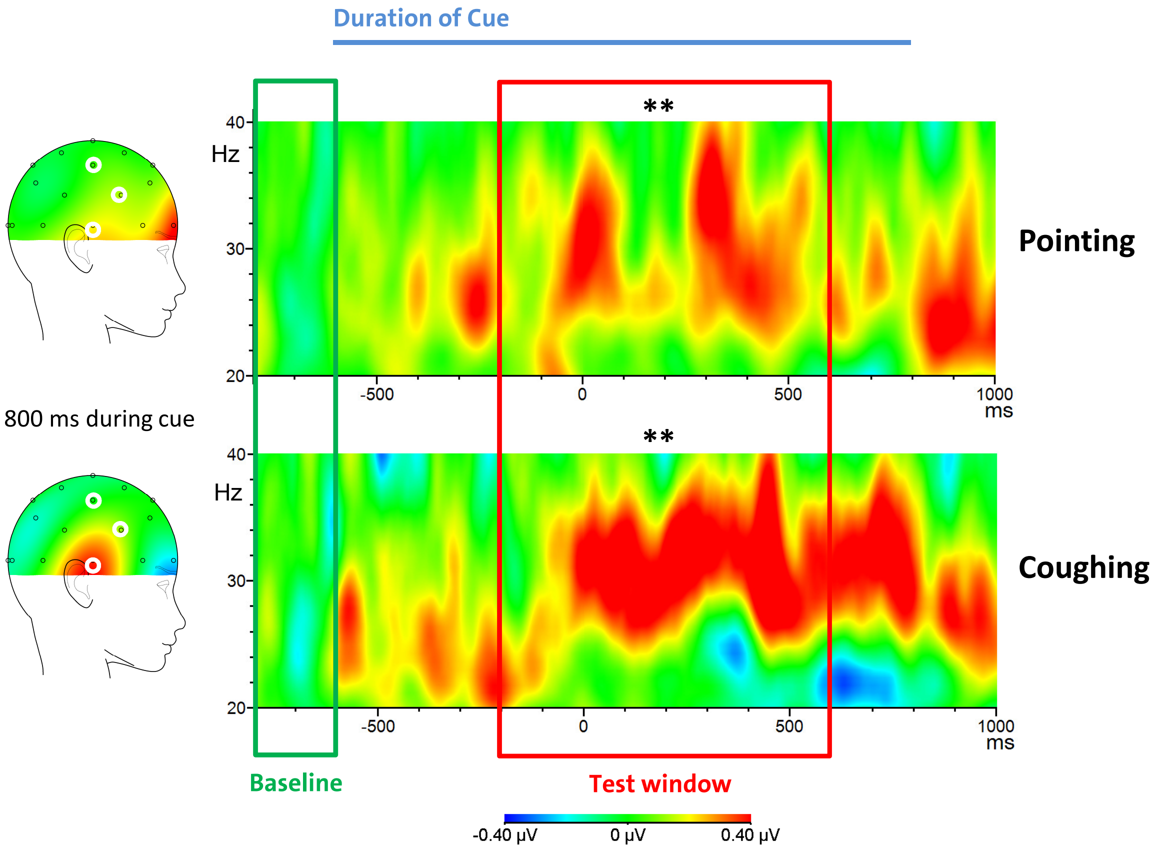


Figure 4.7: Topographic maps (left) and time-frequency plots (right) of gamma activity during the presentation of the cue at right temporal channels C4, FC6 and T8 in the point condition (top) and the control condition (bottom) in 10-month-olds in Experiment 2. Baseline from -800 to -600 ms. Cue starts at -600 ms and ends at 800 ms. Peak of point at 0 ms. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

.007 µV, M$_{\text{“test”}}$ = .043 µV, SD$_{\text{“test”}}$ = .201 µV). A mixed-design ANOVA showed no main effect for time at $F_{(1,36)} = 3.01$, $p = .091$, $n_p^2 = .077$), no main effect for condition $F_{(1,36)} = .25$, $p = .618$, $n_p^2 = .007$) and no interaction between the two $F_{(1,36)} = .28$, $p = .598$, $n_p^2 = .008$).

However, during the same time span, we found increased gamma activity in left prefrontal regions (F3, FC1, Fp1) pertaining to the pointing condition only (pointing condition: M$_{\text{“baseline”}}$= .0004 µV, SD$_{\text{“baseline”}}$= .008 µV, M$_{\text{“test”}}$= .140 µV, SD$_{\text{“test”}}$= .200 µV, see Fig. 4.8, top; control condition: M$_{\text{“baseline”}}$ = -.002 µV, SD$_{\text{“baseline”}}$ = .007 µV, M$_{\text{“test”}}$ = -.074 µV,

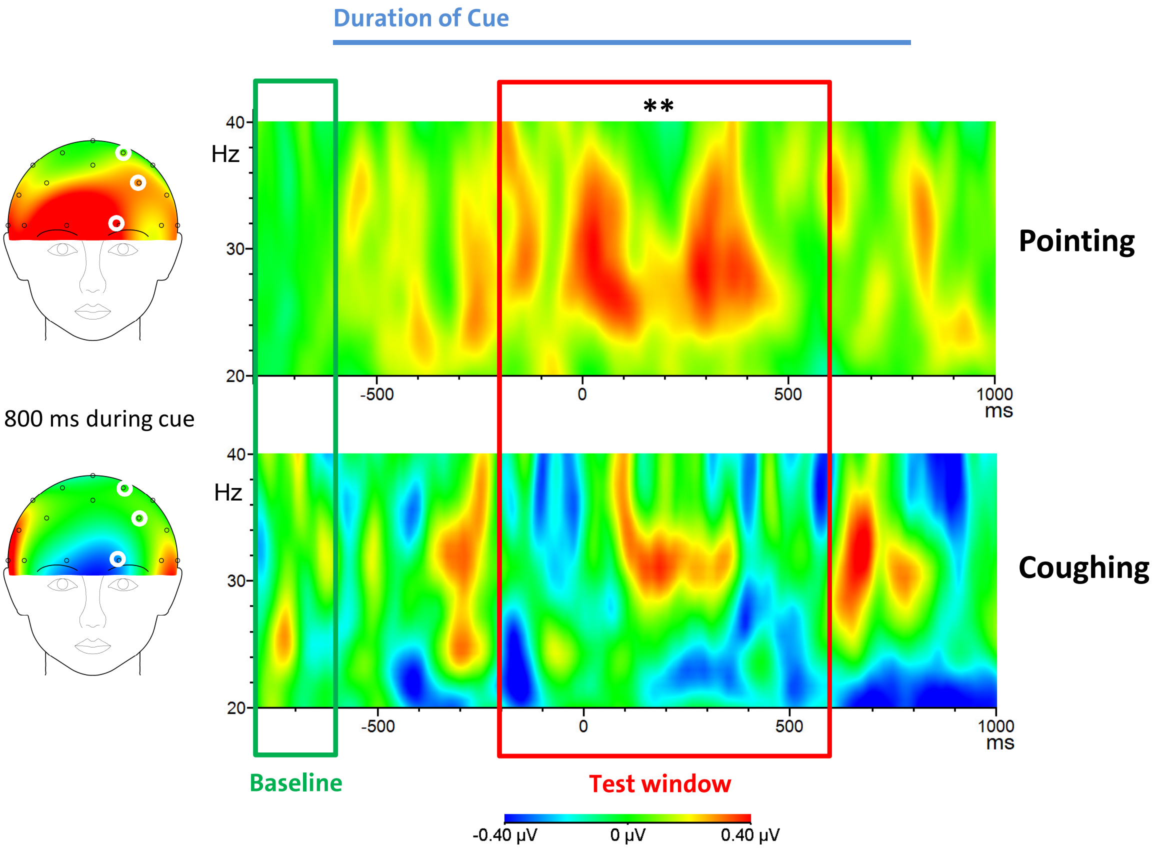


Figure 4.8: Topographic maps (left) and time-frequency plots (right) of gamma activity during the presentation of the cue at left frontal channels F3, FC1 and Fp1 in the point condition (top) and the control condition (bottom) in 10-month-olds in Experiment 2. Baseline from -800 to -600 ms. Cue starts at -600 ms and ends at 800 ms. Peak of point at 0 ms. Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

SD$_{\text{“test”}}$ = .314 µV, see Fig. 4.8, bottom). A mixed-design 2 (time) x 2 (condition) ANOVA showed no main effect for time ($F_{(1,36)} = .66$, $p = .422$, $n_p^2 = .018$), a significant main effect for condition ($F_{(1,36)} = 6.63$, $p = .014$, $n_p^2 = .155$) and a significant interaction between time and condition ($F_{(1,36)} = 6.47$, $p = .015$, $n_p^2 = .152$). Subsequent paired comparisons showed that gamma activity was only significantly increased compared to baseline in the pointing condition ($t_{(20)} = 3.25$, $p = .004$), but not in the control condition ($t_{(16)} = .95$, $p = .357$). An independent-samples t-Test showed that the two conditions were significantly different from each other ($t_{(36)} = 2.56$, $p = .015$).

In right prefrontal regions (F4, FC2, Fp2), activity appeared similarly increased (pointing condition: M“baseline”= .002 µV, SD“baseline”= .005 µV, M“test”= .159 µV, SD“test”= .266 µV; control condition: M“baseline” = -.0008 µV, SD“baseline” = .007 µV, M“test” = -.055 µV, SD“test” = .436 µV). A mixed-design 2 x 2 ANOVA showed no main effect for time ($F_{(1,36)}$ = .81, p = .373, n_p^2 = .022), no significant main effect for condition ($F_{(1,36)}$ = 3.55, p = .068, n_p^2 = .090) and no significant interaction ($F_{(1,36)}$ = 3.41, p = .073, n_p^2 = .086). While these results did not warrant paired comparisons, they give tentative evidence that the effect of the communicative cue may not be restricted to the left hemisphere.

4.4.3 Discussion

Experiment 2 confirms and extends the object maintenance effect found in Experiment 1. We found an increase in gamma activity in right temporal-parietal regions (overlapping with those reported by Kaufman et al. (2003, 2005) and Kampis et al. (2015) 600 – 800 ms post occlusion. The timing and frequency range of this effect was also very similar to the object maintenance effect reported in Experiment 1, whereas the localization may be evidence of a shift from activity focused in the inferior temporal cortex (at 10 months) to activity focused in the frontal cortex (at 12 months, see also Baird et al., 2002). Interestingly, this finding pertained only to the occlusion event in the point condition, but not to the occlusion event in the control (cough) condition. We found another burst of gamma activity in the same ROI before occlusion, during the presentation of the object. Csibra, Davis, et al. (2000) reported bursts of gamma activity in occipito-parietal sites concurrent with the onset and offset of the visual stimuli. This increase in activity during the visibility of the object may therefore be a sign of object processing in the primary visual cortex (Wilcox, Bortfeld, Woods, Wruck, & Boas, 2005). Again, we reported this effect only in the point condition.

In relation to the presentation of the cue, we found that activity in right temporal sites significantly increased during both the communicative and the noncommunicative social cue. There was no difference between conditions in terms of this effect, suggesting that temporal activation was not a result of processing of communicative intent, but rather a sign of the processing of social stimuli in general.

However, we found a significant increase in gamma activity in left frontal regions pertaining to the communicative gesture only, starting shortly before the peak of the pointing gesture and ending shortly before the point was completely retreated (lasting 800 ms in total). In right frontal regions, we found a similar, but less pronounced pattern. The control (cough) gesture did not elicit a similar effect.

4.5 General Discussion

In two experiments, we explored gamma activity pertaining to object occlusion as well as social cues. In Experiment 1, 12-month-old infants showed gamma activation 300-500 ms post occlusion in right frontal regions. In Experiment 2, we report increase in gamma activity in 10-month-olds over right temporal-parietal regions at 600 – 800 ms post occlusion only in the pointing cue condition and not in the control condition. This is more in line with previous studies, who also reported gamma activity in temporal regions following occlusion (Kampis et al., 2015; Kaufman et al., 2003, 2005). The difference between Experiment 1 and Experiment 2 may therefore also lie in the developmental advance between 10 months and 12 months. As shown in a study using pupillometry, 12-month-olds succeed in understanding a pointing gesture as referential, whereas 8-month-olds do not (Pätzold & Liszkowski, 2018).

Additionally, in Experiment 2 we found a burst of gamma activity in the same temporal-parietal region during the presentation of the object, which may reflect the activation of the visual cortex in conjunction with the representation of a colorful toy. Activation during the presentation and after the occlusion of the object pertained to the pointing condition only, which may be an indication that infants represent what the actor is focusing on. Temporal activity during the cue was not different between pointing and control conditions, which suggests that some aspects of the gestures may be processed in similar ways.

Can Experiment 2 fully explain the "social" effect found in Experiment 1? The social cue elicited similar bilateral temporal activation patterns in the two experiments. In Experiment 1, infants showed increase of gamma activity in temporal regions during the social cue, but not the nonsocial cue, right before the cue reached its peak. Similarly, in Experiment 2, bilateral temporal gamma activity was recorded during the social cue, starting around the peak of the cue and ceasing around its end. However, we found this temporal increase in activity in both communicative and non-communicative social conditions. Therefore, we reason that the increase in activity in Experiment 1 is not a "second burst" signifying object representation, but rather a sign of a general social-cognitive processing, as the stimulus of the video of a person is much richer and more salient to the infant than an artificial cue consisting of a light and sound.

However, in Experiment 2 we found significant differences between the pointing condition and the control condition both during the cue, in frontal regions, and in relation to the object maintenance effect, which was restricted to the pointing condition.

Activity in left frontal regions during the pointing gesture indicates that infants selectively processed pointing as a communicative cue different from an accidental gesture. This is in line with NIRS research showing left dorsal prefrontal activation in response to joint attention in five-month-olds (Grossmann & Johnson, 2010; Grossmann et al., 2010, see also Urakawa, Takamoto, Ishikawa, Ono, & Nishijo, 2015, for similar evidence in 7-month-olds during live interaction).

The object maintenance effect was constrained to the referential-communicative context. These findings lead us to believe that infants processed the object very differently in the referential-communicative situation than in the non-referential social situation. From our evidence, it appears that their comprehension of the social situation influenced how the object was represented during presentation and occlusion. We posit that the referential nature of the cue heightened the salience and importance of the object, facilitating its maintenance, whereas the non-referential context did not lead the infant to actively process the object's presence or absence.

This new hypothesis evokes the propositions by Csibra and Gergely (2009) and Csibra (2010), who suggest that *ostensive cues* elicit a referential expectation in infants, which influences how the infant processes the content of a situation. Recent studies investigating the influence of ostensive cues show that the ostensive context can both facilitate (Senju & Csibra, 2008) and hinder (Topál et al., 2008) object processing. However, in the current Experiment 2, we did not explicitly manipulate the ostensive context of the two conditions. In both the pointing and the control condition, infants saw the "Hello" videos, which offered two important ostensive signals (eye contact and infant directed speech). Furthermore, the agent directed attention towards the object using eye gaze in both conditions, which has been described as enough to establish reference (Csibra & Volein, 2008). Rather, the difference between the two conditions was whether the agent tried to initiate joint engagement of the infant with the object (triadic interaction, pointing condition), or whether the communicative attitude was restricted to the occasional greeting of the infant (dyadic interaction, control condition). This manipulated the relevance of the object for the infant, who assigned importance to the events surrounding the appearance and occlusion of the object based on what was transported by the agent.

Our finding may therefore be more in line with the argument brought forward by Kampis et al. (2015): Infants take into account the perspective of their social partner. In the case of the control condition, they find the other person to be uninterested in the object. This leads them to let go of the object representation themselves, similar to other studies which show that object maintenance is upheld only when the object continues to be relevant (Kaufman et al., 2005; Tallon-Baudry et al., 1998). From the study of infant attention, there is evidence that joint engagement helps infant focus their limited attentional resources to the referred-to object compared to having to focus on it without reference (Striano, Reid, & Hoehl, 2006) and that 9-month-old, but not 12-month-old infants preferentially attend to an object that was previously part of joint attention (Striano, Chen, Cleveland, & Bradshaw, 2006). More recently, Michel et al. (2017) presented neurophysiological evidence that even proto-communicative cues (schematic eyes) enhanced object processing in 4-month-olds. While these studies did not investigate object maintenance, they suggest that the way infants initially attend to the novel object may already be different depending on the communicative context.

Here, we suggest that the relevance created by the referential-communicative context also influences how infants represent occluded objects. The current findings therefore shed a new light on the development of object maintenance in relation to its referential context. Further experiments are needed to test how object representation may be facilitated by social relevance over the developmental course, and how social relevance may be elicited by different referential cues.

5 General Discussion

> *"They were not the same eyes with which he had last looked out at this particular scene, and the brain which interpreted the images the eyes resolved was not the same brain. There had been no surgery involved, just the continual wrenching of experience."*
>
> – Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

In this thesis, I used new physiological and neurophysiological measures of infant cognition to provide evidence of referential-communicative understanding and object representation at the end of the first year. In the following discussion, I will put together the findings of the three studies to propose a new hypothesis on the interrelation between referential-communicative cues and object representation.

5.1 Summary of Findings

In Study 1 (Chapter 2), I presented data on object expectation following a pointing gesture. Infants watched videos of an agent pointing towards a closed door. The door then opened to reveal a toy or an empty table. I demonstrated that, at 12 months, infants' pupils dilated when they saw the agent point but no toy appeared, i.e. their expectation towards the appearance of an object was violated. At 8 months, infants did not differentiate between expected and unexpected outcomes in the same experiment. This study delivered direct evidence of referential expectation induced by communication, because a similar nonsocial, attention-directing cue did not elicit an effect of violation of expectation. In contrast to similar studies presenting evidence of referential understanding (Behne et al., 2012; Woodward & Guajardo, 2002), infants had to infer a referent that they had not previously seen, further supporting the notion that they understood the referential intent of the pointing gesture.

In Study 2 (Chapter 3), I explored object representation without social context in a paradigm that was not unlike that of Study 1. Instead of a communicative cue, a logical action sequence elicited the context in which infants had to infer the presence of an object. In a 2 x 2 VOE paradigm, toys were placed behind a door, or moved out of the scene completely. I tested whether infants still expected an object to be behind the door when it reopened (object expectation), or whether there still expected the area to be empty after an object had moved out of it before the door closed (nothingness expectation). Eighteen-month-olds showed a relative increase in pupil size in response to the violation of object expectation, but not in response to the nothingness expectation. Contrary to previous findings in younger infants using a variety of measures (Baillargeon, 1998; Jackson & Sirois, 2009; Kaufman et al., 2003), there was no pupil dilation in infants 10 months of age in response to the VOE conditions. This finding is puzzling considering the current literature on object permanence as an early implicit competence (Kaufman et al., 2003, 2005; Ruffman et al., 2005) and my findings in 12-month-olds in Study 1. However, being able to directly compare the two paradigms of Studies 1 and 2, I find that differences in the communicative context may indeed explain why 10-month-olds fail the paradigm of Study 2.

Finally, in Study 3 (Chapter 4), I assessed activity in the gamma frequency band of the EEG as a marker for object representation in a communicative context. I was able to replicate earlier findings of prolonged gamma activity in the temporal cortex following object occlusion (Kampis et al., 2015; Kaufman et al., 2003, 2005) in 10-month-olds (Experiment 2), but 12-month-olds seemed to show the effect in right frontal regions rather than temporal regions (Experiment 1). Interestingly, the object maintenance effect in Experiment 2 pertained only to the pointing group, but not to the control group who never saw a referential cue towards the container. In addition, I found increase in both frontal and temporal gamma activity during the course of the pointing gesture, but no frontal activity during the control gesture. This led me to suspect a link between referential communication and object representation.

To unravel the common theme of the three studies, I will now discuss the two most surprising findings – the failure of the VOE paradigm to provide evidence of object permanence in 10-month-olds in Study 2 and the limitation of the object maintenance effect to the communicative condition only in Experiment 2 of Study 3 – to form a new hypothesis on early object representation.

5.2 Two Leads to a New Understanding of Early Object Representation

In the introduction, I separately discussed social cognition (referential understanding) and object cognition (object representation during occlusion). I made the proposition that object permanence must be a prerequisite for expecting an object in a communicative context: Only when the infant knows that objects continue to exist even after being occluded, is it plausible to assume that they expect an object to appear based on communicative cues made by others (see Section "The Development of Object Cognition" in Chapter 1). However, from the present data it appears that the two concepts are interlinked in a different way at the end of the first year of life.

The first lead to a new understanding of object representation comes from the two pupillometry studies. In Study 1, I found early evidence for referential pointing comprehension (12 months) but in Study 2, I found late evidence for object permanence (18 months, no evidence at 10 months). Although the experiments were not identical, the general principle was the same. At the start of each trial, an expectation concerning the presence of an object was raised, then fulfilled or violated during the test phase depending on the condition. Age groups who "succeeded" in the VOE condition in either study did so in an obvious, sustained way (compare Figures 2.3 and 3.4), whereas age groups who did not show differences in the VOE condition demonstrated a pupillary response consistent with the visual properties of the stimuli only.

It seems striking that infants should be able to solve a much more difficult problem – infer an object they have never seen from a communicative cue – but not be able to keep track of a scenario in which an object is presented for several seconds, is occluded in full view for a reasonably short time of 2 s, and then fails to reappear. One possible explanation for this divergence may lie in the differences between the two paradigms. In Study 1, the door frame was never presented open before the outcome segment. Therefore, infants had not previously experienced it to be full or empty. In Study 3, all trials started with the opening of the door, which was always empty in the beginning. After the door opened, the animated motion sequence started. Therefore, infants may have learned to associate the door opening to be empty with the start of an interesting event. The main effect for outcome I found in both 18- and 10-month-olds emphasizes this interpretation – the pupil remains relatively larger when the outcome is empty, which is in line with the continued cognitive effort during an unresolved or inconsistent scene (Steinhauer, Condray, & Kasparek, 2000; Verschoor et al., 2013). Nevertheless, 18-month-olds were able to differentiate between an empty outcome constituting a violation or a continuation of the sequence, whereas the 10-month-olds did not. In a recent pupillometry study balancing the opening sequence to be either empty or full (Pätzold & Liszkowski, unpublished data), I found that 12-month-olds showed a tendency for a VOE effect in the test for object expectation in the 1000 ms time window (p = .091), but not the clear results of the 18-

month-olds and no sustained effect.

The more probable explanation, therefore, is that 10-month-olds in Study 2 were missing something that helped them understand the problem posed – something that was obvious enough to the more advanced 18-month-olds – whereas the 12-month-olds in Study 1 had all the contextual cues they needed to make sense of the situation.

Two, in Study 3, Experiment 2, I made the unexpected discovery that only infants who were exposed to the pointing condition, but not the ones exposed to the non-referential control condition, showed the object maintenance effect, combined with prolonged frontal activity during the pointing itself. This holds true even though both pointing and the control cue elicited concurrent temporal gamma activity, indicating that there may be an overlap in how the two cues are processed.

The way the object was presented and occluded in both conditions was completely identical. However, the agent never referred or acted on the object or the process of its reveal or occlusion in the control condition. Gaze direction alone may not have been sufficient to establish reference in this case (Senju & Csibra, 2008) and although coughing occurred time-locked to the event, it may have happened too late to be perceived as contingent (Csibra, 2010). As a social partner, then, the agent appeared uninterested in communicating about the object, which may have led the infant to view the object as irrelevant in this context.

In the pointing condition, however, the communicative gesture emphasized the object's relevance, making it important in this social context to maintain an object representation. Previous studies showed that relevance is needed to induce the object maintenance effect. In the study by Kaufman et al. (2005), the authors differentiated between occlusion and disintegration events. Even though perceptually similar to the occlusion, the disintegration event was incompatible with the continued existence of the object, therefore rendering a maintenance irrelevant. In consequence, infants did not represent the object that had disintegrated. Similarly, in the early adult studies (e.g. Tallon-Baudry et al., 1998), the maintenance of the object was dependent on the relevance of a continued representation for the task. If a representation was not needed for the task, it was not maintained. These examples show that object maintenance is not an automatic process, but depends on the relevance of the representation. How this relevance is created may differ, depending on the context in which the stimuli are delivered. In the case of Experiment 2 in Study 3, it appears I unknowingly not only manipulated the referential intent of the agent, but also the communicatively transported relevance of the object.

5.3 The Communicative Relevance Hypothesis

These two surprising findings in my studies challenge the traditional view of a nonsocial, detached development of cognitive skills in the line of Baillargeon (1998) and Spelke and Kinzler (2007) and support an integrated perspective of social cognition being inseparable from human thinking at large (Tomasello, 2014). The communicative environment influences how the infant processes the object, and whether a representation is formed. It appears that communicative cues heighten the salience of referred-to object, which enables the infant to represent it better than in a noncommunicative context. In contrast to a simple attentional cue, the communicative cue works because it makes the object relevant to the infant for shared attention and a possible interaction with the agent.

This attunement to communication helps infants pick out the objects they are most likely to learn more about, because the social partner may do something interesting with them after communicating about them. This may be how infants start thinking about objects in the physical world. If the social partner does not behave in a communicative manner towards the object, infants may be less inclined to attend to the object, either because they are not able to hold the object in active memory without communicative help, or because the noncommunicative attitude actively reduces the object's relevance. Following the evidence provided here, it appears that at the end of the first year, infants rely strongest on communicative cues, and may not be able to fully use other purely attentional cues. Only when infants have had a certain exposure to the world, or to social interaction, or both, can they peel away from the communicative context and focus on cognitive tasks using other salient cues that are not provided by others. I call this the *communicative relevance hypothesis*. In a way, this proposition is not all that different from Vygotsky's idea of a social foundation of cognition, but I am not making a claim that all cognition is only possible through social construction. Rather, at a specific time in development, when infants are just mastering the use of referential cues, they benefit from a communicatively induced object relevance in a way that helps them perform better in cognitive tasks.

Previous findings often do not report such contextual dependence on communicative relevance. For example, concerning the object maintenance effect, Kaufman et al. (2003, 2005) found object maintenance in the absence of communicative cues. Kampis et al. (2015) included a social context very similar to my control condition (gaze direction, but no communication), and still reported the object maintenance effect. How can these diverging findings be integrated with my suggestion?

Kaufman et al. (2003) actually found heightened gamma activity when a hand came into view to remove an occluder. They explained this finding as evidence for sustained object representation because the infant associated the hand with the reveal of the object. Kampis et al. (2015) argued that this activation may indeed reflect the attribution of an object representation to another person, in this case, the owner of the hand. I however present another explanation: As the hand came into view, the infants expected the agent to refer to or handle the hidden object, which helped them to re-establish an object representation. The infants did not need to have a representation of the agent's mental state to expect an object; they only needed to understand the communicative relevance of the act.

Kampis et al. (2015) found gamma activity both when the infant represented an object and when the infant knew that the agent was representing an object even though it was no longer in view. They interpreted these findings as evidence that infants encoded other agents' mental states, including meta-representations. However, the results can also be interpreted in terms of relevance for the infant: When the object was removed from the agent's view, but not the infant's, it was no longer available for shared attention. Similarly, in the false belief condition, the infant may have still expected the agent to act upon the object that she could no longer see, maintaining its relevance for future interaction. The infant may be ready to show the other person that the container is now empty, which would be equivalent to sharing an exciting visual experience that happened outside the other's view (compare Camaioni et al., 2004).

Why is the communicative context so seldom addressed when reporting infants' cognitive achievements? If fairly considered, a testing situation in the lab is never fully non-communicative, as researchers design the experimental set-up to provide a fun, pleasant atmosphere in which the infant will show the best possible behavior. It is hard to imagine such a set-up without thinking of the caregiver and the research assistants who do everything to make the situation interesting for the child, which often, if not always, includes a good amount of infant-directed speech, joint engagement, and communicative attention-directing. Even during the presentation of stimuli, many cues may be picked up by the infant as communicative. For example, in the study by Kaufman et al. (2003), stimuli were presented contingent with infants' looking, so that a new stimulus was presented if the infant looked away for more than 2 s. This is also standard protocol in many visual habituation paradigms and looking time studies (Colombo, McCardle, & Freund, 2009). In the study by Kaufman et al. (2005), the experimenter played sounds at will to reengage the infant's attention. Both of these measures served to heighten infants' interest in the stimuli and make them more relevant to them.

5.3.1 Integration into Theories

The hypothesis made above fits into the *shared intentionality hypothesis* recently put forward by Tomasello (2014), which posits that human cognition is profoundly shaped by the way humans interact and cooperate with each other. Following this view, the development of all cognitive processes has to be considered in the social context in which they are embedded. With these broad claims, my communicative relevance hypothesis may be a piece of the puzzle of how exactly the social context influences cognition in a concrete developmental situation.

The consideration of communicative relevance also evokes Csibra and Gergely's 2009 *Natural Pedagogy*, which suggests that communication induces referential expectations in infants. However, Csibra and Gergely constrain the formation of expectation to a specific ostensive environment, in which direct eye contact, communicative cues and temporal contingency interplay. Crucially, they reject the induction of referential expectation in situations where an ostensive element is missing (Farroni, Mansfield, Lai, & Johnson, 2003; Senju & Csibra, 2008). In my studies, however, communicative cues were not prepared for their ostensiveness but for the presence of at least one recognizable element of communication (such as an index-finger point). Therefore, direct eye contact was missing in the studies presented in Study 1, Study 2 and Experiment 1 in Study 3. Infant-direct speech was used sparsely in Chapter 1 and Chapter 3, since audio signals were secondary in stimulus design. Contingent responsivity is up for debate – communicative cues were time-locked in stimulus presentation by design, but not contingent on the infants' behavior.

Does this mean my results conflict with Csibra and Gergely's theory? A recent study by Szufnarowska, Rohlfing, Fawcett, and Gredebäck (2014) used a design similar to that reported in the study by Senju and Csibra (2008) to test whether "nonostensive" social cues could elicit gaze following in the same way ostensive cues do. They found that a movement combined with shivering and gaze at the object, but no direct gaze towards the infant, elicited gaze following in the same way that direct gaze did. Szufnarowska et al. discuss this finding in the light of the importance of attention. They argue that not a specific set of cues, but the successful capturing of the infant's attention is crucial to the learning environment that creates referentiality. This finding supports my notion of the importance of the salience of the context, with the "shrugging" condition (deliberate action Szufnarowska et al., 2014) transporting enough intention and purpose, and the "coughing" condition (undeliberate action, Study 3) failing to do exactly that. Therefore, I think *Natural Pedagogy* may be defined too rigidly, and may align with communicative relevance if further explored.

In another line of studies (Reid, Striano, Kaufman, & Johnson, 2004; Reid & Striano, 2005; Striano, Reid, & Hoehl, 2006), the authors explored the influence of social cues on infants' online object processing, leading to the development of the *directed attention model of infant social cognition* (Reid & Striano, 2007). According to this model, infants use social cues as a filter for all incoming information from very early on in infancy. Most recently, Michel et al. (2017) presented neurophysiological evidence that even proto-communicative cues, in this case, schematic eyes, enhanced object processing in 4-month-olds. While the authors did not investigate object representation during occlusion, and used drastically simplified social cues, they suggest that the way infants initially attend to novel objects may already depend on the communicative context.

Although the evidence surrounding the directed attention model are more reminiscent of cueing paradigms such as the ones discussed in Chapter 1 (e.g. Bertenthal et al., 2014, see Section "Current Findings from the Study of Referential Understanding") than the communicatively complex studies presented here, these studies come to similar conclusions as I do concerning the unique role of social cues in heightening infants' attention towards objects. Maybe the biological foundation that helps infants attune towards communication and benefit from them later on lies in this early sensitivity towards social cues.

5.3.2 A Cross-cultural Perspective on Communicative Relevance

In the current studies, I tested a Western demographic, specifically a German urban sample that fits into the WEIRD characteristics (Henrich et al., 2010). Accordingly, the communicative relevance hypothesis is built on a very specific social group and transfer to other cultures has to be carefully considered. If I base this hypothesis on a biological adaptation, as Tomasello (2008) and Csibra and Gergely (2009) would suggest, the consequence would be that a similar facilitation through heightened communicative relevance may be present in all cultures, but might occur at different ages due the actual prevalence of social-communicative cues. It would be plausible that infants in cultures with very little joint engagement between caregiver and infant would show certain object cognition milestones at a later age, because they do not receive the beneficial input and have to exert more cognitive resources to finding relevant information through other contextual clues.

If I base the communicative relevance hypothesis on a constructivist approach, in the tradition of Vygotsky (1981), the opposite would be true: An infant receiving a lot of referential input and social attention would rely on it to get the information they need about objects' relevance, and show a delayed responsiveness to noncommunicatively-induced relevance. Therefore, an infant from a culture with less input would actually be faster in reaching cognitive milestones without social context than their Western counterpart, because they have more experience directing their own attentional resources without relying on communicative cues.

These predictions about the social relevant hypothesis could be tested in cross-cultural studies, or even in studies where the amount of joint engagement is different between groups within one culture. To an extent, even an experimental manipulation of communicative input, as in a training study, might be a conceivable way to test this.

5.3.3 Cross-species Reasonings

Are humans uniquely disposed to make use of communicative relevance, or do other animals also able to benefit from the attentional heightening of a communicatively relevant content? As described in Chapter 1 (Section "The Pointing Gesture"), other great apes do not engage in referential communication, which may be explained by their lack of motivation to engage in cooperative interaction (Tomasello, 2014) or their inability to comprehend another's intention (R. Moore, 2013). They are, however, capable of tracking objects very much in the same way as humans (Call, 2001), an ability which does not seem to be impaired in the slightest by their lack of communicative understanding. In fact, many other animals perform well in object permanence tasks (e.g. birds, see Salwiczek, Emery, Schlinger, & Clayton, 2009 or dogs, see Zentall & Pattison, 2016), without receiving referential cues from conspecifics. Therefore, it appears that the interrelation between referential understanding and object cognition is unique to humans.

5.3.4 Absence Representation in Communicative Context

Following the line of thought presented above, I will attempt to put absence representation into the communicative context. Study 2 showed that 18-month-olds do not form a representation of an empty set in a nonsocial, audiovisual presentation, as demonstrated by their failure to react with surprise when a formerly empty space turns out to be occupied by an object. However, we do know that early in the second year of life, infants use verbal expressions like "(all) gone" to represent absence or search for an object, or denoting an empty container (Gopnik, 1984). At the same age, around 15 months, the majority of infants in an Italian sample produced conventional gestures denoting emptiness as part of their preverbal gestural communication (throwing hands up or shrugging; Caselli, Rinaldi, Stefanini, & Volterra, 2012). A recent eye-tracking study (Liszkowski & Ramenzoni,

2015) showed that 18-month-olds were able to infer the absent referent from a choice of two objects after having watched a person point to the empty place where the referent had previously been. At first glance, these competences seem to point towards a sort of absence representation at 18 months: The infant encodes the absence of an item or stuff from a container or place.

Despite the focus on the emptiness, however, these results may actually indicate object representation (something is gone that should be there) rather than emptiness representation (something is there even though it should not be there). A true absence representation seems to be only possible with the magical appearance task (Wynn & Chiang, 1998). How may the magical appearance task be made easier, in consonance with the approach detailed above? Communicative attention towards the event may help establish the empty container as relevant, enabling the child to encode its empty state. It may not be enough to share attention towards it, but the relevance of the emptiness should be somehow augmented, for instance by way of assigning a function to it. A container could be deliberately kept empty for a later role, so having contents appear would be a violation to that role. This would be an artificial way to assign the "placeholder" status to an empty set, maybe similarly to how adults have incorporated zero into their magnitude scale (Nieder, 2016).

It is unclear how much the context would change the representation of what the empty container stands for in the child's mind, but a suitable control condition could help rule out alternative explanations. In this example, in a control condition, the empty container would also be assigned a functional role, but one that is not violated by having a content. If the infant's reaction is different from the condition in which having content violates the status, this would be evidence for the child's focus on the violation of the role, and not on the emptiness. However, if the children performed the same in both conditions, but different in a control condition where the emptiness is preserved, this would be evidence for emptiness representation.

5.4 Methodological Implications

Here, I employed two novel methodologies to answer questions concerning infant cognition. Pupillometry proved a useful measure of cognitive arousal. Data collection was identical to eye tracking studies. Infants at the end of the first year of life and well into the second year of life showed a similar performance in reaching minimal looking times and providing pupil data, which makes pupillometry a great tool for longitudinal studies. However, I used pupillometry specifically with a VOE paradigm, which is useful in measuring arousal following a violation, but is not designed to track online processing. This renders it more difficult to make statements concerning the emergence of object representations or object maintenance during a manipulation. Other studies have been successfully applying pupillometry to track online changes without using violation of expectation (Fawcett et al., 2017; Verschoor et al., 2015). It is also conceivable to combine both a VOE paradigm and an online measure of change of arousal over time in the same experiment, if the usual luminance constraints are observed.

With the limitations of the VOE paradigm in mind, I specifically applied the EEG power band analysis as a tool to investigate the online representation of objects. However, I found implementation much more tedious and less rewarding than pupillometry. Restrained mobility, directed attention and time spent in a small room imposed high demands on the infants. This led to high attrition rates which detract from the explanatory power of the experiments. One way to improve utility of the EEG might be to use a net/saline system instead of the cap/gel system implemented here, cutting down on preparation time and making recording sessions more successful on average (Bell & Cuevas, 2012; DeBoer et al., 2007). Also, the nature of my video stimuli with two separate events in each trial led to a lower number of valid trials, and contributed to the high attrition rate. In the future, stimuli may be further simplified to allow the analysis of more trials per infant, and include a larger number of infants in the final data set.

Still, the EEG study presented in this thesis offered an important contribution to exploring the context of the object maintenance effect beyond the context of simple object occlusion, advancing the field of infant cognition.

5.5 Limitations

In the studies reported here, I used methods that have not been employed before to answer questions of referential understanding and object expectation. Therefore, the experimental designs were novel and the implications we can take away from these data should be regarded as preliminary. Converging evidence concerning the same questions from different laboratories, using similar, or even different, methodologies, are needed to further explore the role of communicative relevance in regard to object cognition. In particular, the results of the 10-month-olds in Study 2 (Chapter 3) are difficult to explain in the context of other literature surrounding object permanence. While positive evidence of object permanence from global looking time measures may be regarded as susceptible to lower-level explanations, many other studies using more robust measures and better balanced paradigms (Kaufman et al., 2003; Ruffman et al., 2005; Jackson & Sirois, 2009) have also posited positive evidence of object permanence or comparable object cognition abilities by 10 months of age. It remains unclear why 18-month-olds succeeded in the paradigm used in Study 2, but 10-month-olds failed. The communicative relevance hypothesis offers an interesting post-hoc explanation, but the social-communicative context was not directly manipulated in that study. Currently, 12-month-olds and 14-month-olds are tested in the lab to further explore the utility of the VOE paradigm from Study 2 in detecting object permanence.

In regard to all of my results, the interpretation concerning a developmental course has to be considered with caution. The royal road to tracking developmental change is by devising longitudinal studies in which the same participants are tested at different points in time, avoiding false conclusions stemming from inter-individual variability between different groups of subjects (Casey, Tottenham, Liston, & Durston, 2005). Nevertheless, I decided to use a cross-sectional approach here instead, for two reasons. One, I was interested in one developmental stage in particular, the end of the first year of life, because this is the age were behavioral evidence and cognitive interpretations are most at odds. Two, due to the novelty of the methods and the exploratory nature of my paradigms, a cross-sectional approach allowed me to revise my paradigms, and establish an effect with older infants, before testing my target age groups. In the future, a truly developmental approach may help put the findings concerning communicative relevance into the context of other social-cognitive milestones such as language development and ToM competencies.

5.6 Implications for Future Research

In the studies described above, object representation was tested under two premises – one, that an object was occluded by an outside force, without the implication of a social context (Study 2) and two, that the object was occluded in the presence of an agent, who acted on the object either before it was shown (Study 1) or after it disappeared (Study 3). I found that object expectations were influenced by the actions of the communicative agent, which led me to propose the communicative relevance hypothesis.

An option that has not been tested thus far would be to turn the infant into the acting protagonist of the object occlusion, for example in a gaze-contingent eye tracking set-up. Would an infant be surprised to find a box empty, after he or she filled it with a toy themself? My explanatory theme for the findings has been that of communicative relevance: social partners make a situation more relevant, and their interest in an object may heighten its relevance in the child's eyes. The communicative relevance may also work the other way around. A communicative partner may be present while the infant is acting on an object, and the infant may hope to share his or her experience during or after occlusion with that partner. Following this line of thought, I would hypothesize that a communicative relevance induced by providing an enthusiastic recipient of the infant's shared attention would help an infant acting on an object to represent it during occlusion, while an unresponsive or absent partner would hamper the infant's ability to track his or her own action.

5.7 Concluding Remarks

What do infants represent when the see someone point? What do infants represent when a toy has been hidden out of sight? The studies presented here shed a new light on the nature of the mental representations infants form when they process a referential communicative act or an object occlusion event. I provided pupillometric evidence that 12-month-old infants expect the referential content of a simple communicative interaction. I also provided neurophysiological evidence of object permanence, replicating the object maintenance effect in the gamma band in 10- and 12-month-olds. However, an unexpected finding was that the object maintenance effect in 10-month-olds was contingent on the communicative context. 10-month-olds in the noncommunicative condition, and 10-month-olds in a purely nonsocial object permanence task using pupillometry, did not seem to represent the object during occlusion. This led me to draft a new hypothesis on social-cognitive development: Communicative relevance, as elicited by referential cues and other inductions of meaningful communicative contexts, drive object representation at the end of the first year. This new understanding of early object cognition may serve as puzzle piece in our comprehension of the development of human cognition as a whole.

References

Adamson, L. B., & Bakeman, R. (1991). The development of shared attention during infancy. In R. Vasta (Ed.), *Annals of Child Development, Vol. 8* (pp. 1–41). London, England: Jessica Kingsley Publishers.

Aguiar, A., & Baillargeon, R. (1999). 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive Psychology*, *39*(2), 116–157. DOI: `10.1006/cogp.1999.0717`

Ahmed, O. J., & Cash, S. S. (2013). Finding synchrony in the desynchronized EEG: The history and interpretation of gamma rhythms. *Frontiers in Integrative Neuroscience*, *7*, 58. DOI: `10.3389/fnint.2013.00058`

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4). DOI: `10.1167/14.4.1`

Amso, D., & Johnson, S. P. (2008). Development of visual selection in 3- to 9-month-olds: Evidence from saccades to previously ignored locations. *Infancy*, *13*(6), 675–686. DOI: `10.1080/15250000802459060`

Aslin, R. N. (2012). Infant eyes: A window on cognitive development. *Infancy*, *17*(1), 126–140. DOI: `10.1111/j.1532-7078.2011.00097.x`

Aston-Jones, G., & Bloom, F. E. (1981). Norepinephrine-containing locus coeruleus neurons in behaving rats exhibit pronounced responses to non-noxious environmental stimuli. *Journal of Neuroscience*, *1*(8), 887–900. DOI: `10.1523/JNEUROSCI.01-08-00887.1981`

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. DOI: `10.1146/annurev.neuro.28.061604.135709`

Aston-Jones, G., Ennis, M., Pieribone, V. A., Nickell, W. T., & Shipley, M. T. (1986). The brain nucleus locus coeruleus: Restricted afferent control of a broad efferent network. *Science*, *234*(4777), 734–737. DOI: `10.1126/science.3775363`

Aston-Jones, G., Rajkowski, J., Kubiak, P., & Alexinsky, T. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *Journal of Neuroscience*, *14*(7), 4467–4480. DOI: `10.1523/JNEUROSCI.14-07-04467.1994`

Babinsky, E., Braddick, O., & Atkinson, J. (2012). Infants and adults reaching in the dark. *Experimental Brain Research*, *217*(2), 237–249. DOI: `10.1007/s00221-011-2984-5`

Baillargeon, R. (1987). Object permanence in 3½-and 4½-month-old infants. *Developmental Psychology*, *23*(5), 655. DOI: `10.1037/0012-1649.23.5.655`

Baillargeon, R. (1998). Infants' understanding of the physical world. In M. Sabourin, F. Craik, & M. Robert (Eds.), *Advances in psychological science, Vol. 2: Biological and cognitive aspects* (pp. 503–529). Hove, England: Psychology Press/Erlbaum (UK) Taylor & Francis.

Baillargeon, R. (1999). Young infants' expectations about hidden objects: A reply to three challenges. *Developmental Science*, *2*(2), 115–132. DOI: `10.1111/1467-7687.00061`

Baillargeon, R., & DeVos, J. (1991). Object permanence in young infants: Further evidence. *Child Development*, *62*(6), 1227–1246. DOI: `10.1111/j.1467-8624.1991.tb01602`

.x

Baillargeon, R., Graber, M., DeVos, J., & Black, J. (1990). Why do young infants fail to search for hidden objects? *Cognition*, *36*(3), 255–284. DOI: 10.1016/0010-0277(90)90059-S

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191–208. DOI: 10.1016/0010-0277(85)90008-3

Baillargeon, R., Stavans, M., Di Wu, Gertner, Y., Setoh, P., Kittredge, A. K., & Bernard, A. (2012). Object individuation and physical reasoning in infancy: An integrative account. *Language Learning and Development*, *8*(1), 4–46. DOI: 10.1080/15475441.2012.630610

Baird, A. A., Kagan, J., Gaudette, T., Walz, K. A., Hershlag, N., & Boas, D. A. (2002). Frontal lobe activation during object permanence: Data from near-infrared spectroscopy. *NeuroImage*, *16*(4), 1120–1126. DOI: 10.1006/nimg.2002.1170

Baldwin, J. M. (1884). *Mental development in the child and the race: Methods and processes*. New York, NY: Macmillan.

Barlow, H. B., & Levick, W. R. (1969). Changes in the maintained discharge with adaptation level in the cat retina. *The Journal of Physiology*, *202*(3), 699–718. DOI: 10.1113/jphysiol.1969.sp008836

Barrera, M. E., & Maurer, D. (1981). Recognition of mother's photographed face by the three-month-old infant. *Child Development*, *52*(2), 714. DOI: 10.2307/1129196

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly of Behavior and Development*, *21*(3), 205–226.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276. DOI: 10.1037/0033-2909.91.2.276

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo,

L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology 2nd Edition* (Vol. 2, pp. 142–162). Cambridge, England: Cambridge Univ. Press.

Behne, T., Liszkowski, U., Carpenter, M., & Tomasello, M. (2012). Twelve-month-olds' comprehension and production of pointing. *British Journal of Developmental Psychology*, *30*(Pt 3), 359–375. DOI: 10.1111/j.2044-835X.2011.02043.x

Bell, M. A., & Cuevas, K. (2012). Using EEG to study cognitive development: Issues and practices. *Journal of Cognition and Development*, *13*(3), 281–294. DOI: 10.1080/15248372.2012.691143

Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research. Brain Research Reviews*, *42*(1), 33–84. DOI: 10.1016/S0165-0173(03)00143-7

Bertenthal, B. I. (1996). Origins and early development of perception, action, and representation. *Annual Review of Psychology*, *47*, 431–459. DOI: 10.1146/annurev.psych.47.1.431

Bertenthal, B. I., Boyer, T. W., & Harding, S. (2014). When do infants begin to follow a point? *Developmental Psychology*, *50*(8), 2036–2048. DOI: 10.1037/a0037152

Bertrand, O., & Tallon-Baudry, C. (2000). Oscillatory gamma activity in humans: A possible role for object representation. *International Journal of Psychophysiology*, *38*(3), 211–223. DOI: 10.1016/S0167-8760(00)00166-5

Bialystok, E., & Codd, J. (2000). Representing quantity beyond whole numbers: Some, none, and part. *Canadian Journal of Experimental Psychology*, *54*(2), 117–128. DOI: 10.1037/h0087334

Bogartz, R. S., Shinskey, J. L., & Speaker, C. J. (1997). Interpreting infant looking: The event set × event set design. *Developmental Psychology*, *33*(3), 408–422. DOI: 10.1037/0012-1649.33.3.408

Bower, T. G. R., Broughton, J., & Moore, M. K. (1971). Development of the object concept as manifested in changes in the tracking behavior of infants between 7 and 20 weeks of age. *Journal of Experimental Child Psychology*, *11*(2), 182–193. DOI: 10.1016/0022-0965(71)90074-9

Bower, T. G. R., & Wishart, J. G. (1972). The effects of motor skill on object permanence. *Cognition*, *1*(2-3), 165–172. DOI: 10.1016/0010-0277(72)90017-0

Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication.* Sunderland, MA: Sinauer.

Bradshaw, J. L. (1968). Pupil size and problem solving. *The Quarterly Journal of Experimental Psychology*, *20*(2), 116–122. DOI: 10.1080/14640746808400139

Brentano, F. (1973/1924). *Psychology from an empirical standpoint (Second Edition).* London: Routledge & Kegan Paul.

Bretherton, I. (1991). Intentional communication and the development of an understanding of mind. In D. Frye & C. Moore (Eds.), *Children's theories of mind: Mental states and social understanding* (pp. 49–75). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., & Sirois, S. (2013). Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. *Behavior Research Methods*, *45*(4), 1322–1331. DOI: 10.3758/s13428-013-0327-0

Bronson, G. W. (1994). Infants' transitions toward adult-like scanning. *Child Development*, *65*(5), 1243–1261. DOI: 10.1111/j.1467-8624.1994.tb00815.x

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology*, *130*, 67–78. DOI: 10.1016/j.jecp.2014.09.010

Bruner, J. S. (1975). The ontogenesis of speech acts. *Journal of Child Language*, *2*(01), 1–19. DOI: 10.1017/S0305000900000866

Brunetti, M., Zappasodi, F., Marzetti, L., Perrucci, M. G., Cirillo, S., Romani, G. L., ... Aureli, T. (2014). Do you know what I mean? Brain oscillations and the understanding of communicative intentions. *Frontiers in Human Neuroscience*, *8*. DOI: 10.3389/fnhum.2014.00036

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, *4*(3), 121–138. DOI: 10.1080/17470215208416611

Butler, S. C., Caron, A. J., & Brooks, R. (2009). Infant understanding of the referential nature of looking. *Journal of Cognition and Development*, *1*(4), 359–377. DOI: 10.1207/S15327647JCD0104_01

Butterworth, G. E. (2003). Pointing is the royal road to language for babies. In S. Kita (Ed.), *Pointing* (pp. 9–33). Mahwah, NJ: Erlbaum.

Butterworth, G. E., & Cochran, E. (1980). Towards a mechanism of joint visual attention in human infancy. *International Journal of Behavioral Development*, *3*(3), 253–272. DOI: 10.1177/016502548000300303

Butterworth, G. E., & Grover, L. (1988). The origins of referential communication in human infancy. In L. Weiskrantz (Ed.), *Thought without language.* Oxford and New York: Clarendon Press and Oxford University Press.

Butterworth, G. E., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*(1), 55–72. DOI: 10.1111/j.2044-835X.1991.tb00862.x

Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, *304*(5679), 1926–1929. DOI: 10.1126/science.1099745

Buzsáki, G., & Wang, X.-J. (2012). Mechanisms of gamma oscillations. *Annual Review of Neuroscience*, *35*, 203–225. DOI: 10.1146/annurev-neuro-062111-150444

Byers-Heinlein, K., Morin-Lessard, E., & Lew-Williams, C. (2017). Bilingual infants control their languages as they listen. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(34), 9032–9037. DOI: 10.1073/pnas.1703220114

Call, J. (2001). Object permanence in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and children (*Homo sapiens*). *Journal of Comparative Psychology*, *115*(2), 159–171. DOI: 10.1037/0735-7036.115.2.159

Call, J., & Tomasello, M. (1994). Production and comprehension of referential pointing by orangutans (*Pongo pygmaeus*). *Journal of Comparative Psychology*, *108*(4), 307–317. DOI: 10.1037/0735-7036.108.4.307

Camaioni, L., Perucchini, P., Bellagamba, F., & Colonnesi, C. (2004). The role of declarative pointing in developing a theory of mind. *Infancy*, *5*(3), 291–308. DOI: 10.1207/s15327078in0503_3

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1-2), 179–213. DOI: 10.1016/S0010-0277(00)00154-2

Carpendale, J. I. M., & Carpendale, A. B. (2010). The development of pointing: From personal directedness to interpersonal direction. *Human Development*, *53*(3), 110–126. DOI: 10.1159/000315168

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *The Behavioral and brain sciences*, *27*(01). DOI: 10.1017/S0140525X04000032

Carpenter, M., & Call, J. (2013). How joint is the joint attention of apes and human infants? In J. Metcalfe & H. S. Terrace (Eds.), *Agency and joint attention* (pp. 49–61). Oxford University Press. DOI: 10.1093/acprof:oso/9780199988341.003.0003

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G. E., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i. DOI:

10.2307/1166214

Caselli, M. C., Rinaldi, P., Stefanini, S., & Volterra, V. (2012). Early action and gesture "vocabulary" and its relation with word comprehension and production. *Child Development*, *83*(2), 526–542. DOI: 10.1111/j.1467-8624.2011.01727.x

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: What have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*(3), 104–110. DOI: 10.1016/j.tics.2005.01.011

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1966). *Cartesian linguistics*. New York, NY: Harper and Row.

Chomsky, N. (1983). Mental representations. *Syracuse Scholar (1979-1991)*, *4*(2), 2.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620539

Clarke, R. J., & Ikeda, H. (1985). Luminance and darkness detectors in the olivary and posterior pretectal nuclei and their relationship to the pupillary light reflex in the rat. *Experimental Brain Research*, *57*(2), 224–232. DOI: 10.1007/BF00236527

Clifton, R. K., Muir, D. W., Ashmead, D. H., & Clarkson, M. G. (1993). Is visually guided reaching in early infancy a myth? *Child Development*, *64*(4), 1099–1110. DOI: 10.1111/j.1467-8624.1993.tb04189.x

Clifton, R. K., Perris, E. E., & McCall, D. D. (1999). Does reaching in the dark for unseen objects reflect representation in infants? *Infant Behavior and Development*, *22*(3), 297–302. DOI: 10.1016/S0163-6383(99)00017-X

Cohen, L. B., & Cashon, C. H. (2006). Infant perception and cognition. In I. B. Weiner (Ed.), *Handbook of psychology* (Vol. 34, p. 1691). Hoboken, NJ: Wiley. DOI: 10.1002/0471264385.wei0603

Cohen, L. B., & Marks, K. S. (2002). How infants process addition and subtraction events. *Developmental Science*, *5*(2), 186–201. DOI: 10.1111/1467-7687.00220

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. Cambridge, MA: MIT Press.

Colombo, J. (2002). Infant attention grows up: The emergence of a developmental cognitive neuroscience perspective. *Current directions in psychological science*, *11*(6), 196–200. DOI: 10.1111/1467-8721.00199

Colombo, J., McCardle, P. D., & Freund, L. (Eds.). (2009). *Infant pathways to language: Methods, models, and research disorders*. New York, NY: Psychology Press.

Colonnesi, C., Stams, G. J. J., Koster, I., & Noom, M. J. (2010). The relation between pointing and language development: A meta-analysis. *Developmental Review*, *30*(4), 352–366. DOI: 10.1016/j.dr.2010.10.001

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, *58*(3), 306–324. DOI: 10.1016/j.neuron.2008.04.017

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology*, *34*(1), 28–38. DOI: 10.1037/0012-1649.34.1.28

Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science*, *11*(2), 239–264. DOI: 10.1177/1745691615621279

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1431), 447–458. DOI: 10.1098/rstb.2002.1235

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, *25*(2), 141–168. DOI: 10.1111/j.1468-0017.2009.01384.x

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. DOI: 10.1126/science.290.5496.1582

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. DOI: 10.1016/j.tics.2009.01.005

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *366*(1567), 1149–1157. DOI: 10.1098/rstb.2010.0319

Csibra, G., Tucker, L. A., & Johnson, M. H. (1998). Neural correlates of saccade planning in infants: A high-density ERP study. *International Journal of Psychophysiology*, *29*(2), 201–215. DOI: 10.1016/S0167-8760(98)00016-6

Csibra, G., Tucker, L. A., Volein, A., & Johnson, M. H. (2000). Cortical development and saccade planning: The ontogeny of the spike potential. *Neuroreport*, *11*(5), 1069–1073. DOI: 10.1097/00001756-200004070-00033

Csibra, G., & Volein, A. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology*, *26*(1), 1–11. DOI: 10.1348/026151007X185987

DeBoer, T., Scott, L. S., & Nelson, C. A. (2007). Methods for acquiring and analyzing infant event-related potentials. In M. de Haan (Ed.), *Infant EEG and event-related potentials* (Vol. 500, pp. 5–37). Hove, England, and New York, NY: Psychology Press.

DeCasper, A., & Fifer, W. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, *208*(4448), 1174–1176. DOI: 10.1126/science.7375928

Dennett, D. C. (1996). *The intentional stance* (6. printing ed.). Cambridge, MA: MIT Press.

D'Entremont, B., Hains, S., & Muir, D. (1997). A demonstration of gaze following in 3- to 6-month-olds. *Infant Behavior and Development*, *20*(4), 569–572. DOI: 10.1016/S0163-6383(97)90048-5

Diamond, A. (1990). Developmental time course in human infants and infant monkeys, and the neural bases of, inhibitory control in reaching. *Annals of the New York Academy of Sciences*, *608*(1), 637–676. DOI: 10.1111/j.1749-6632.1990.tb48913.x

Dörrenberg, S., Rakoczy, H., & Liszkowski, U. (2018). How (not) to measure infant theory of mind: Testing the replicability and validity of four non-verbal measures. *Cognitive Development*. DOI: 10.1016/j.cogdev.2018.01.001

Dunn, K., & Bremner, J. G. (2017). Investigating looking and social looking measures as an index of infant violation of expectation. *Developmental Science*, *20*(6). DOI: 10.1111/desc.12452

E. Herrmann, & Tomasello, M. (2006). Apes' and children's understanding of cooperative and competitive motives in a communicative situation. *Developmental Science*, *9*(5), 518–529. DOI: 10.1111/j.1467-7687.2006.00519.x

E. L. Kaufman, Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*(4), 498. DOI: 10.2307/1418556

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience*. DOI: 10.1016/j.dcn.2016.11.001

Einhäuser, W. (2017). The pupil as marker of cognitive processes. In Q. Zhao (Ed.), *Computational and cognitive neuroscience of vision* (pp. 141–169). Singapore: Springer Singapore. DOI: 10.1007/978-981-10-0213-7_7

Ellis, C. J. (1981). The pupillary light reflex in normal subjects. *British Journal of Ophthalmology*, *65*(11), 754–759. DOI: 10.1136/bjo.65.11.754

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–9605. DOI: 10.1073/pnas.152159999

Farroni, T., Mansfield, E. M., Lai, C., & Johnson, M. H. (2003). Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology*, *85*(3), 199–212. DOI: 10.1016/S0022-0965(03)00022-5

Fawcett, C., Arslan, M., Falck-Ytter, T., Roeyers, H., & Gredebäck, G. (2017). Human eyes with dilated pupils induce pupillary contagion in infants. *Scientific Reports*, *7*(1), 9601. DOI: 10.1038/s41598-017-08223-3

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science*, *6*(5), 568–584. DOI: 10.1111/1467-7687.00313

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, *97*(3), 295–313. DOI: 10.1016/j.cognition.2004.09.010

Fernald, A. E., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics* (Vol. 44, pp. 97–135). Amsterdam and Philadelphia: John Benjamins Pub. DOI: 10.1075/lald.44.06fer

Fitzgerald, H. E. (1968). Autonomic pupillary reflex activity during early infancy and its relation to social and nonsocial visual stimuli. *Journal of Experimental Child Psychology*, *6*(3), 470–482. DOI: 10.1016/0022-0965(68)90127-6

Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT Press.

Franco, F., & Butterworth, G. (1996). Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language*, *23*(2), 307–336. DOI: 10.1017/S0305000900008813

Friese, U., Köster, M., Hassler, U., Martens, U., Trujillo-Barreto, N., & Gruber, T. (2013). Successful memory encoding is associated with increased cross-frequency coupling between frontal theta and posterior gamma oscillations in human scalp-recorded EEG. *NeuroImage*, *66*, 642–647. DOI: 10.1016/j.neuroimage.2012.11.002

Gandhi, N. J., & Katnani, H. A. (2011). Motor functions of the superior colliculus. *Annual Review of Neuroscience*, *34*, 205–231. DOI: `10.1146/annurev-neuro-061010-113728`

Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant pupil diameter changes in response to others' positive and negative emotions. *PLoS ONE*, *6*(11), e27132. DOI: `10.1371/journal.pone.0027132`

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective & Behavioral Neuroscience*, *10*(2), 252–269. DOI: `10.3758/CABN.10.2.252`

Glick, J. (2013). Recovering the organismic: Werner and Kaplan's "Symbol formation" 50 years later. *Culture & Psychology*, *19*(4), 441–452. DOI: `10.1177/1354067X13500327`

Gliga, T., & Csibra, G. (2009). One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*, *20*(3), 347–353. DOI: `10.1111/j.1467-9280.2009.02295.x`

Gliga, T., Volein, A., & Csibra, G. (2010). Verbal labels modulate perceptual object processing in 1-year-old children. *Journal of Cognitive Neuroscience*, *22*(12), 2781–2789. DOI: `10.1162/jocn.2010.21427`

Golinkoff, R. M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on Psychological Science*, *8*(3), 316–339. DOI: `10.1177/1745691613484936`

Gontier, N. (2013). Pointing and the evolution of language: An applied evolutionary epistemological approach. In: Pointing: Where embodied cognition meets the symbolic mind. *Humana. Mente: Journal of Philosophical Studies*, *24*, 1–26.

Gooley, J. J., Ho Mien, I., St Hilaire, M. A., Yeo, S.-C., Chua, E. C.-P., van Reen, E., . . . Lockley, S. W. (2012). Melanopsin and rod-cone photoreceptors play different roles in mediating pupillary light responses during exposure to continuous light in humans. *Journal of Neuroscience*, *32*(41), 14242–14253. DOI: 10.1523/JNEUROSCI.1321-12.2012

Gopnik, A. (1984). The acquisition of gone and the development of the object concept. *Journal of Child Language*, *11*(02), 473. DOI: 10.1017/S0305000900005778

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544–549.

Gredebäck, G., Eriksson, M., Schmitow, C., Laeng, B., & Stenberg, G. (2012). Individual differences in face processing: Infants' scanning patterns and pupil dilations are influenced by the distribution of parental leave. *Infancy*, *17*(1), 79–101. DOI: 10.1111/j.1532-7078.2011.00091.x

Gredebäck, G., Johnson, S. P., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*(1), 1–19. DOI: 10.1080/87565640903325758

Gredebäck, G., & Melinder, A. M. D. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition*, *114*(2), 197–206. DOI: 10.1016/j.cognition.2009.09.004

Gredebäck, G., Melinder, A. M. D., & Daum, M. M. (2010). The development and neural basis of pointing comprehension. *Social Neuroscience*, *5*(5-6), 441–450. DOI: 10.1080/17470910903523327

Grice, H. P. (1957). Meaning. *The Philoshophical Review*, *66*(3), 377–388. DOI: 10.2307/2182440

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Academic Press.

Grossmann, T., & Johnson, M. H. (2010). Selective prefrontal cortex responses to joint attention in early infancy. *Biology Letters*, *6*(4), 540–543. DOI: 10.1098/rsbl.2009.1069

Grossmann, T., Johnson, M. H., Farroni, T., & Csibra, G. (2007). Social perception in the infant brain: Gamma oscillatory activity in response to eye gaze. *Social Cognitive and Affective Neuroscience*, *2*(4), 284–291. DOI: 10.1093/scan/nsm025

Grossmann, T., Parise, E., & Friederici, A. D. (2010). The detection of communicative signals directed at the self in infant prefrontal cortex. *Frontiers in human neuroscience*, *4*, 201. DOI: 10.3389/fnhum.2010.00201

Gruber, T., Müller, M. M., & Keil, A. (2002). Modulation of induced gamma band responses in a perceptual learning task in the human EEG. *Journal of Cognitive Neuroscience*, *14*(5), 732–744. DOI: 10.1162/08989290260138636

de Haan, M. (Ed.). (2007). *Infant eeg and event-related potentials*. Hove, England and New York, NY: Psychology Press.

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, *21*(2), 167–179. DOI: 10.1016/S0163-6383(98)90001-7

Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (2011). Collaboration encourages equal sharing in children but not in chimpanzees. *Nature*, *476*(7360), 328–331. DOI: 10.1038/nature10278

Hare, B., & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, *68*(3), 571–581. DOI: 10.1016/j.anbehav.2003.11.011

Harris, C. M., Hainline, L., Abramov, I., Lemerise, E., & Camenzuli, C. (1988). The distribution of fixation durations in infants and naive adults. *Vision Research*, *28*(3), 419–432. DOI: 10.1016/0042-6989(88)90184-8

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York, NY: Wiley & Sons.

Henderson, L. M., Yoder, P. J., Yale, M. E., & McDuffie, A. (2002). Getting the point: Electrophysiological correlates of protodeclarative pointing. *International Journal of Developmental Neuroscience*, *20*(3-5), 449–458. DOI: 10.1016/S0736-5748(02)00038-2

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2-3), 61-83; discussion 83-135. DOI: 10.1017/S0140525X0999152X

Hepach, R., Vaish, A., Grossmann, T., & Tomasello, M. (2016). Young children want to see others get the help they need. *Child Development*, *87*(6), 1703–1714. DOI: 10.1111/cdev.12633

Hepach, R., Vaish, A., & Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychological Science*, *23*(9), 967–972. DOI: 10.1177/0956797612440571

Hepach, R., & Westermann, G. (2013). Infants' sensitivity to the congruence of others' emotions and actions. *Journal of Experimental Child Psychology*, *115*(1), 16–29. DOI: 10.1016/j.jecp.2012.12.013

Hepach, R., & Westermann, G. (2016). Pupillometry in infancy research. *Journal of Cognition and Development*, *17*(3), 359–377. DOI: 10.1080/15248372.2015.1135801

Herrmann, C. S., Grigutsch, M., & Busch, N. A. (2005). EEG oscillations and wavelet analysis. In T. C. Handy (Ed.), *Event-related potentials*. Cambridge, Mass: MIT Press.

Herrmann, C. S., & Mecklinger, A. (2000). Magnetoencephalographic responses to illusory figures: Early evoked gamma is affected by processing of stimulus features. *International Journal of Psychophysiology*, *38*(3), 265–281. DOI: 10.1016/S0167-8760(00)00170-7

Herrmann, C. S., Munk, M. H., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: Memory match and utilization. *Trends in Cognitive Sciences*, *8*(8), 347–355. DOI: 10.1016/j.tics.2004.06.006

Hespos, S. J., & Baillargeon, R. (2001). Reasoning about containment events in very young infants. *Cognition*, *78*(3), 207–245. DOI: 10.1016/S0010-0277(00)00118-9

Hespos, S. J., Gredebäck, G., von Hofsten, C., & Spelke, E. S. (2009). Occlusion is hard: Comparing predictive reaching for visible and hidden objects in infants and adults. *Cognitive Science*, *33*(8), 1483–1502. DOI: 10.1111/j.1551-6709.2009.01051.x

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science*, *132*(3423), 349–350. DOI: 10.1126/science.132.3423.349

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190–1192. DOI: 10.1126/science.143.3611.1190

Heyes, C. M. (2014a). False belief in infancy: A fresh look. *Developmental Science*, *17*(5), 647–659. DOI: 10.1111/desc.12148

Heyes, C. M. (2014b). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, *9*(2), 131–143. DOI: 10.1177/1745691613518076

Hilton, S. M., & Smith, P. R. (1984). Ventral medullary neurones excited from the hypothalamic and mid-brain defence areas. *Journal of the Autonomic Nervous System*, *11*(1), 35–42. DOI: 10.1016/0165-1838(84)90006-7

Hochmann, J.-R., & Papeo, L. (2014). The invariance problem in infancy: A pupillometry study. *Psychological Science*, *25*(11), 2038–2046. DOI: 10.1177/0956797614547918

Hoehl, S., Michel, C., Reid, V. M., Parise, E., & Striano, T. (2014). Eye contact during live social interaction modulates infants' oscillatory brain activity. *Social Neuroscience*, *9*(3), 300–308. DOI: 10.1080/17470919.2014.884982

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Develop-*

mental Neuropsychology, *37*(3), 187–209. DOI: 10.1080/87565641.2011.627958

Hoehl, S., Wahl, S., & Pauen, S. (2014). Disentangling the effects of an adult model's eye gaze and head orientation on young infants' processing of a previously attended object. *Infancy*, *19*(1), 53–64. DOI: 10.1111/infa.12035

Holmqvist, K. (2015). *Eye tracking: A comprehensive guide to methods and measures.* Oxford, England: Oxford University Press.

Hood, B., & Willatts, P. (1986). Reaching in the dark to an object's remembered position: Evidence for object permanence in 5-month-old infants. *British Journal of Developmental Psychology*, *4*(1), 57–65. DOI: 10.1111/j.2044-835X.1986.tb00998.x

Hublin, J.-J., Ben-Ncer, A., Bailey, S. E., Freidline, S. E., Neubauer, S., Skinner, M. M., ... Gunz, P. (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*. *Nature*, *546*(7657), 289–292. DOI: 10.1038/nature22336

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69–95.

Itakura, S. (2004). Gaze-following and joint visual attention in nonhuman animals. *Japanese Psychological Research*, *46*(3), 216–226. DOI: 10.1111/j.1468-5584.2004.00253.x

Jackson, I. R., & Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Developmental Science*, *12*(4), 670–679. DOI: 10.1111/j.1467-7687.2008.00805.x

Jessen, S., Altvater-Mackensen, N., & Grossmann, T. (2016). Pupillary responses reveal infants' discrimination of facial emotions independent of conscious perception. *Cognition*, *150*, 163–169. DOI: 10.1016/j.cognition.2016.02.010

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*, *89*(1), 221–234. DOI: 10.1016/j.neuron.2015.11.028

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585. DOI: 10.1126/science.154.3756.1583

Kahneman, D., & Beatty, J. (1967). Pupillary responses in a pitch-discrimination task. *Perception & Psychophysics*, *2*(3), 101–105. DOI: 10.3758/BF03210302

Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during a mental task. *Science*, *157*(3785), 218–219. DOI: 10.1126/science.157.3785.218

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219. DOI: 10.1016/0010-0285(92)90007-O

Kahneman, D., & Wright, P. (1971). Changes of pupil size and rehearsal strategies in a short-term memory task. *The Quarterly Journal of Experimental Psychology*, *23*(2), 187–196. DOI: 10.1080/14640747108400239

Káldy, Z., & Sigala, N. (2004). The neural mechanisms of object working memory: What is where in the infant brain? *Neuroscience and biobehavioral reviews*, *28*(2), 113–121. DOI: 10.1016/j.neubiorev.2004.01.002

Kampe, K. K. W., Frith, C. D., & Frith, U. (2003). "Hey John": Signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *Journal of Neuroscience*, *23*(12), 5258–5263. DOI: 10.1523/JNEUROSCI.23-12-05258.2003

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. *Proceedings of the Royal Society of London B: Biological Sciences*, *282*(1819). DOI: 10.1098/rspb.2015.1683

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2016). On potential ocular artefacts in infant electroencephalogram: A reply to comments by Köster. *Proceedings of the Royal Society of London B: Biological Sciences*, *283*(1835). DOI: 10.1098/

rspb.2016.1285

Katzner, K. (2011). *The languages of the world.* London, England; New York, NY: Routledge.

Kaufman, J., Csibra, G., & Johnson, M. H. (2003). Representing occluded objects in the human infant brain. *Proceedings of the Royal Society B: Biological Sciences*, *270*(Suppl_2), S140–S143. DOI: 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 of the United States of America*, *102*(42), 15271–15274. DOI: 10.1073/pnas.0507626102

Keren, A. S., Yuval-Greenberg, S., & Deouell, L. Y. (2010). Saccadic spike potentials in gamma-band EEG: Characterization, detection and suppression. *NeuroImage*, *49*(3), 2248–2263. DOI: 10.1016/j.neuroimage.2009.10.057

Kirchhofer, K. C., Zimmermann, F., Kaminski, J., & Tomasello, M. (2012). Dogs (*Canis familiaris*), but not chimpanzees (*Pan troglodytes*), understand imperative pointing. *PLoS ONE*, *7*(2), e30913. DOI: 10.1371/journal.pone.0030913

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*(1), 78–100. DOI: 10.1037/0033-2909.100.1.78

Kourouyan, H. D., & Horton, J. C. (1997). Transneuronal retinal input to the primate edinger-westphal nucleus. *Journal of Comparative Neurology*, *381*(1), 68–80. DOI: 10.1002/(SICI)1096-9861(19970428)381:1<68::AID-CNE6>3.0.CO;2-I

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*(6012), 1830–1834. DOI: 10.1126/science.1190792

Kozicz, T., Bittencourt, J. C., May, P. J., Reiner, A., Gamlin, P. D. R., Palkovits, M., . . . Ryabinin, A. E. (2011). The Edinger-Westphal nucleus: A historical, structural, and functional perspective on a dichotomous terminology. *Journal of Comparative*

Neurology, *519*(8), 1413–1434. DOI: 10.1002/cne.22580

Krauzlis, R. J., Lovejoy, L. P., & Zénon, A. (2013). Superior colliculus and visual spatial attention. *Annual Review of Neuroscience*, *36*, 165–182. DOI: 10.1146/annurev-neuro-062012-170249

Krehm, M., Onishi, K. H., & Vouloumanos, A. (2014). I see your point: Infants under 12 months understand that pointing is communicative. *Journal of Cognition and Development*, *15*(4), 527–538. DOI: 10.1080/15248372.2012.736112

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. DOI: 10.1177/1745691611427305

Laming, D. (2008). Weber's law. In P. Rabbitt (Ed.), *Inside psychology: A science over 50 years.* Oxford University Press.

Larsen, R. S., & Waters, J. (2018). Neuromodulatory correlates of pupil dilation. *Frontiers in neural circuits*, *12*, 21. DOI: 10.3389/fncir.2018.00021

Leavens, D. A. (2004). Book review of S. Kita (Ed.), Pointing: Where language, culture and cognition meet, Lawrence Erlbaum Associates, 2003. *Cognitive Systems Research*(5), 157–165. DOI: 10.1016/j.cogsys.2004.01.002

Leavens, D. A. (2012). Joint attention: Twelve myths. In A. Seemann (Ed.), *Joint attention* (pp. 43–72). Cambridge, MA and London, England: MIT Press.

Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception*, *13*(3), 287–305. DOI: 10.1068/p130287

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*(3), 265–288. DOI: 10.1016/S0010-0277(87)80006-9

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing 'what' and 'where' systems. *Trends in Cognitive Sciences*, *2*(1), 10–18. DOI: 10.1016/S1364-6613(97)01113-3

Leung, E. H., & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology*, *17*(2), 215–220. DOI: `10.1037/0012-1649.17.2.215`

Leung, S., Mareschal, D., Rowsell, R., Simpson, D., Iaria, L., Grbic, A., & Kaufman, J. (2016). Oscillatory activity in the infant brain and the representation of small numbers. *Frontiers in Systems Neuroscience*, *10*, 4. DOI: `10.3389/fnsys.2016.00004`

Levy, E. T., & McNeill, D. (2013). Narrative development as symbol formation: Gestures, imagery and the emergence of cohesion. *Culture & Psychology*, *19*(4), 548–569. DOI: `10.1177/1354067X13500328`

Liebal, K., Carpenter, M., & Tomasello, M. (2010). Infants' use of shared experience in declarative pointing. *Infancy*, *15*(5), 545–556. DOI: `10.1111/j.1532-7078.2009.00028.x`

Liszkowski, U. (2006). Infant pointing at twelve months: Communicative goals, motives, and social-cognitive abilities. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality* (pp. 153–178). New York, NY: Berg.

Liszkowski, U. (2011). Three lines in the emergence of prelinguistic communication and social cognition. *Journal of Cognitive Education and Psychology*, *10*(1), 32–43. DOI: `10.1891/1945-8959.10.1.32`

Liszkowski, U. (2017). Emergence of shared reference and shared minds in infancy. *Current Opinion in Psychology*, *23*, 26–29. DOI: `10.1016/j.copsyc.2017.11.003`

Liszkowski, U., Brown, P., Callaghan, T., Takada, A., & de Vos, C. (2012). A prelinguistic gestural universal of human communication. *Cognitive Science*, *36*(4), 698–713. DOI: `10.1111/j.1551-6709.2011.01228.x`

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science*, *7*(3), 297–307. DOI: `10.1111/j.1467-7687.2004.00349.x`

Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development*, *7*(2), 173–187. DOI: 10.1207/s15327647jcd0702_2

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007a). Pointing out new news, old news, and absent referents at 12 months of age. *Developmental Science*, *10*(2), F1–F7. DOI: 10.1111/j.1467-7687.2006.00552.x

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007b). Reference and attitude in infant pointing. *Journal of Child Language*, *34*(1), 1–20. DOI: 10.1017/S0305000906007689

Liszkowski, U., & Ramenzoni, V. C. (2015). Pointing to nothing? Empty places prime infants' attention to absent objects. *Infancy*, *20*(4), 433-444. DOI: 10.1111/infa.12080

Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (2009). Prelinguistic infants, but not chimpanzees, communicate about absent entities. *Psychological Science*, *20*(5), 654–660. DOI: 10.1111/j.1467-9280.2009.02346.x

Liszkowski, U., & Tomasello, M. (2011). Individual differences in social, cognitive, and morphological aspects of infant pointing. *Cognitive Development*, *26*(1), 16–29. DOI: 10.1016/j.cogdev.2010.10.001

Lœvenbruck, H., Dohen, M., & Vilain, C. E. (2009). Pointing is 'special'. In S. Fuchs, H. Lœvenbruck, D. Pape, & P. Perrier (Eds.), *Some aspects of speech and the brain* (pp. 211–258). Peter Lang.

Loewenfeld, I. E., & Lowenstein, O. (1993). *The pupil: Anatomy, physiology, and clinical applications*. Wiley-Blackwell.

Loewy, A. D., Wallach, J. H., & McKellar, S. (1981). Efferent connections of the ventral medulla oblongata in the rat. *Brain Research Reviews*, *3*(1), 63–80. DOI: 10.1016/0165-0173(81)90012-6

Lum, J. A. G., Youssef, G. J., & Clark, G. M. (2017). Using pupillometry to investigate sentence comprehension in children with and without specific language impairment. *Journal of speech, language, and hearing research : JSLHR*, *60*(6), 1648–1660. DOI: 10.1044/2017_JSLHR-L-16-0158

Lutzenberger, W., Pulvermüller, F., Elbert, T., & Birbaumer, N. (1995). Visual stimulation alters local 40-hz responses in humans: An EEG-study. *Neuroscience Letters*, *183*(1-2), 39–42. DOI: 10.1016/0304-3940(94)11109-V

Mandler, J. M. (1998). Representation. In D. Kuhn & R. S. Siegler (Eds.), *Cognition, perception, and language: Handbook of child psychology 5th ed.* New York, NY: Wiley.

Marcovitch, S., & Zelazo, P. D. (1999). The A-Not-B error: Results from a logistic meta-analysis. *Child Development*, *70*(6), 1297–1313. DOI: 10.1111/1467-8624.00095

Mareschal, D., & Johnson, M. H. (2003). The "what" and "where" of object representations in infancy. *Cognition*, *88*(3), 259–276. DOI: 10.1016/S0010-0277(03)00039-8

Marino, A. C., & Scholl, B. J. (2005). The role of closure in defining the "objects" of object-based attention. *Perception & Psychophysics*, *67*(7), 1140–1149. DOI: 10.3758/BF03193547

Marno, H., Davelaar, E. J., & Csibra, G. (2014). Nonverbal communicative signals modulate attention to object properties. *Journal of Experimental Psychology. Human Perception and Performance*, *40*(2), 752–762. DOI: 10.1037/a0035113

Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, *1*(1), 1289. DOI: 10.5334/joc.18

Mathôt, S., Grainger, J., & Strijkers, K. (2017). Pupillary responses to words that convey a sense of brightness or darkness. *Psychological Science*, *28*(8), 1116–1124. DOI: 10.1177/0956797617702699

Mathôt, S., & van der Stigchel, S. (2015). New light on the mind's eye. *Current Directions* *in Psychological Science*, *24*(5), 374–378. DOI: 10.1177/0963721415593725

Matthews, D., Behne, T., Lieven, E., & Tomasello, M. (2012). Origins of the human pointing gesture: A training study. *Developmental Science*, *15*(6), 817–829. DOI: 10.1111/j.1467-7687.2012.01181.x

McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, *15*(11), 776–781. DOI: 10.1111/j.0956-7976.2004.00755.x

McDougal, D. H., & Gamlin, P. D. (2010). The influence of intrinsically-photosensitive retinal ganglion cells on the spectral sensitivity and response dynamics of the human pupillary light reflex. *Vision research*, *50*(1), 72–87. DOI: 10.1016/j.visres.2009.10.012

McDougal, D. H., & Gamlin, P. D. (2015). Autonomic control of the eye. *Comprehensive Physiology*, *5*(1), 439–473. DOI: 10.1002/cphy.c140014

McGillion, M., Pine, J. M., Herbert, J. S., & Matthews, D. (2017). A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse socioeconomic status backgrounds. *Journal of Child Psychology and Psychiatry*, *58*(10), 1122–1131. DOI: 10.1111/jcpp.12725

van der Meer, E., Beyer, R., Horn, J., Foth, M., Bornemann, B., Ries, J., . . . Wartenburger, I. (2010). Resource allocation and fluid intelligence: Insights from pupillometry. *Psychophysiology*, *47*(1), 158–169. DOI: 10.1111/j.1469-8986.2009.00884.x

Meltzoff, A. N. (2007). 'Like me': A foundation for social cognition. *Developmental Science*, *10*(1), 126–134. DOI: 10.1111/j.1467-7687.2007.00574.x

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*(4312), 75–78. DOI: 10.1126/science.198.4312.75

Meltzoff, A. N., & Moore, M. K. (1998). Object representation, identity, and the paradox of early permanence: Steps toward a new framework. *Infant Behavior and Development*, *21*(2), 201–235. DOI: 10.1016/S0163-6383(98)90003-0

Merritt, D. J., & Brannon, E. M. (2013). Nothing to it: Precursors to a zero concept in preschoolers. *Behavioural Processes*, *93*, 91–97. DOI: 10.1016/j.beproc.2012.11.001

Michel, C., Wronski, C., Pauen, S., Daum, M. M., & Hoehl, S. (2017). Infants' object processing is guided specifically by social cues. *Neuropsychologia*. DOI: 10.1016/j.neuropsychologia.2017.05.022

Miller, D. J., Cohen, L. B., & Hill, K. T. (1970). A methodological investigation of Piaget's theory of object concept development in the sensory-motor period. *Journal of Experimental Child Psychology*, *9*(1), 59–85. DOI: 10.1016/0022-0965(70)90101-3

Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science*, *7*(1), F1-F9. DOI: 10.1111/j.1467-7687.2004.00315.x

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, *16*(4), 495–500. DOI: 10.1016/0163-6383(93)80007-U

Moore, C. (1996). Theories of mind in infancy. *British Journal of Developmental Psychology*, *14*(1), 19–40. DOI: 10.1111/j.2044-835X.1996.tb00691.x

Moore, C. (1999). Intentional relations and triadic interactions. In P. D. Zelazo & Astington, J. W., Olson, D. R. (Eds.), *Developing theories of intention: Social understanding and self-control* (pp. 43–61). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Moore, C., & Corkum, V. (1994). Social understanding at the end of the first year of life. *Developmental Review*, *14*(4), 349–372. DOI: 10.1006/drev.1994.1014

Moore, C., & Corkum, V. (1998). Infant gaze following based on eye direction. *British*

Journal of Developmental Psychology, *16*(4), 495–503. DOI: 10.1111/j.2044-835X.1998.tb00767.x

Moore, C., & D'Entremont, B. (2001). Developmental changes in pointing as a function of attentional focus. *Journal of Cognition and Development*, *2*(2), 109–129. DOI: 10.1207/S15327647JCD0202{_}1

Moore, M. K., Borton, R., & Darby, B. L. (1978). Visual tracking in young infants: Evidence for object identity or object permanence? *Journal of Experimental Child Psychology*, *25*(2), 183–198. DOI: 10.1016/0022-0965(78)90076-0

Moore, R. (2013). Evidence and interpretation in great ape gestural communication. In: Pointing: Where embodied cognition meets the symbolic mind. *Humana. Mente: Journal of Philosophical Studies*, *24*, 27–51.

Moore, R., Call, J., & Tomasello, M. (2015). Production and comprehension of gestures between orang-utans (*Pongo pygmaeus*) in a referential communication game. *PLoS ONE*, *10*(6), e0129726. DOI: 10.1371/journal.pone.0129726

Mou, Y., & vanMarle, K. (2014). Two core systems of numerical representation in infants. *Developmental Review*, *34*(1), 1–25. DOI: 10.1016/j.dr.2013.11.001

Müller, M. M., Bosch, J., Elbert, T., Kreiter, A., Sosa, M. V., Sosa, P. V., & Rockstroh, B. (1996). Visually induced gamma-band responses in human electroencephalographic activity — a link to animal studies. *Experimental Brain Research*, *112*(1), 96–102. DOI: 10.1007/BF00227182

Munakata, Y. (1997). Perseverative reaching in infancy: The roles of hidden toys and motor history in the AB task. *Infant Behavior and Development*, *20*(3), 405–416. DOI: 10.1016/S0163-6383(97)90011-4

Mundy, P., Card, J., & Fox, N. (2000). EEG correlates of the development of infant joint attention skills. *Developmental Psychobiology*, *36*(4), 325–338. DOI: 10.1002/(SICI)1098-2302(200005)36:4<325::AID-DEV7>3.0.CO;2-F

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, *16*(5), 269–274. DOI: 10.1111/j.1467-8721.2007.00518.x

Murphy, M., & Messer, D. J. (1977). Mothers, infants and pointing: A study of a gesture. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 325–354). London, England: Academic Press.

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with bold activity in human locus coeruleus. *Human Brain Mapping*, *35*(8), 4140–4154. DOI: 10.1002/hbm.22466

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology*, *48*(11), 1532–1543. DOI: 10.1111/j.1469-8986.2011.01226.x

Naber, M., & Nakayama, K. (2013). Pupil responses to high-level image content. *Journal of Vision*, *13*(6). DOI: 10.1167/13.6.7

Nieder, A. (2016). Representing something out of nothing: The dawning of zero. *Trends in Cognitive Sciences*, *20*(11), 830–842. DOI: 10.1016/j.tics.2016.08.008

Noles, N. S., Scholl, B. J., & Mitroff, S. R. (2005). The persistence of object file representations. *Perception & Psychophysics*, *67*(2), 324–334. DOI: 10.3758/BF03206495

Oakes, L. M. (2010). Using habituation of looking time to assess mental processes in infancy. *Journal of Cognition and Development*, *11*(3), 255–268. DOI: 10.1080/15248371003699977

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 255–258. DOI: 10.1126/science.1107621

Pätzold, W., & Liszkowski, U. (2018). *Pupillometry reveals communication-induced object expectations in 12- but not 8-month-old infants.* Manuscript submitted for publication.

Pelé, M., Dufour, V., Thierry, B., & Call, J. (2009). Token transfers among great apes (*Gorilla gorilla*, *Pongo pygmaeus*, *Pan paniscus*, and *Pan troglodytes*): Species differences, gestural requests, and reciprocal exchange. *Journal of Comparative Psychology*, *123*(4), 375–384. DOI: 10.1037/a0017253

Penttonen, M., & Buzsáki, G. (2003). Natural logarithmic relationship between brain oscillators. *Thalamus and Related Systems*, *2*(02), 145. DOI: 10.1017/S1472928803000074

Perner, J. (1991). *Learning, development, and conceptual change. understanding the representational mind.* Cambridge, MA: MIT Press.

Perner, J., Huemer, M., & Leahy, B. (2015). Mental files and belief: A cognitive theory of how children represent belief and its intensionality. *Cognition*, *145*, 77–88. DOI: 10.1016/j.cognition.2015.08.006

Perreault, C., & Mathew, S. (2012). Dating the origin of language using phonemic diversity. *PLoS ONE*, *7*(4), e35289. DOI: 10.1371/journal.pone.0035289

Phillips, J., Ong, D. C., Surtees, A. D. R., Xin, Y., Williams, S., Saxe, R., & Frank, M. C. (2015). A second look at automatic theory of mind: Reconsidering Kovács, Téglás, and Endress (2010). *Psychological Science*, *26*(9), 1353–1367. DOI: 10.1177/0956797614558717

Piaget, J. (1954). *The construction of reality in the child.* New York, NY: Basic Books.

Pika, S., & Mitani, J. C. (2006). Referential gestural communication in wild chimpanzees (*Pan troglodytes*). *Current Biology*, *16*(6), R191-2. DOI: 10.1016/j.cub.2006.02.037

Pika, S., & Mitani, J. C. (2010). The directed scratch: Evidence for a referential gesture in chimpanzees. In C. Knight & R. P. Botha (Eds.), *The prehistory of language* (Vol. 1, pp. 166–181). Oxford, England: Oxford University Press.

Powell, B. B. (2009). *Writing: Theory and history of the technology of civilization.* John

Wiley & Sons.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*(1-2), 127–158. DOI: 10.1016/S0010-0277(00)00156-6

Range, F., & Virányi, Z. (2011). Development of gaze following abilities in wolves (*Canis lupus*). *PloS one*, *6*(2), e16888. DOI: 10.1371/journal.pone.0016888

Reid, V. M., Csibra, G., Belsky, J., & Johnson, M. H. (2007). Neural correlates of the perception of goal-directed action in infants. *Acta Psychologica*, *124*(1), 129–138. DOI: 10.1016/j.actpsy.2006.09.010

Reid, V. M., & Striano, T. (2005). Adult gaze influences infant attention and object processing: Implications for cognitive neuroscience. *European Journal of Neuroscience*, *21*(6), 1763–1766. DOI: 10.1111/j.1460-9568.2005.03986.x

Reid, V. M., & Striano, T. (2007). The directed attention model of infant social cognition. *European Journal of Developmental Psychology*, *4*(1), 100–110. DOI: 10.1080/17405620601005648

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. *Neuroreport*, *15*(16), 2553–2555. DOI: 10.1097/00001756-200411150-00025

Reiner, A., Fitzgerald, M. E. C., & Li, C. (2012). Neural control of ocular blood flow. In L. Schmetterer & J. Kiel (Eds.), *Ocular blood flow* (pp. 243–309). Springer Science & Business Media. DOI: 10.1007/978-3-540-69469-4_12

Renner, L. F., & Włodarczak, M. (2017). When a dog is a cat and how it changes your pupil size: Pupil dilation in response to information mismatch. In *Proc. Interspeech 2017* (pp. 674–678). DOI: 10.21437/Interspeech.2017-353

Richer, F., & Beatty, J. (1987). Contrasting effects of response uncertainty on the task-evoked pupillary response and reaction time. *Psychophysiology*, *24*(3), 258–262. DOI: 10.1111/j.1469-8986.1987.tb00291.x

Rivera, S. M., Wakeley, A., & Langer, J. (1999). The drawbridge phenomenon: Representational reasoning or perceptual preference? *Developmental Psychology*, *35*(2), 427–435. DOI: 10.1037/0012-1649.35.2.427

Roach, B. J., & Mathalon, D. H. (2008). Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia Bulletin*, *34*(5), 907–926. DOI: 10.1093/schbul/sbn093

Rohlfing, K. J., Longo, M. R., & Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. *Developmental Science*, *15*(3), 426–435. DOI: 10.1111/j.1467-7687.2012.01139.x

Rolfs, M. (2009). Microsaccades: Small steps on a long way. *Vision Research*, *49*(20), 2415–2441. DOI: 10.1016/j.visres.2009.08.010

Ruffman, T., Slade, L., & Redman, J. (2005). Young infants' expectations about hidden objects. *Cognition*, *97*(2), B35-43. DOI: 10.1016/j.cognition.2005.01.007

Ruskell, G. L. (2003). Access of autonomic nerves through the optic canal, and their orbital distribution in man. *The Anatomical Record. Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology*, *275*(1), 973–978. DOI: 10.1002/ar.a.10108

Saby, J. N., & Marshall, P. J. (2012). The utility of EEG band power analysis in the study of infancy and early childhood. *Developmental Neuropsychology*, *37*(3), 253–273. DOI: 10.1080/87565641.2011.614663

Salomo, D., & Liszkowski, U. (2013). Sociocultural settings influence the emergence of prelinguistic deictic gestures. *Child Development*, *84*(4), 1296–1307. DOI: 10.1111/cdev.12026

Salwiczek, L. H., Emery, N. J., Schlinger, B., & Clayton, N. S. (2009). The development of caching and object permanence in western scrub-jays (*Aphelocoma californica*): Which emerges first? *Journal of Comparative Psychology*, *123*(3), 295–303.

Samar, V. J., Bopardikar, A., Rao, R., & Swartz, K. (1999). Wavelet analysis of neuroelectric waveforms: A conceptual tutorial. *Brain and Language*, *66*(1), 7–60. DOI: 10.1006/brln.1998.2024

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, Sarah E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1255–1266. DOI: 10.1037/a0018729

Santiesteban, I., Catmur, C., Coughlan Hopkins, S., Bird, G., & Heyes, C. (2013). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology. Human Perception and Performance*. DOI: 10.1037/a0035175

Sara, S. J., & Bouret, S. (2012). Orienting and reorienting: The locus coeruleus mediates cognition through arousal. *Neuron*, *76*(1), 130–141. DOI: 10.1016/j.neuron.2012.09.011

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266. DOI: 10.1038/253265a0

Schaffer, H. R. (1984). The child's entry into a social world. *Behavioural Development: A Series of Monographs*.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718–730. DOI: 10.1016/j.neuropsychologia.2005.07.017

Schilling, T. H. (2000). Infants' looking at possible and impossible screen rotations: The role of familiarization. *Infancy*, *1*(4), 389–402. DOI: 10.1207/S15327078IN0104_2

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1-2), 1–46. DOI: 10.1016/S0010-0277(00)00152-9

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, *80*(1-2), 159–177. DOI: 10.1016/S0010-0277(00)00157-8

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language* (Vol. 626). Cambridge, England: Cambridge University Press.

Searle, J. R. (1995). *The construction of social reality.* New York, NY: Free Press.

Sebastián-Gallés, N. (2013). Eyes wide shut: Linking brain and pupil in bilingual and monolingual toddlers. *Trends in Cognitive Sciences*, *17*(5), 197–198. DOI: 10.1016/j.tics.2013.03.002

Seifert, J. (2008). *Ereigniskorrelierte EEG-Aktivität.* Lengerich, Germany: Pabst.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, *18*(9), 668–671. DOI: 10.1016/j.cub.2008.03.059

Shatz, M., & O'Reilly, A. W. (1990). Conversational or communicative skill? A reassessment of two-year-olds' behaviour in miscommunication episodes. *Journal of Child Language*, *17*(01), 131. DOI: 10.1017/S0305000900013143

Shinskey, J. L. (2012). Disappearing décalage: Object search in light and dark at 6 months. *Infancy*, *17*(3), 272–294. DOI: 10.1111/j.1532-7078.2011.00078.x

Simon, T. J., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development*, *10*(2), 253–269. DOI: 10.1016/0885-2014(95)90011-X

Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679–692. DOI: 10.1002/wcs.1323

Sirois, S., & Jackson, I. R. (2012). Pupil dilation and object permanence in infants. *Infancy*, *17*(1), 61–78. DOI: 10.1111/j.1532-7078.2011.00096.x

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*(8), 343–348. DOI: 10.1016/S1364-6613(03)00156-6

Smith, L. B., Thelen, E., Titzer, R., & McLin, D. (1999). Knowing in the context of acting:

The task dynamics of the "A-not-B" error. *Psychological Review*, *106*(2), 235.

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*(6), 907–912. DOI: 10.1111/j.1467-7687.2009.00946.x

Southgate, V., Csibra, G., Kaufman, J., & Johnson, M. H. (2008). Distinct processing of objects and faces in the infant brain. *Journal of Cognitive Neuroscience*, *20*(4), 741–749. DOI: 10.1162/jocn.2008.20052

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, *14*(1), 29–56. DOI: 10.1207/s15516709cog1401_3

Spelke, E. S. (2000). Core knowledge. *American Psychologist*, *55*(11), 1233. DOI: 10.1037/0003-066X.55.11.1233

Spelke, E. S., Kestenbaum, R., Simons, D. J., & Wein, D. (1995). Spatiotemporal continuity, smoothness of motion and object identity in infancy. *British Journal of Developmental Psychology*, *13*(2), 113–142. DOI: 10.1111/j.2044-835X.1995.tb00669.x

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. DOI: 10.1111/j.1467-7687.2007.00569.x

St. Louis, E. K., Frey, L. C., & Britton, J. W. (Eds.). (2016). *Electroencephalography (EEG): An introductory text and atlas of normal and abnormal findings in adults, children, and infants.* Chicago, IL: American Epilepsy Society.

Steinhauer, S. R., Condray, R., & Kasparek, A. (2000). Cognitive modulation of midbrain function: Task-induced reduction of the pupillary light reflex. *International Journal of Psychophysiology*, *39*(1), 21–30. DOI: 10.1016/S0167-8760(00)00119-7

Stets, M., Stahl, D., & Reid, V. M. (2012). A meta-analysis investigating factors underlying attrition rates in infant ERP studies. *Developmental Neuropsychology*, *37*(3), 226–252. DOI: 10.1080/87565641.2012.654867

Striano, T., Chen, X., Cleveland, A., & Bradshaw, S. (2006). Joint attention social cues influence infant learning. *European Journal of Developmental Psychology*, *3*(3), 289–299. DOI: 10.1080/17405620600879779

Striano, T., Reid, V. M., & Hoehl, S. (2006). Neural mechanisms of joint attention in infancy. *European Journal of Neuroscience*, *23*(10), 2819–2823. DOI: 10.1111/j.1460-9568.2006.04822.x

Szufnarowska, J., Rohlfing, K. J., Fawcett, C., & Gredebäck, G. (2014). Is ostension any more than attention? *Scientific reports*, *4*, 5304. DOI: 10.1038/srep05304

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, *3*(4), 151–162. DOI: 10.1016/S1364-6613(99)01299-1

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. *Journal of Neuroscience*, *16*(13), 4240–4249. DOI: 10.1523/JNEUROSCI.16-13-04240.1996

Tallon-Baudry, C., Bertrand, O., Peronnet, F., & Pernier, J. (1998). Induced γ-band activity during the delay of a visual short-term memory task in humans. *Journal of Neuroscience*, *18*(11), 4244–4254. DOI: 10.1523/JNEUROSCI.18-11-04244.1998

Tamási, K., McKean, C., Gafos, A., Fritzsche, T., & Höhle, B. (2017). Pupillometry registers toddlers' sensitivity to degrees of mispronunciation. *Journal of Experimental Child Psychology*, *153*, 140–148. DOI: 10.1016/j.jecp.2016.07.014

Tempelmann, S., Kaminski, J., & Liebal, K. (2013). When apes point the finger: Three great ape species fail to use a conspecific's imperative pointing gesture. *Interaction Studies*, *14*(1), 7–23. DOI: 10.1075/is.14.1.02tem

Tomasello, M. (1998). The concept of reference in the cognitive sciences. *Pragmatics & Cognition*, *6*(1-2), 229–243. DOI: 10.1075/pc.6.1-2.12tom

Tomasello, M. (2008). *Origins of human communication.* Cambridge, MA: MIT Press.

Tomasello, M. (2014). *A natural history of human thinking.* Cambridge, MA: Harvard

University Press.

Tomasello, M., & Call, J. (2008). Assessing the validity of ape-human comparisons: A reply to Boesch (2007). *Journal of Comparative Psychology*, *122*(4), 449–452. DOI: 10.1037/0735-7036.122.4.449

Tomasello, M., Call, J., & Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development*, *68*(6), 1067–1080. DOI: 10.1111/j.1467-8624.1997.tb01985.x

Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, *55*(4), 1063–1069. DOI: 10.1006/anbe.1997.0636

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *The Behavioral and Brain Sciences*, *28*(5), 675-91; discussion 691-735. DOI: 10.1017/S0140525X05000129

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*(3), 705–722. DOI: 10.1111/j.1467-8624.2007.01025.x

Topál, J., Gergely, G., Erdőhegyi, Á., Csibra, G., & Miklósi, Á. (2009). Differential sensitivity to human communication in dogs, wolves, and human infants. *Science*, *325*(5945), 1269–1272. DOI: 10.1126/science.1176960

Topál, J., Gergely, G., Miklósi, Á., Erdőhegyi, Á., & Csibra, G. (2008). Infants' perseverative search errors are induced by pragmatic misinterpretation. *Science*, *321*(5897), 1831–1834. DOI: 10.1126/science.1161437

Urakawa, S., Takamoto, K., Ishikawa, A., Ono, T., & Nishijo, H. (2015). Selective medial prefrontal cortex responses during live mutual gaze interactions in human infants: An fNIRS study. *Brain Topography*, *28*(5), 691–701. DOI: 10.1007/s10548-014-0414-2

Veà, & Sabater-Pi. (1998). Spontaneous pointing behaviour in the wild pygmy chimpanzee (*Pan paniscus*). *Folia primatologica*, *69*(5), 289–290. DOI: 10.1159/000021640

Verschoor, S. A., Paulus, M., Spapé, M. M., Bíró, S., & Hommel, B. (2015). The developing cognitive substrate of sequential action control in 9- to 12-month-olds: Evidence for concurrent activation models. *Cognition*, *138*, 64–78. DOI: 10.1016/j.cognition.2015.01.005

Verschoor, S. A., Spapé, M. M., Bíró, S., & Hommel, B. (2013). From outcome prediction to action selection: Developmental change in the role of action-effect bindings. *Developmental Science*, *16*(6), 801–814. DOI: 10.1111/desc.12085

Vygotsky, L. (1981). The genesis of higher mental functions. In J. V. Wertsch (Ed.), *The concept of activity in soviet psychology* (pp. 144–188). Armonk, N.Y.: M.E. Sharpe.

de Waal, F. (1982). *Chimpanzee politics: Sex and power among apes.* London, England: Cape.

Wang, C.-A., Boehnke, S. E., White, B. J., & Munoz, D. P. (2012). Microstimulation of the monkey superior colliculus induces pupil dilation without evoking saccades. *Journal of Neuroscience*, *32*(11), 3629–3636. DOI: 10.1523/JNEUROSCI.5512-11.2012

Wang, C.-A., & Munoz, D. P. (2015). A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology*, *33*, 134–140. DOI: 10.1016/j.conb.2015.03.018

Wang, S.-H., Baillargeon, R., & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition*, *93*(3), 167–198. DOI: 10.1016/j.cognition.2003.09.012

Wellman, H. M., & Miller, K. F. (1986). Thinking about nothing: Development of concepts of zero. *British Journal of Developmental Psychology*, *4*(1), 31–42. DOI: 10.1111/j.2044-835X.1986.tb00995.x

Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and the expression of thought.* New York, NY: Wiley.

Wetzel, N., Buttelmann, D., Schieler, A., & Widmann, A. (2016). Infant and adult pupil dilation in response to unexpected sounds. *Developmental Psychobiology*, *58*(3), 382–392. DOI: 10.1002/dev.21377

Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition*, *72*(2), 125–166. DOI: 10.1016/S0010-0277(99)00035-9

Wilcox, T., & Baillargeon, R. (1998a). Object individuation in infancy: The use of featural information in reasoning about occlusion events. *Cognitive Psychology*, *37*(2), 97–155. DOI: 10.1006/cogp.1998.0690

Wilcox, T., & Baillargeon, R. (1998b). Object individuation in young infants: Further evidence with an event-monitoring paradigm. *Developmental Science*, *1*(1), 127–142. DOI: 10.1111/1467-7687.00019

Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., & Boas, D. A. (2005). Using near-infrared spectroscopy to assess neural activation during object processing in infants. *Journal of Biomedical Optics*, *10*(1), 11010. DOI: 10.1117/1.1852551

Willatts, P. (1984). Stages in the development of intentional search by young infants. *Developmental Psychology*, *20*(3), 389–396. DOI: 10.1037/0012-1649.20.3.389

Wishart, J. G., Bower, T. G. R., & Dunkeld, J. (1978). Reaching in the dark. *Perception*, *7*(5), 507–512. DOI: 10.1068/p070507

Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, *17*(1), 1061–1084. DOI: 10.1016/S0885-2014(02)00074-6

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749–750. DOI: 10.1038/358749a0

Wynn, K., & Chiang, W.-C. (1998). Limits to infants' knowledge of objects: The case of magical appearance. *Psychological Science*, *9*(6), 448–455. DOI: 10.1111/1467-9280.00084

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15-B25. DOI: 10.1016/S0010-0277(03)00050-7

Xu, F. (2005). Categories, kinds, and object individuation in infancy. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time.* Mahwah, NJ: L. Erlbaum Associates.

Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, *30*(2), 111–153. DOI: 10.1006/cogp.1996.0005

Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I., & Deouell, L. Y. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron*, *58*(3), 429–441. DOI: 10.1016/j.neuron.2008.03.027

Zentall, T. R., & Pattison, K. F. (2016). Now you see it, now you don't: Object permanence in dogs. *Current Directions in Psychological Science*, *25*(5), 357–362. DOI: 10.1177/0963721416664861

Zimmermann, F., Zemke, F., Call, J., & Gómez, J. C. (2009). Orangutans (*Pongo pygmaeus*) and bonobos (*Pan paniscus*) point to inform a human about the location of a tool. *Animal Cognition*, *12*(2), 347–358. DOI: 10.1007/s10071-008-0194-8

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Fakultät für
Psychologie und
Bewegungswissenschaft

Institut für Bewegungswissenschaft
Institut für Psychologie

Erklärung gemäß *(bitte Zutreffendes ankreuzen)*

☐ **§ 4 (1c) der Promotionsordnung**

des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010

☐ **§ 5 (4d) der Promotionsordnung**

des Instituts für Psychologie der Universität Hamburg vom 20.08.2003

Hiermit erkläre ich,

______________________________ (Vorname, Nachname),

dass ich mich an einer anderen Universität oder Fakultät noch keiner Doktorprüfung unterzogen oder mich um Zulassung zu einer Doktorprüfung bemüht habe.

______________________ Ort, Datum

______________________ Unterschrift

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Fakultät für
Psychologie und
Bewegungswissenschaft

Institut für Bewegungswissenschaft
Institut für Psychologie

Eidesstattliche Erklärung nach ***(bitte Zutreffendes ankreuzen)***

☐ **§ 7 (4) der Promotionsordnung des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010**

☐ **§ 9 (1c und 1d) der Promotionsordnung des Instituts für Psychologie der Universität Hamburg vom 20.08.2003**

Hiermit erkläre ich an Eides statt,

1. dass die von mir vorgelegte Dissertation nicht Gegenstand eines anderen Prüfungsverfahrens gewesen oder in einem solchen Verfahren als ungenügend beurteilt worden ist.
2. dass ich die von mir vorgelegte Dissertation selbst verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die wörtlich oder inhaltlich übernommenen Stellen habe ich als solche kenntlich gemacht.

______________________________ ______________________________

Ort, Datum Unterschrift